IMPROVING TEXT CATEGORIZATION PERFORMANCE BY COMBINING FEATURE SELECTION METHODS

by

Ece Özbilen

B.S., Computer Engineering, Beykent University, 2008

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Computer Engineering

Boğaziçi University

2011

IMPROVING TEXT CATEGORIZATION PERFORMANCE BY COMBINING FEATURE SELECTION METHODS

APPROVED BY:

Assoc. Prof. Tunga Güngör . . . .
(Thesis Supervisor)

Prof. Fikret Gürgen . . . .

Prof. Bülent Sankur . . . .

DATE OF APPROVAL:

ACKNOWLEDGEMENTS

I would like to express my special gratitude to my thesis advisor, Assoc. Prof. Tunga Güngör, for his mentorship, guidance and support throughout the whole process of this thesis.

I would like to thank to Prof. Fikret Gürgen and Prof. Bülent Sankur for their participation to my thesis jury among their heavy program and also for their valuable comments.

I also would like to thank to TÜBİTAK (Türkiye Bilimsel ve Teknolojik Araştırma Kurumu) for their financial support (BIDEB-2210 Fellowship) during my master study.

The last but not the least, I would like to thank my family for their unconditional love and unwavering support.

ABSTRACT

IMPROVING TEXT CATEGORIZATION PERFORMANCE BY COMBINING FEATURE SELECTION METHODS

Even though the arrival of the machine learning methods in text categorization is one of the essential factors that improves the effectiveness of text categorization, high dimensionality is still a challenge for classification performance. There are several ways to reduce the dimension of input vector in classification and feature selection is one of the most popular and effective methods of reducing dimension. Various researches have been done to improve the performance of feature selection methods on text categorization but they mostly deal with how to advance the performance of the individual feature selection methods whereas we know that combining the outputs of multiple algorithms/classifiers is one of the promising strategies that has been studied extensively in information retrieval.

With this motivation, we present a comprehensive analysis of the comparison between the feature selection methods and their varied binary combinations for text categorization with a comparative discussion. We analyze the performances of five common feature selection methods with their combinations on five standard datasets with varied skewness in both global and local policies by using SVM. Comparing the performance of the individual methods with the performance of the combination methods shows that combining two feature selection methods significantly improves the performance of the individual methods. In addition, rank combination achieves better performance in the case of global policy on the other hand score combination significantly achieves better performance in the case of local policy.

In this thesis, the main concern is to investigate the effectiveness of combining the individual metrics on the performances of text categorization. Thus, we also propose new combination methods that some of them clearly outperform the success of the score and rank combinations.

ÖZET

ÖZNİTELİK SEÇME METODLARINI BİRLEŞTİREREK METİN SINIFLANDIRMA PERFORMANSININ İYİLEŞTİRİLMESİ

Makine öğrenmesi yöntemlerinin metin sınıflandırmada kullanılmaya başlanması, sınıflandırma performansını arttıran önemli bir faktör olmasına rağmen yüksek boyutluluk sınıflandırma başarısı için hala önemli bir problem. Sınıflandırmada doküman vektörlerinin boyutunu azaltmak için birçok yöntem önerilmektedir. Öznitelik seçme yöntemi de boyut azaltmada kullanılan en yaygın ve etkili yöntemlerden biridir. Öznitelik seçme metodlarının sınıflandırmadaki performansını arttırmak için birçok araştırma yapılmış ve yapılıyor olmasına rağmen, incelenen öznitelik seçme metodlarının bir arada kullanılması ile ilgili araştırmalar dokuman sınıflandırma alanında çok kısıtlı.

Farklı yöntemleri birleştirerek bilgi erişim alanında başarılı sonuçlar elde edilmesi, bizi bu çalışmada öznitelik seçme metodlarını birleştirerek metinleri sınıflandırmaya yöneltti. Bu amaçla, bu çalışmada özellik seçme yöntemlerinin ve bu yöntemlerin çeşitli ikili birleşimlerinin karşılaştırılmasına yönelik kapsamlı bir araştırma sunuyoruz. Beş farklı öznitelik seçme metodu ve birleşimlerini farklı özellikteki beş veri kümesi üzerinde yerel ve genel politika kapsamında SVM sınıflandırıcısı ile analiz edildi. Analiz sonucunda, birleştirilen öznitelik seçme metodlarının metodların tek kullanılmasına göre daha başarılı sonuçlar elde ettiğini gördük. Özellikle yöntemlerin skor değerlerini birleştirmek yerel politikada belirgin şekilde başarılıyı arttırırken, sıra değerlerini birleştirmek genel politikada daha başarılı sonuçlar elde edilmesini sağladı.

Bu tezde amacımız öznitelik seçme metodlarını birleştirmenin metin sınıflandırma performansındaki başarısını incelemek ve karşılaştırmaktır. Bu kapsamda skor ve sıra birleştirme yöntemlerinin yanında yeni birleştirme yöntemleri de tezde önerildi ve incelendi. Çalışma sonucunda önerilen bazı yöntemlerin skor ve sıra birleştirme yöntemlerinin başarısını da geliştirdiği gözlemlendi.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

c	Category
c_i	Category i
C_i	Proposed combination method i
C_{rank}	Rank Combination
C_{score}	Score Combination
d	Document vector
df	Document frequency
d_j	Document j
f_{max}	Function that selects the maximum of the inputs
(g)	Global policy
(l)	Local policy
M	Total number of feature selection method
n	Total number of documents
n_i	Number of documents term i appears in
$Rank_i$	Rank Vector of the feature selection method i
r_i	Rank of the term i
$Score_i$	Score Vector of the feature selection method i
$Score_i'$	Normalized Score Vector of the feature selection method i
s_i	Score of the term i
tf	Term frequency
tf_i	Frequency of term i
tf_{ij}	Frequency of term i in document j
t_k	Term k
w_i	Weight of term i
w_{ij}	Weight of term i in document j
x^2	CHI-square distribution
π	Precision
ρ	Recall

$P(\overline{t_k}, c_i)$	Percentage of documents belonging to class c_i in which term t_k does not occur
$P(t_k\ \overline{c_i})$	Percentage of documents not belonging to class c_i in which term t_k occurs
$P(\overline{c_i})$	Percentage of documents not belonging to class c_i
$P(\overline{t_k}, \overline{c_i})$	Percentage of documents not belonging to class c_i in which term t_k does not occur
$P(\overline{t_k})$	Percentage of documents in which term t_k does not occur
$P(c_i)$	Percentage of documents belonging to class c_i
$P(t_k, c_i)$	Percentage of documents belonging to class c_i in which term t_k occurs
$P(t_k)$	Percentage of documents in which term t_k occurs
Acc	Accuracy
Acc2	Accuracy2
BNS	Bi-Normal Separation
CFA	Combinatorial Fusion Analysis
CHI	Chi-square Statistics
DF	Document Frequency
FN	False Negatives
FP	False Positives
GR	Gain Ratio
HTML	Hyper Text Markup Language
IDF	Inverse Document Frequency
IG	Information Gain
IR	Information Retrieval
k-NN	*k*-nearest neighbor
LLSF	Linear Least Squares Fit
MI	Mutual Information
ModApte	Modified Apte
NB	Naive Bayes
NNet	Neural network

Odds	Odds Ratio
RCV1-v2	Reuters Corpus Volume 1 Version 2
SMART	System for the Mechanical Analysis and Retrieval of Text
s-test	Micro Sign Test
STW	Supervised Term Weighting
SVM	Support Vector Machines
tf-idf	Term frequency-inverse document frequency
TN	True Negatives
TP	True Positives
TREC	Text Retrieval Conference
TS	Term Strength

1. INTRODUCTION

Since the early 1990s, the accessibility and abundance of digital documents make text categorization an important and necessary research field. Today, text categorization is being applied in many contexts in order to organize and manipulate the documents. Arrival of the machine learning methods in text categorization is one of the essential factors that improves the effectiveness of text categorization in information retrieval systems.

Text categorization is a supervised learning task that assigns the predefined category labels to new documents based on the likelihood derived from a set of labeled training documents. In order to classify documents, each document should be transformed into a model that preserves as much of the original information as possible. The *bag of words* representation is one of the simple and preferred models that represents a document as a set of distinct words by ignoring the order and meaning of words. When the number of words in documents is considered, high dimensionality may become an inevitable problem. Since the data in text categorization are high-dimensional, naturally dimensionality reduction becomes a necessity for efficiency and accuracy.

Feature selection is one of the well-known processes that reduces the dimensionality by ranking all features according to their importance estimated by a method and then selecting ones with the highest values. Feature selection not only reduces time and storage requirements but also improves the efficiency and accuracy of the classifiers. Feature selection makes applying classifiers on data more efficient by reducing the size of the effective features. In addition, feature selection often improves classification accuracy by eliminating noise features that are non-informative or misleading for classification and lead to incorrect generalization, *overfitting,* from the training documents.

In text categorization there are two main policies to apply feature selection: local policy and global policy. The local policy, where a different set of features is selected from each class independent from other classes, gives equal weight to each class. Thus, it tends to optimize the classification performance on frequent and infrequent classes by selecting the most important features for each class. On the other hand, the global policy, where a single set of features is selected from all classes, provides a global view of the entire dataset by extracting a single global score from the local scores. Thus, the global policy tends to penalize the infrequent classes in highly skew datasets by selecting the most important features for the entire dataset

One of the main drawbacks in text categorization is imbalance data distribution since the rare classes are dominated by common classes. As the performances of the classifiers are directly affected by the skewness of the datasets, the performance is commonly measured by two different alternatives: micro-averaged and macro-averaged F-measures. First gives equal weight to each document and therefore it tends to be dominated by the classifier's performance on common categories while reflects the overall accuracy better. On the other hand, second gives equal weight to each category regardless of its frequency and thus it is influenced more by the classifier's performance on rare categories.

1.1. Related Works

The arrival of the machine learning methods in the text categorization field is one of the most important factors that accelerates the improvement in this field by strong theoretical motivations. A growing number of machine learning methods have been used for text categorization, including probabilistic classifiers, decision trees, regression methods, nearest neighbor classifiers, Rocchio method, neural networks, example-based classifiers, etc. [1].

In 1995 new machine learning method Support Vector Machines (SVMs) were introduced by Vapnik [40, 42]. In later years, many studies have explored the use of SVMs for text categorization with promising results [3, 14, 15, 18, 31, 32, 37, 39]. One of the most basic studies that introduces SVMs for text categorization is presented by Joachims in 1998. In the study the performance of SVM using non-linear model is compared with four popular machine learning algorithms (Naïve Bayes (NB) classifier, Rocchio method, *k*-nearest neighbor (*k*-NN) classifier and C4.5 decision tree) on Reuters and Ohsumed datasets. The analysis concludes that SVM is very well suited for test categorization and significantly outperforms other methods. In the same year, Dumais et al. compare the effectiveness of five different machine learning algorithms (a variant of Rocchio's method, decision trees, NB, Bayes Nets and SVM) on Reuters dataset for text categorization and conclude that the accuracy of their simple linear SVM is among the best reported for the Reuters similar to Joachims study but the linear SVM is especially promising because it is a much simpler and more efficient than Joachims' non-linear model [14]. In 1999 a controlled study with statistical significant tests on five machine learning algorithms (SVM, *k*-NN, neural network (NNet) method, NB and Linear Least-Square Fit (LLSF) mapping) is conducted by Yang and Liu. Their results show that SVM is one of the most successful machine learning algorithms. Sebastiani presents a survey that covers the main machine learning approaches in text categorization [15].

Reducing dimensionality is another critical issue in text categorization. Feature selection is one of the effective methods that improves the efficiency and accuracy of the classifiers by selecting only more discriminative terms in a dataset as features. In the literature, various feature selection methods have been presented and analyzed [16].

Yang and Petersen, 1997, compare five feature selection methods (information gain (IG), document frequency (DF), term strength (TS), mutual information (MI) and x^2-test (CHI)) on the Reuters and Ohsumed datasets by using *k*-NN and LLSF classification algorithms in the case of global policy. They define that IG and CHI are the most successful methods [2]. In 2003 Forman present an empirical comparison of twelve feature selection methods (CHI, IG, Odds Ratio (Odds), DF, Accuracy (Acc), Acc2, Bi-Normal Separation (BNS), etc.) on a benchmark that are gathered from Reuters, TREC and Ohsumed, etc. by using SVM in the case of local policy. This study reports that the proposed method BNS shows outstanding performance for accuracy and F-measure, especially in highly skew datasets but still IG yields the best results for precision [3]. In the same year, Debole and Sebastiani propose supervised term weighting (STW) scheme using three feature selection scores (IG, CHI and gain ratio (GR)) with tf-idf weighting on Reuters dataset with SVM in both local and global policies. They conclude that GR outperforms other methods and has given very good results as a STW function, especially in macro-averaged F-measure [4].

Özgür et al., 2005, compare tf-idf weighting with boolean weighting on Reuters dataset with SVM again in both local and global policies. They find out that tf-idf weighting performs better than boolean weighting and the global policy performs better for large number of keywords but for small number of keywords, local policy outperforms the global policy [5]. In a later study, Özgür and Güngör continue to examine the performance of these two policies with these two weightings on six standard document collections (Classic3, Hitech, LA1, Reviews, Reuters and Wap) with different skewness properties and with different number of keywords by using SVM in more detail in 2006. In addition, they prove that the results of the previous study can be generalize by the global policy performs better for large number of keywords while local policy performs better for small number of keywords and in skew datasets [6]. Taşçı and Güngör, 2006, expend the analysis with various existing feature selection methods (tf-idf, IG, CHI, Acc2 and DF) and four new proposed feature selection methods which resemble the Acc2 metric. They have compared these feature selection methods on six standard document collections by varying number of selected features from 10 to 2000 in both local and global policy. Furthermore, they conclude that Acc2 is the best metric among the existing metrics, especially with a few number of features. The success of Acc2 becomes clear in local policy on skew datasets. On the other hand, the proposed metric M_1 is more successful than the best existing metric Acc2 in the experiments [7, 8].

Liu et al., 2007, focus on the data imbalance problem in text categorization by presenting a probability based term weighing scheme inspired by different feature selection methods. They conclude that using the probability based term weighting scheme can improve classification performance on rare classes [17].

1.2. Motivation

As seen, various researches have been done to improve the performance of feature selection methods on text categorization but they mostly deal with how to improve the performance of the individual feature selection methods whereas we know that combining the outputs of multiple classifiers is one of the promising strategies that has been studied extensively in information retrieval area. Combining the outputs of classifiers or search systems, most of time, demonstrates better performance than a single algorithm [9, 12, 13, 28, 29, 49, 50].

Lee, 1995, proposes that significant improvement in retrieval performance can be achieved by combining multiple weighting schemes and concludes that combining rankings based on cosine normalization of the tf-idf weight with other normalization schemes is clearly effective [49]. Two years later, 1997, Lee, combines the results of the six selected retrieval systems and finds that combining the outputs of search systems using ranks of the documents gives better results than the scores of the documents if there is a different rank-score curve [50].

Furthermore, although many feature selection methods exist in text categorization, it is hard to state one is in general superior to others since the success of the methods depends on various variables. It is more likely that combining different feature selection methods obtains more effective performance in text categorization. From this point of view, we decide to apply the idea of combining feature selection methods in text categorization. When we investigate the literature, it has been seen that there are a few studies that focus on this strategy in recent years. The papers, are given below in detail, study the use of combining feature selection methods in text categorization.

In the first study, Olsson and Oard, 2006, combine multiple feature selection methods (DF, IG, CHI, CHI_{max} and CHI_{avg}) by using the scores themselves or the rank ordering of the features. In order to combine two or more feature selection metrics, they introduce three combination methods: highest rank (HR), lowest rank (LR) and average rank (AR) combination. They compare individual and combination methods by using *k*-NN with symmetric Okapi term weighting on 23,149 documents from RCV1-v2 dataset in local policy. The experiments show that CHI based individual feature selection methods and combinations of two CHI based feature selection methods show the best performance. More importantly, they conclude that even the result of a simple combination method is better than the results of the individual metrics [10].

In the second study, Li et al. ,2009, analyze the score and rank combinations of five feature selection methods (DF, IG, CHI, MI and TS) by using the Combinatorial Fusion Analysis (CFA). CFA states when and how feature selection methods can be combined to achieve better performance by using rank-score function and rank-score graph. It is seem that the combination of two feature selection methods is more likely to have a better result when these two methods also show a good performance individually and their rank-score graphs are quite different. They evaluate these five feature selection methods by performing the NB classifier on two datasets. The first dataset, denoted by REC, is extracted from the Newsgroups data set. The second dataset, denoted by TOPIC, is extracted from the TOPICS category sets of the Reuters-21578. For REC dataset, the best 2-combinations are: MI & IG and MI & DF rank combinations which are higher than the best individual metric MI and the best 3-combination is: MI & DF & IG rank combination. For TOPIC dataset, the best 2-combination is: TS & DF rank combination which is higher than the best individual feature selection method TS and the best 3-combination is: TS & DF & CHI which is better than each of TS, DF and CHI. The results show that combining multiple feature selection methods can improve the performance of text categorization in terms of the average F-measure results. The performance of the rank combination is significantly better than the score combination. In addition, as the number of the feature selection method in the combination increases, the performance of the combination decreases. The best performances are achieved by the combination of two methods [11].

However these previous two studies compare the performance of the individual and combination methods, the methods are analyzed on data which are constructed by selecting documents from the standard datasets. The second study prefers to test the methods on homogenous two datasets include, respectively, only 4 and 5 classes. Thus it is hard to compare these results with other studies. Secondly, both studies prefer to share their experimental results by demonstrating them in figures but it is a necessity to see the experimental results in more detail with numeric values in order to understand the results. The second study gives the results of the average and maximum F-measures but only compares the results based on the average F-measure values. There is not any information about the success of the combination methods in terms of maximum F-measure values. Finally the first study classifies documents with k-NN and the second uses Naïve Bayes classifier whereas we know that SVM outperforms these two machine learning algorithms in many studies in the literature. Although these limitations should be noted, the experimental results demonstrate that combining feature selection methods improves the effectiveness of text categorization in both studies. This promising result motivates us to perform the combination of the feature selection methods in a more comprehensive and comparative manner.

In this study, the main concern is to investigate the effectiveness of combining the individual metrics on the performances of text categorization. Firstly we analyze the performance of five common feature selection methods on five standard datasets with varied skewness in both global and local policies, and then evaluate the performance of all possible binary score and rank combinations of these five feature selection methods with the same experimental settings in order to determine the most appropriate features for classification, and finally compare the performance of the individual methods with the performance of the combination methods. The results of the experiments show that combining two feature selection methods can improve the performance of the individual metrics. In addition, rank combination achieves better performance in the case of global policy and score combination significantly achieves better performance in the case of local policy.

When we see that the combination of different feature selection methods can improve the performance of text categorization, secondly, we propose varied combination methods derived from score and rank values of the terms. These experiments also show that some of them also outperform the success of the score and rank combinations.

1.3. Thesis Organization

The rest of the thesis is organized as follows: Section 2 describes our document presentation and preprocessing steps. Section 3 overviews the five individual feature selection methods that are used in this study together with local and global policies. Section 4 describes how the outputs of the different feature selection methods can be combined, introduces score and rank combinations and then proposes new seven methods for combining different feature selection methods. Section 5 describes the system architecture step by step that designed for this study. In Section 6 we describe our experimental settings; the classifier, the datasets, evaluation metrics. In Section 7 we show through experimental results that significant improvement can be obtained by combining two feature selection methods with detailed analysis. Finally, Section 8 concludes the thesis and gives the future research.

2. DOCUMENT PREPROCESSING AND REPRESENTATION

Document representation is the process of transforming the unstructured text into a structured data as a vector in order to classify the text documents by applying machine learning techniques. The most widely used method for document representation is the "*vector space model*" introduced by Salton and associates in 1975 [18]. In vector space model, each document is represented as a vector *d* and each dimension in the vector corresponds to distinct term in the term space of the document collection [5].

In this study, we use "*bag of words*" in vector space model which define each term as a distinct single word. The *bag of words* representation is a very simple and preferred method that makes the representation and learning highly efficient and easy by ignoring the order and meaning of distinct words [20]. Although it is a simple method, high dimensionality becomes an important issue when terms are defined as single words in the feature space. In order to reduce the high dimensionality, we apply some preprocessing methods which described by the following tasks:

2.1. Parsing the Documents and Case-folding

In the first step, all the HTML mark-up tags and non-alphabetic characters such as numerals, special characters and date are removed from the documents in the dataset. Then case-folding is applied to convert all characters into same case "*lower case*" in order to avoid the duplication of the same words.

2.2. Removing Stopwords

Overly common words, such as pronouns, prepositions and conjunctions in English, like 'it', 'in' and 'and', occur so frequently that they cannot give any useful information about the content and be discriminatory for a specific class. These words are so called *"stopwords"*. We use the stopword list was built by Salton and Buckley for the SMART system at Cornell University to eliminate common words. The list consists of 571 words is given in Appendix A [21].

2.3. Stemming

Removing stopwords causes an efficient reduction in the dimensionality of the feature space but we also need stemming word to reduce the dimensionality of the feature space to a reasonable number. *Stemming* is a preprocessing for finding the root morphemes of the words. In order to stem the words, we use Porter's Stemmer which is the most widely used algorithm for word stemming in English. Porter's Stemming Algorithm is a process for removing the commoner morphological and inflexional affixes from words [22, 23]. In other words, it is based on only morphological issues that are completely independent from the syntactic and semantic structure of the sentence. For example, the words "computer", "computers", "computing" and "computes" are stemmed the same root "comput". After stemming, terms that left a single character are also removed since they cannot give any information about the content of a document.

2.4. Term Weighting

As already mentioned at the beginning of this section we represent each document as a vector *d*

$$\mathrm{d} = (w_1, w_2, \ldots, w_n) \tag{2.1}$$

where w_i is the weight of term *i* of document d. There are several various ways to compute these term weights [24]. There are three main assumptions that are valid for all computations [25]:

1. Rare terms are no less important than frequent terms,
2. Multiple appearances of a term in a document are no less important than single appearances,
3. Long documents are no more important than short documents.

The *term frequency-inverse document frequency (tf-idf) weighting* is one of the widely used weighting methods that takes into account these properties. *df* formula meets the first assumption, *tf* formula meets the second assumption and length-normalization meets the third assumption which given above. Thus we apply *tf-idf weighting* method in this study whose formula is given below:

$$w_{ij} = tf_{ij} \,.\log(\frac{N}{df_{ij}}) \tag{2.2}$$

where w_{ij} is the weight of a term i in document j, tf_{ij} denotes the frequency of the term *i* in document j, df_{ij} denotes *the number of documents* in which a term *i* occurs in the whole document and *N* is the total number of documents.

The *tf-idf* weighting considers that if a term more often occurs in a document, it is more discriminative whereas if it appears in most of the documents, then it is less discriminative for the content.

3. FEATURE SELECTION

Text categorization is a supervised learning task that assigns the predefined category labels to new documents based on the likelihood derived from a set of labeled training documents. In order to classify documents, each document should be transformed into a model that preserves as much of the original information as possible. The *bag of words* representation is one of the simple and preferred models that represents a document as a set of distinct words by ignoring the order and meaning of words. When the number of words in documents is considered, high dimensionality may become an inevitable problem. Since the data in text categorization are high-dimensional, naturally dimensionality reduction becomes a necessity for efficiency and accuracy.

Feature selection is one of the well-known processes that reduces the dimensionality by ranking all features according to their importance estimated by a metric and then selecting ones with the highest values. Feature selection not only reduces time and storage requirements but also improves the efficiency and accuracy of the classifiers. Feature selection makes applying classifiers on data more efficient by reducing the size of the effective features. In addition, feature selection often improves accuracy of the classification by eliminating noise features that are non-informative and misleading for classification and lead to incorrect generalization *(overfitting)* from the training set.

3.1. Local and Global Policy

In text categorization there are two main policies to apply feature selection: local policy and global policy. In the first policy, a different set of features is selected from each category. In the second policy, a single set of features is selected from all categories.

The local policy, where a different set of features is selected from each class independent from other classes, gives equal weight to each class. Thus, the local policy tends to optimize the classification performance on frequent and infrequent classes by selecting the most important features for each class. On the other hand, the global policy, where a single set of features is selected from all classes, provides a global view of the entire dataset by extracting a single global score from the local scores. Thus, the global policy tends to penalize the infrequent classes in highly skew datasets by selecting the most important features for the entire dataset [4, 26].

There are several ways to obtain global score from the local scores: *maximization, averaging, weighted averaging* and *weighted maximum* are the most popular globalization techniques. *Maximization, averaging* and *weighted averaging* were presented by Yang and Pedersen in 1997 and *weighted maximum* was proposed by Calvo and Ceccatto in 2000 [2, 27]. We selected *maximization,* which formulation given below, as a globalization technique, since it consistently outperformed other globalization techniques in the study of Debole and Sebastiani. In their paper, the success of the *maximization* was explained that it prefers to select terms that are good separator even on a single category rather than terms that are only fair separators on many categories.

$$f_{max}\ (t_k) = \ max_{i=1}^{|C|}\{\ f(t_k, c_i)\} \tag{3.1}$$

3.2. Feature Selection Metrics

In this study, first of all the five widely used feature selection metrics: term frequency-inverse document frequency *(tf-idf),* chi-square statistics *(CHI),* information gain *(IG), Accuracy2 (Acc2)* and document frequency thresholding *(DF)* are analyzed in order to compare with combination methods. The following section describes these metrics appearing in the literature.

3.2.1. Term Frequency-inverse Document Frequency (tf-idf)

The idea of the tf-idf feature selection is to selects the words with the highest tf-idf scores. This method gives the highest scores to the terms that appear in a few documents with a high frequency. In other words, if a term more often occurs in a document, this means it is more discriminative whereas if it appears in most of the documents, then it is less discriminative for the content.

The formulation of the tf-idf feature selection method is obtained from the tf-idf weighting described in Section 2.4 in detail:

$$w_{ij} = tf_{ij} \,.\log(\frac{N}{df_{ij}}) \tag{3.2}$$

$$tf - idf(w_i) = \sum_{j=1}^{N} w_{ij} \tag{3.3}$$

3.2.2. Chi-square Statistics (CHI)

In experimental sciences, chi-square statistics is frequently used to measure how the observation results differ from the expected results. In other words, it measures the independence of two random variables.

$$CHI = \sum_{ij} \frac{(Observed_{ij} - Expected_{ij})^2}{Expected_{ij}} \tag{3.4}$$

Chi-square statistics is also widely used in text categorization [4, 3, 7, 11]. In text categorization, the two random variables are occurrence of term t_k and occurrence of class c_i and chi-square statistics measures the independence between t_k and c_i. The formula for chi-square score is:

$$CHI\,(t_k, c_i) = N\;x\;\frac{[P(t_k, c_i)P(\overline{t_k}, \overline{c_i}) - P(\overline{t_k}, c_i)P(t_k\,\overline{c_i})]^2}{P(t_k)P(\overline{t_k})P(c_i)P(\overline{c_i})} \tag{3.5}$$

where $P(t_k)$ is the percentage of documents in which term t_k occurs, $P(\overline{t_k})$ is the percentage of documents in which term t_k does not occur, $P(c_i)$ is the percentage of documents belonging to class c_i, $P(\overline{c_i})$ is the percentage of documents not belonging to class c_i, $P(t_k, c_i)$ is the percentage of documents belonging to class c_i in which term t_k occurs, $P(\overline{t_k}, \overline{c_i})$ is the percentage of documents not belonging to class c_i in which term t_k does not occur, $P(\overline{t_k}, c_i)$ is the percentage of documents belonging to class c_i in which term t_k does not occur and $P(t_k\,\overline{c_i})$ is the percentage of documents not belonging to class c_i in which term t_k occurs.

If chi-square score of a term t_k is low value, this means t_k is independent from the class c_i and if chi-square score of a term t_k is high value, this means t_k is dependent of the class c_i . Thus the chi-square feature selection method selects the terms with the highest chi-square score which are more informative for classification.

3.2.3. Information Gain (IG)

Another popular feature selection method in text categorization is information gain [4, 3, 7, 11]. It is measure the decrease in entropy by existence or absence of the term in a document. Information gain score will be null for two independent variables and it will be high because of the dependence between two variables. The information gain score of a term t_k is calculated by the formula:

$$IG(t_k, c_i) = \sum_{C \in \{c_i, \overline{c_i}\}} \sum_{t \in \{t_i, \overline{t_i}\}} P(t, c) \,.\log\left(\frac{P(t, c)}{P(t).P(c)}\right) \tag{3.6}$$

Information gain feature selection method selects the terms with the highest information gain scores which contains much information about the classes.

3.2.4. Accuracy2 (Acc2)

The final method Accuracy2 has showed a promising performance compare to other feature selection metrics in the previous studies [3, 7]. In this method, only the number of documents in which the term occurs is taken into account without considering the number of actual documents. It measures the difference of the distributions of a term in the documents belonging to a class and in the documents not belonging to that class. Thus, the term t_k that never occurs in a class c_i can be selected as a feature for c_i . Below is the formula for calculation of accuracy2 score:

$$Acc2\ (t_k, c_i) = |P(t_k, c_i) - P(t_k\ \overline{c_i})| \tag{3.7}$$

3.2.5. Document Frequency (DF)

Document frequency is a very simple and popular method that measures the number of documents in which the term appears without class labels [3, 7, 11]. Purpose of the method is to eliminate the rare words which are assumed non-informative and misleading for classification. Document frequency feature selection method selects the terms with the scores. Its formula is:

$$DF(t_k, c_i) = P(t_k, c_i) \tag{3.8}$$

4. COMBINATION OF FEATURE SELECTION METHODS

This section discusses methods for combining the outputs of the individual feature selections metrics. Various researches have been done to improve the performance of feature selection methods on text categorization but they mostly deal with how to improve the performance of the individual feature selection methods whereas we know that combining the outputs of classifiers one of the promising strategy that has been studied extensively in information retrieval area.

Furthermore, although many feature selection methods exist in text categorization, it is hard to state one is in general superior to others since the success of the methods depends on various variables. It is more likely that combining different feature selection methods obtains more effective performance in text categorization.

From this point of view, we apply some combination strategies to feature selection methods described in Section 3. There are many ways of combining the outputs of the various methods. The most popular ways to combine corresponding outputs are linear combining methods such as averaging and weighted averaging and non-linear combining methods such as ranking and voting. Tumer and Ghosh, 1999, state that simply averaging among them is the most common combining strategy in information retrieval [12]. More importantly, Fox and Shaw, 1994, conclude that best combining strategy is based on summing the outputs of the algorithms which is equivalent to averaging when compared six combining strategies [28]. Beside this conclusion, Hull et al., 1996, find that the performance of the simple averaging strategy outperforms the complex combinations of the classifiers [29]. Considering these results, we decide to apply the averaging strategy in two ways: score combination and rank combination.

4.1. Score and Rank Combinations of Feature Selection Methods

The previous studies show that the success of the combination and the number of feature selection method involved in combination are inversely related. As the number of the feature selection method in the combination increases, the performance of the combination decreases. In addition, it states that the best performances are achieved by the combination of two feature selection methods [10, 11]. Thus we only consider combining two distinct feature selection methods. In the study, we evaluate the performance of all possible binary-combinations (2-combinations) of five feature selection methods: *tf-idf, CHI, IG, Acc2* and *DF* metrics which are *tf-idf & CHI, tf-idf & IG, tf-idf & Acc2, tf-idf & DF, CHI & IG, CHI & Acc2, CHI & DF, IG & Acc2, IG & DF* and *Acc2& DF.*

The principle of the feature selection methods can separate into two steps. First is scoring that is giving higher scores to the terms which considered more informative for classification and second is selection that is selecting the terms with the highest scores. In order to combine the outputs of varied feature selection methods in scoring step, scores of each term from the varied feature selection methods are normalized using the maximum and minimum scores according to the below formula:

$Score = (s_1, s_2, ..., s_n)$ where s_i score of the i^{th} term, n is the total number of term.

$$score' = \frac{Score - min(Score)}{max(Score) - min(Score)} \tag{4.1}$$

By normalization, the scores fall in the same range [0, 1] and scores of the terms from the varied feature selection methods are represented equally which is allowing for meaningful comparisons between the methods.

4.1.1. Score Combination

Score combination is averaging the normalized term scores of the different feature selection methods.

$$C_{score} = \sum_{i=1}^{M} \frac{Score_i'}{M} \tag{4.2}$$

where M is the number of feature selection methods which is 2 for this study.

4.1.2. Rank Combination

Rank Combination is averaging the term ranks of the different feature selection methods obtained from the term scores.

$Rank = (r_1, r_2, ..., r_n)$ where r_i rank of the ith term, n is the total number of term.

$$C_{rank} = \sum_{i=1}^{M} \frac{Rank_i}{M} \tag{4.3}$$

There are various ways for assigning rankings such as standard competition ranking, modified competition ranking, dense ranking, ordinal ranking and fractional ranking. In this study, we rank the terms according to the descending order of their scores with standard competition ranking strategy. In competition ranking *("1, 2, 2, 4" ranking),* terms that have the same score get the same ranking number and then a gab is left in the ranking numbers.

4.2. Proposed Combinations of Feature Selection Methods

In this study, the main concern is to investigate the effectiveness of combining the individual metrics on the performances of text categorization. In order to find the best combinations, we also present seven new methods to combine the feature selection metrics beside the score and rank combinations. In Section 7, the results of the proposed combination methods are also compared and discussed in detail.

First two proposed combinations are similar to score combination that can be seen as the different modifications of the score combination. On the other hand, the other five can be seen as a product combination that are used both score and rank value of the terms.

4.2.1. C1 Combination – Logarithmic Combination

Score combination is simply averaging the term scores of the feature selection methods. This proposed combination is based on the same principle but the logarithmic scores of the terms are averaged instead of term scores. The basic idea is to enlarge the interval between the highest and the lowest scores taking advantage of the logarithm. Thus, it is possible to increase the difference between the high score and the low scores. The interval between the scores increases as the scores decrease and it decreases as the scores increase. The natural logarithm is used as a logarithm function and if the score is zero, then it will replace the value with 0.00001 in order to compute the logarithm. The calculation of the C_1 Combination is given by the following formula:

$$C_1 = \sum_{i=1}^{M} \frac{ln\,(Score_i^{'})}{M} \tag{4.4}$$

4.2.2. C2 Combination – Square Combination

The second proposed combination is also based on the score combination but the square of the term scores are averaged instead of term scores. Whereas the interval between the highest and the lowest scores is not changed, the interval between the scores exponentially increases as the scores increase. Unlike previous methods, the difference between the scores increases as the scores increase and decreases as the scores decrease. In other words, if the term is considered important for classification, this method increases the importance of the term among the other terms before the combination. The formula is:

$$C_2 = \sum_{i=1}^{M} \frac{(Score_i^{'})^2}{M} \tag{4.5}$$

4.2.3. C3 Combination – Product Combination with Fraction

In the third method firstly the rank of the term is multiplied by the scores of the terms and sums the results of the multiplication of the each feature selection methods and then divides one by the sum in order to combine the outputs of the feature selection metrics. If the score is 1, then it will replace the value with 0.99999 in order to avoid division by zero. In feature selection step the terms are selected with the highest values. The formula is:

$$C_3 = \frac{1}{\sum_{i=1}^{M} Rank_i \; x \; (1 - Score_i^{'})} \tag{4.6}$$

When evaluated the performances of the combinations in classification in Section 7, we see that both score and rank combinations of the feature selection methods improve the performance of the individual methods. Thus, we decide to use both rank and score values in the same function. The main aim of the "product combinations" C3-C4-C5-C6 and C7 is to take advantage of the score and rank of the terms while determining the most discriminative terms. In these methods the rank of the term is as important factor as the score of the term in the combination. For instance if the scores of terms of the different feature selection methods are same, then the term with highest rank is considered more important. Following methods are derived from the same principle with varied versions of the product combination.

4.2.4. C4 Combination – Product Combination with Fraction

Differ from the third proposed method, the forth method multiplies the logarithm of the rank by the square of the score of the term. The calculation of the C_4 Combination is given by the following formula:

$$C_4 = \frac{1}{\sum_{i=1}^{M} \ln(Rank_i)\ x\ (1 - Score_i')^2} \tag{4.7}$$

4.2.5. C5 Combination – Product Combination with Fraction

Differ from the third proposed method; the fifth method multiplies the square root of the rank by the square of the score of the term. The calculation of the C_5 Combination is given by the following formula:

$$C_5 = \frac{1}{\sum_{i=1}^{M} (Rank_i)^{1/2}\ x\ (1 - Score_i')^2} \tag{4.8}$$

4.2.6. C6 Combination – Logarithmic Product Combination

The sixth method multiplies the logarithm of the rank by the logarithm of the score of the term and then *sums the results* of multiplication of the each feature selection method in order to combine the different feature selection methods. The calculation of the C_6 Combination is given by the following formula:

$$C_6 = \sum_{i=1}^{M} ln(Rank_i)\, x \ln(Score_i^{'}) \tag{4.9}$$

4.2.7. C7 Combination – Product Combination

The final method multiplies the square root of the rank by the logarithm of the score of the term and then sums the results of multiplication of the each feature selection method in order to combine the different feature selection methods. The formula is:

$$C_7 = \sum_{i=1}^{M} (Rank_i)^{1/2}\, x \ln(Score_i^{'}) \tag{4.10}$$

5. SYSTEM ARCHITECTURE

In this section we present our text categorization system consists of the different steps. We characterize the functionality of the each step and describe interactions between individual steps. Figure 5.1 demonstrates the system architecture in detail. The main steps of the system are the following:

5.1. Preprocessing

The first step is the preprocessing of the datasets. In this step each document is parsed, non-alphabetic characters and mark-up tags are discarded, case-folding is performed, stopwords are eliminated according to the stopword list of the SMART system and then each word is stemmed using Porter's stemmer. At the end of these processes the category list, term list, term matrix and category matrix of the training documents and term matrix and category matrix of the test documents are created. Finally the tf-idf weighting is calculated for each remaining word in the documents.

5.2. Feature Combination

The second step combines the feature selection metrics with the varied combination approaches.

Firstly, each feature selection metric computes the scores of the all terms and gives the higher scores to terms considered more informative for classification. In order to combine the outputs of different feature selection methods, the scores of each term from the varied feature selection methods are normalized.

Secondly the ranks of the all terms are computed by ranking the scores of the terms with standard competition ranking. In this step all possible binary-combinations of different feature selection methods are generated.

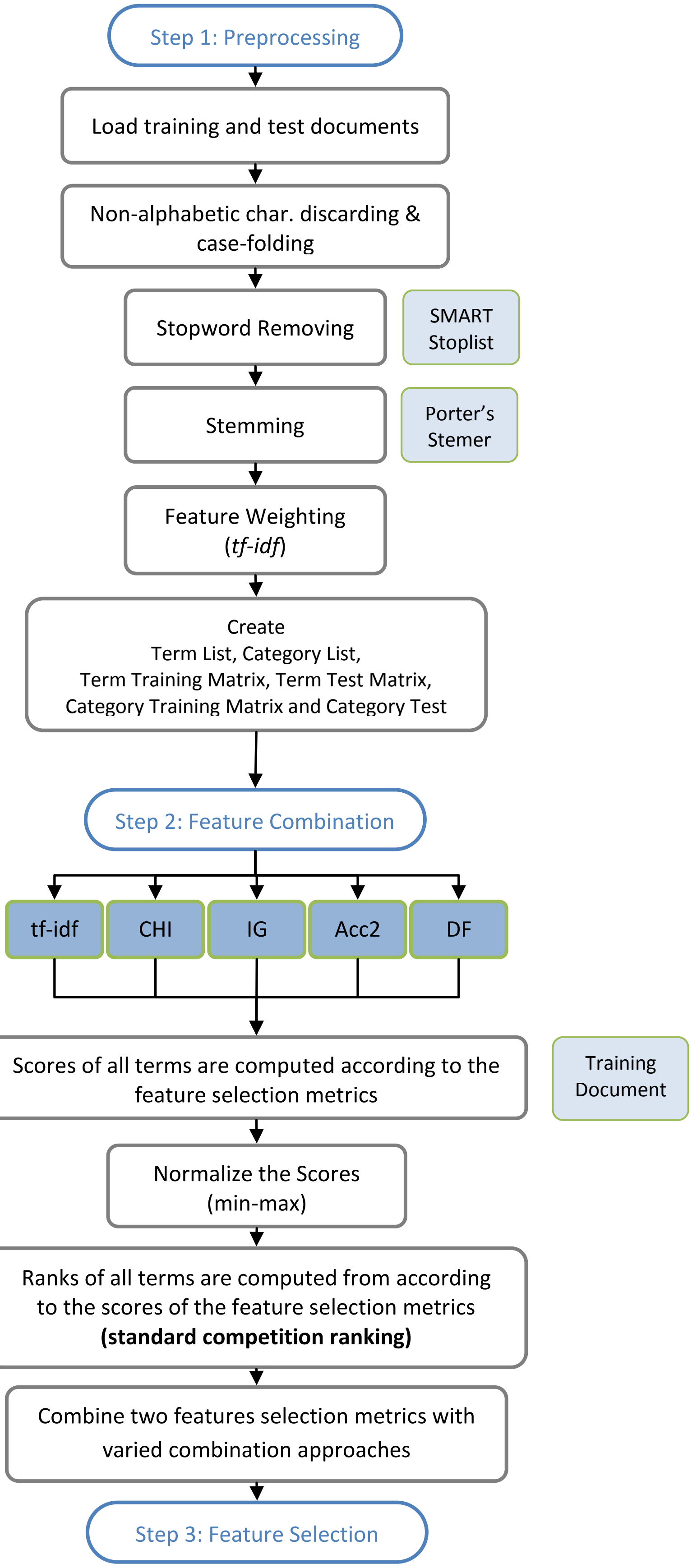


Figure 5.1. Preprocessing and combination steps

5.3. Feature Selection

The third step is feature selection that reduces the dimensionality by ranking all terms according to their importance estimated by combination and then selecting a given number of terms from the term list with the highest values. After selecting, the topic list, term list, and topic and term matrixes of the documents are reformed according to the selected features.

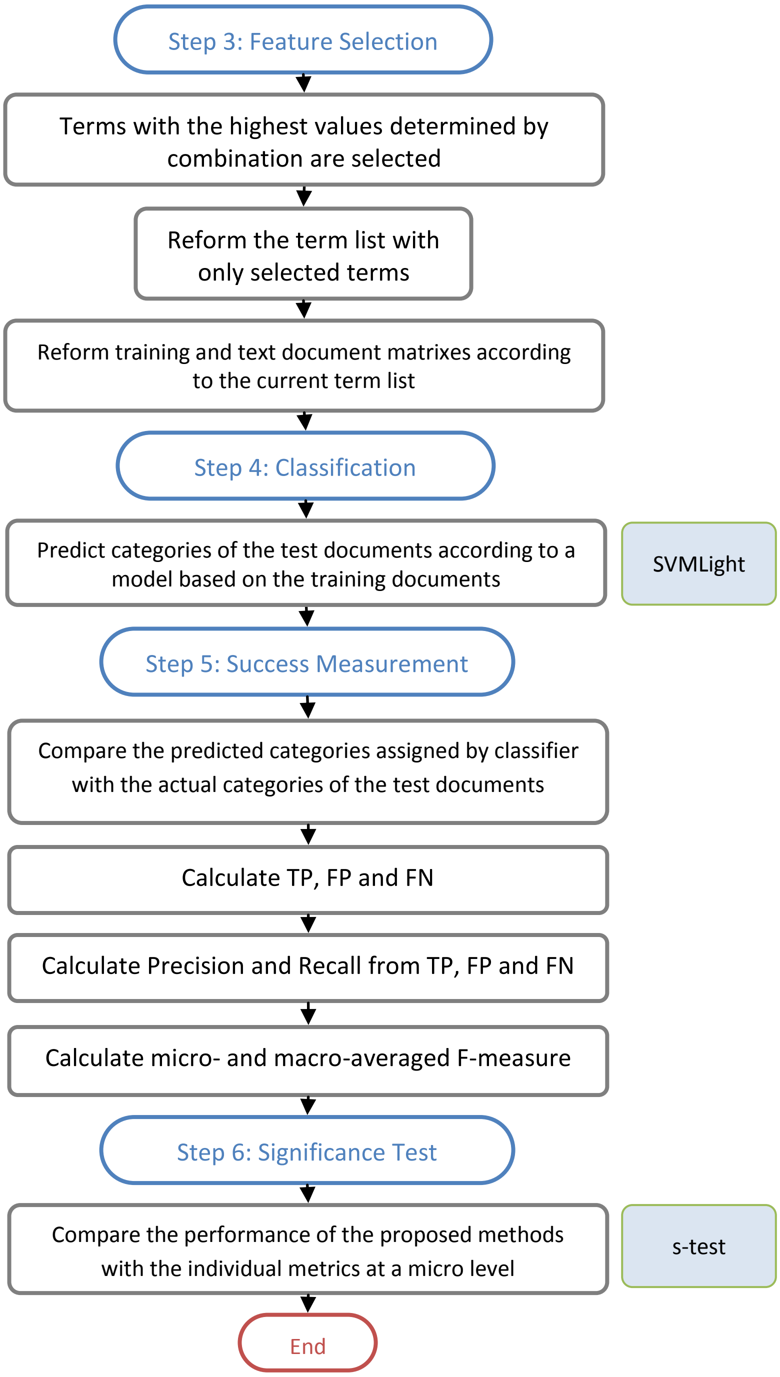


Figure 5.2. Feature selection, classification and success measurement steps

5.4. Classification

In the forth step, the category of the each test document is predicted according to the model derived from the training documents by using SVM_{Light}.

5.5. Success Measurement

In the fifth step the performance of the classifier is evaluated according to the F-measure results. In order to compare the predicted categories assigned by classifier with the actual categories of the test documents, first of all the number of True Positives, False Negatives and False Positives are determined, then precision and recall is computed using these values and finally the micro- and macro-averaged F-measures are calculated from precision and recall.

5.6. Significance Test

In the final step the best individual metrics are compared with the proposed combinations that improve the performance of the individual metrics. In order to measure the significance of the improvements in the proposed methods, micro sign test (s-test) is used. Micro sign test is designed to evaluate the performance of the two different systems at a micro level based on their binary decisions on all the document/category pairs. We follow Yang and Liu to perform the s-test [15].

In this significance test the correctness of the system decisions are computed by counting *n*, the number of times that systems A and system B differ and *k*, the number of times that system A is better than system B. The significance level (P-value) can be computed using the binomial distribution if $n \leq 12$ and if $n > 12$, the P-value (1-sided) can be computed using the normal distribution.

$$Z = \frac{k - 0.5.n}{0.5\sqrt{n}} \tag{5.1}$$

Standard z values are calculated and the corresponding confidence levels are determined according to the standard normal distribution.

6. EXPERIMENTAL SETTINGS

6.1. Classifier

In this study, we use SVM as a classifier which outperformed other classification methods in text categorization consistently in previous studies [3, 14, 15, 18, 31, 32, 37, 39].

Support Vector Machine was introduced as a statistical learning theory by Vapnik in 1995 at AT&T Bell Laboratories [40, 42]. It is based on the Structural Risk Minimization principle from computational learning theory [41]. The basic idea of this principle is to find a hypothesis for guarantee the minimum true error. In here, true error means the probability that a hypothesis will make an error on an unseen and randomly selected test example [18]. SVM is designed for solving two-class problems and the idea behind SVM is to find a hyperplane in n-dimensional space that separates the positive training examples from the negative examples with the largest possible margin in order to determine the best separation between the two classes. In text categorization, learned hyperplanes separates the documents in input space that belong to different topics.

One of the reasons the success of the SVM in text categorization is its capability in very high dimensional feature vectors, given that these vectors are sparse [32]. Because the learning process of SVM is independent from the dimensionality of the feature space while it measures the complexity of hypothesis according to the margin which means it separates the data instead of the number of feature [18]. Another feature that distinguishes SVM from other classifiers is the *generalization ability*. The decision function is determined by assuming that the training data which belongs to different classes does not overlap with each other. So the distance from the training data is maximized and in this way SVM prevents overfitting to training data [31]. In addition to, SVM provide a fast and effective classification that can easily incorporate new documents [40]. Thus, we can say that SVM ideally suitable for text categorization.

In the study, we use *SVM*^{light} with the default parameter settings that a linear kernel has been used. The *SVM*^{light} system is a very efficient implementation of SVMs that was developed by Joachims, 1999, at the University of Dortmund and has been commonly used in previous studies.

6.2. Datasets

We perform our experiments on five standard datasets, widely used in text categorization research. The properties of these datasets are summarized in Table 6.1. We divide these five datasets into 3 categories according to their skewness. The skewness is calculated by dividing the standard deviation of the class distribution by the mean of the distribution. The first dataset Classic3 is a homogenous dataset where all the classes are nearly equally well represented in the training set and each class is disjoint from each other clearly. Hitech and LA1 are categorized as skew datasets in our study because they are neither homogenous as Classic3 nor highly skew as the Wap and Reuters. Finally, Wap and Reuters-21578 are categorized as a highly skew dataset with varying class distributions. These datasets are particularly hard to categorize since the rare classes are dominated by the common classes. In addition, there is a strong semantic overlap between the topics since both Wap and Reuters consist of general topics that are very close to each other and share many common terms. In order to divide the Reuters-*21578* dataset into training and test sets, we use ModApte splitting method that has been mostly used in the literature. In Section 7, we discuss the property of each dataset in more detail.

Dataset	# of documents	# of training documents	# of test documents	# of Classes	# of Terms	min class size	max class size	Skewness (sd/mean)
Classic3	3891	2699	1192	3	10930	1033	1460	homogenous (0.13)
Hitech	2300	1530	770	6	18867	116	603	skew (0.45)
LA1	3204	2134	1070	6	25024	273	943	skew (0.45)
Wap	1560	1047	513	20	8064	5	341	highly skew (0.96)
Reuters-21578	12902	9603	3299	90	20308	2	3964	highly skew (3.32)

Table 6.1. Properties of the datasets

6.3. Performance Measures

To evaluate the performance of the keyword selection approaches, we use the commonly used F-measure metric which is equal to the harmonic mean of recall (ρ) and precision (π) [35]. They are defined as:

$$\pi = \frac{TP}{TP + FP}, \ \rho = \frac{TP}{TP + FN} \tag{6.1}$$

$$F = \frac{2\pi\rho}{\pi + \rho}$$

The idea behind the F-measure can be explained in Figure 6.1. The right circle represents the all defective set and the left represents the set that were classified as defective by a classifier. The intersection between these sets represents the true positive (TP) while the remaining parts represent false negative (FN) and false positive (FP). Accuracy of the classifier is defined by measuring the extent of the intersection between the two sets [33].

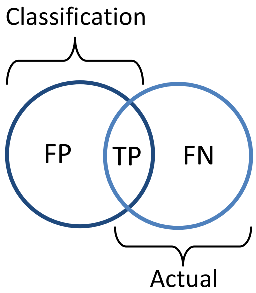


Figure 6.1. Demonstration of the F-measure

Since the absolute size is not meaningful, this value should be normalized by the proportional area. The F-measure is defined as:

$$\text{F} = \frac{2(TP)}{FP + FN + 2(TP)} \tag{6.2}$$

The F-measure values are in the interval [0-1]. When the two sets are identical, F-measure obtains the highest value and it obtains the lowest value when the two sets are mutually exclusive. Thus, larger F-measure values correspond to higher classification quality.

F-measure can be computed by two different alternatives, micro-averaged F-measure and macro-averaged F-measure. In this way, the overall F-measure score of the entire classification problem can be computed by using these different types of averaging methods. [35]

Micro-averaged F-measure gives equal weight to each document and therefore it tends to be dominated by the classifier's performance on common categories while reflects the overall accuracy better. Precision and recall are obtained by summing over all individual decision:

$$\pi = \frac{TP}{TP+FP} = \frac{\sum_{i=1}^{C} TPi}{\sum_{i=1}^{C} TPi + FPi} \quad , \quad \rho = \frac{TP}{TP+FN} = \frac{\sum_{i=1}^{C} TPi}{\sum_{i=1}^{C} TPi + FNi}$$

where C indicates the number of categories.

$$\textit{Micro-averaged F-measure} = \frac{2\pi\rho}{\pi+\rho} \qquad (6.3)$$

On the other hand *Macro-averaged F-measure* gives equal weight to each category regardless of its frequency and thus it is influenced more by the classifier's performance on rare categories. Precision and recall are first computed locally for each category and then F-measure is computed globally by averaging over the decisions of all categories:

$$\pi_i = \frac{TPi}{TPi + FPi} \, , \; \rho_i = \frac{TPi}{TPi + FNi}$$

$$F_{\mathrm{i}} = \frac{2\pi i\rho i}{\pi i + \rho i}$$

$$\textit{Macro-averaged F-measure} = \frac{\sum_{t=1}^{M} Fi}{M} \qquad (6.4)$$

In text classification, TP_i is the number of documents that are assigned correctly to class *i*. FP_i is the number of documents that are assigned incorrectly to class *i* by the classifier but which actually do not belong to class *i* and *FNi* is the number of documents not assigned to class *i* by the classifier but which actually belong to class *i*.

7. RESULTS AND DISCUSSION

In this section, we show the experimental results of the study with a comprehensive and comparative discussion. The experimental results of each dataset (Classic3*,* Hitech*,* LA1*,* Wap *and* Reuters) are demonstrated as follow: firstly the properties of the datasets are described in detail, secondly the performance of the five feature selection methods (term frequency-inverse document frequency *(tf-idf),* chi-square statistics *(CHI),* information gain *(IG),* Accuracy2 *(Acc2)* and document frequency thresholding *(DF)*) are analyzed on these datasets, then the results of the binary score and rank combinations of the feature selection methods are compared with the results of these five individual methods that analyzed and finally we present the results of the proposed combination methods. In addition, we summarize and overview all the results of the study at the end of the section.

Tables 7.1, 7.14, 7.27, 7.41 and 7.55 show the micro- and macro-averaged F-measure results of the five feature selection methods with global and local policy as a function of different number of keywords ranging from 10 to 2000 for the Classic3*,* Hitech*,* LA1*,* Wap and Reuters datasets, respectively. In the tables, the global and local versions of the existing metrics are denoted by (g) and (l), respectively. The highest score among the methods for each keyword number is shown in red font and the success rate without feature selection is shown as a last column "ALL" in these tables. The highest values for each keyword number are indicated as "MAX" rows and the averages of the each keyword number are also indicated as "AVERAGE" rows at the tables.

Tables 7.2, 7.15, 7.28, 7.42 and 7.56 show the micro-averaged F-measure results of the score and rank combinations with global policy, Tables 7.3, 7.16, 7.29, 7.43 and 7.57 show the micro-averaged F-measure results of the score and rank combinations with global policy, Tables 7.4, 7.17, 7.30, 7.44 and 7.58 show the micro-averaged F-measure results of the score and rank combinations with local policy and Tables 7.5, 7.18, 7.31, 7.45 and 7.59 show the macro-averaged F-measure results of the score and rank combinations with local policy for the datasets. All possible 2-combinations of the feature selection methods are presented in these tables. If the combination F-measure value is higher or equal to the highest F-measure of the individual metrics, it is shown in bold and underlined for each number of keywords ranging from 10 to 2000. If it is also the highest value in the column, then it is also shown in red. In addition if the combination F-measure value is lower than the highest F-measure of the individual metrics but it is higher than the value of the feature selection metrics that constitute it, then it is shown in blue.

Tables 7.8, 7.21, 7.34, 7.48 and 7.62 show the micro-averaged F-measure results of the proposed combination methods with global policy, Tables 7.9, 7.22, 7.35, 7.49 and 7.63 show the macro-averaged F-measure results of the proposed combination methods with global policy, Tables 7.12, 7.25, 7.38, 7.52 and 7.66 show the micro-averaged F-measure results of the proposed combination methods with local policy and Tables 7.13, 7.26, 7.39, 7.53 and 7.67 show the micro-averaged F-measure results of the proposed combination methods with local policy for the datasets. . If the combination F-measure value is higher or equal to the highest F-measure of the individual metrics, it is shown in bold and underlined for each keyword number. If it is also the highest value in the column, then it is also shown in red. While if it is higher than the highest F-measure value of the score and rank combinations, then it is shown in red, bold and double underlined. In addition if the combination F-measure value is lower than the highest F-measure of the individual metrics but it is higher than the value of the feature selection metrics that constitute it, then it is shown in blue.

In this study, the average of the F-measures results is computed by only averaging the keyword numbers from 100 to 2000 while the performance of the classification with the keywords less than 100 is very low.

In each figure, we also show the best individual metric results as a red straight line in order to compare the results of the individual metrics with their combinations. In addition, if the gap between the performance of the best individual metric and the second best individual metric is high, the second best individual metric is also shown as a violet straight line in some figures.

7.1. Homogenous Dataset

7.1.1. The Classic 3 Dataset

7.1.1.1. Property of the Dataset

The Classic3 dataset is a well known collection of documents composed of 3,891 abstracts from 3 disjoint research fields as shown in Figure 1: 1,398 CRANFIELD documents from aeronautical system papers, 1,033 MEDLINE documents from medical journals, and 1,460 CISI documents from information retrieval papers.

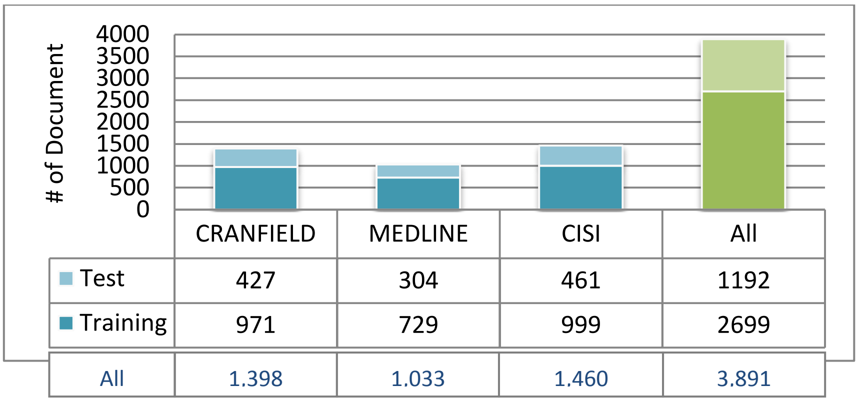


Figure 7.1. Property of the Classic3 dataset

Classic3 has been used by many researchers [6, 7, 8, 44] in text mining and it is chosen as a *"homogenous dataset"* in our study, where all the classes are nearly equally well represented in the training set. First two thirds of each class is selected for the training set and the remaining one third is used for testing.

The most significant feature of the Classic3 dataset is that each class is disjoint from each other clearly, which means about 50 percent of the terms occur in only one class and the documents that share many common terms belong to the same class in the dataset. Since the classes are disjoint from each other, the Classic3 dataset is relatively easy to classify among other datasets in our study.

7.1.1.2. Analysis of the Existing Metrics

In this section, we discuss the results of the five well-known feature selection metrics used in the text categorization domain on homogenous datasets. Classic3 is a good example of a homogenous dataset. Table 7.1 shows the micro- and macro-averaged F-measure results for the Classic3 dataset. We compare both the local and global version of the feature selection metrics on the Classic3 dataset to these analyses.

The Classic3 dataset has the highest F-measure results among the datasets since other datasets consist of many classes with skew class distributions. As we mentioned above, the Classic3 dataset contains sufficient number of documents from three well-separated classes. Thus the results of the micro- and macro-averaged F-measures are very close to each other.

When the test documents are classified with all words without applying any feature selection method, the classifier achieves the highest micro-F and macro-F values of 99.4%. This situation can be explained by the fact that all classes have enough training documents and most of the terms occur in only one class. This means that almost all terms are important in the classification of the documents and improve the classifier's performance individually. This fact also explains why the classifier achieves better performance as the number of keywords increases from 10 to 2000 in homogenous datasets. But still feature selection is a necessary process if we want to overcome time and space limitations.

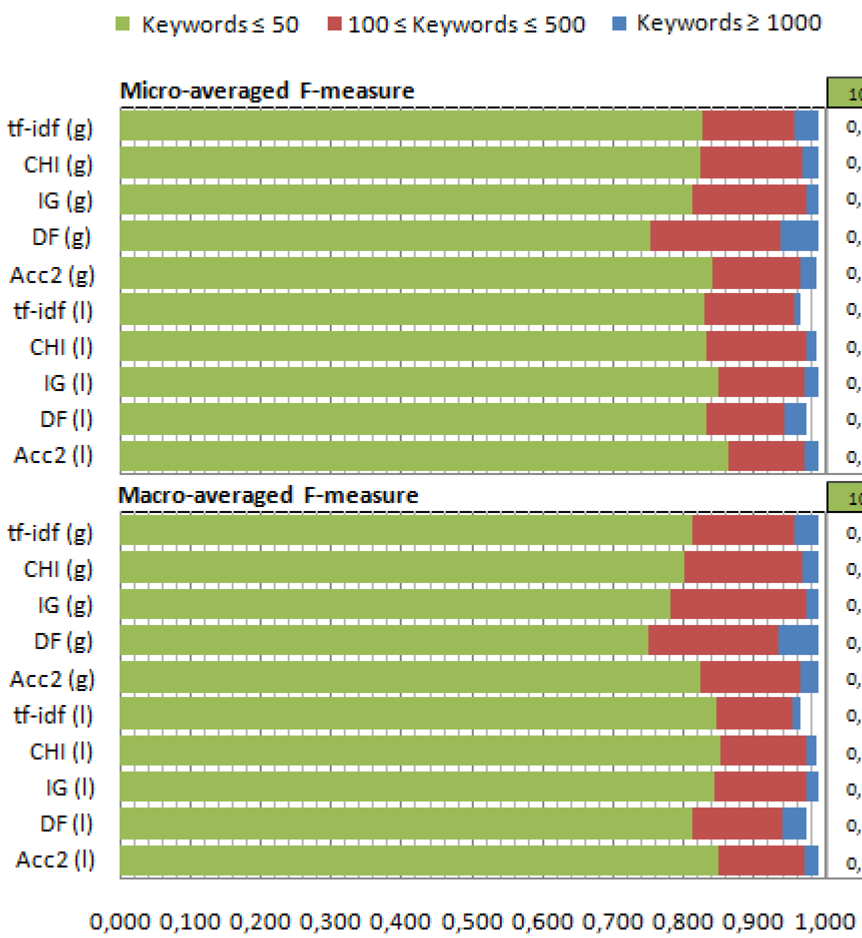
Another observation about homogenous datasets is that the local policy performs significantly better than the global policy when the keyword number is low. For our homogenous dataset Classic3, this range is between 10 to 200 keywords. On the other hand, when the keyword number is high (more than 200 keywords), the global policy performs better than the local policy in homogenous datasets. This is due to the fact that in corpus-based approach, a single set of keywords is selected for all classes while in class-based approach, a distinct set of keywords is selected for each class [5]. The aim of the global policy is to find general keywords for all classes; on the other hand the local policy tries to find a small but crucial portion of keywords for each class. For this reason, the local policy is more successful with a few number of keywords by selecting more important and useful keywords for classification. The global policy is not successful finding keywords that carry critical information for all classes when the keyword number is low although the accuracy increases gradually as the number of keywords increases. But it can be noted that there is not much difference between the global and local policies in homogenous datasets.

Micro-F	10	30	50	100	200	500	1000	1500	2000	All
tf-idf (g)	0,701	**0,873**	0,901	0,937	0,956	0,981	0,988	**0,992**	**0,992**	0,994
CHI (g)	0,732	0,848	0,890	**0,956**	0,956	**0,989**	**0,991**	**0,992**	**0,992**	0,994
IG (g)	0,702	0,848	0,886	**0,956**	**0,974**	0,988	**0,991**	0,990	**0,992**	0,994
DF (g)	0,622	0,800	0,833	0,894	0,943	0,970	0,986	**0,992**	**0,992**	0,994
Acc2 (g)	**0,736**	0,867	**0,916**	0,944	0,967	0,984	0,988	0,989	0,991	0,994
MAX	0,736	0,873	0,916	0,956	0,974	0,989	0,991	0,992	0,992	
Macro-F	10	30	50	100	200	500	1000	1500	2000	**All**
tf-idf (g)	0,665	**0,871**	0,898	0,936	0,953	0,980	0,989	0,992	**0,992**	0,994
CHI (g)	**0,709**	0,821	0,870	**0,956**	0,956	**0,990**	**0,991**	**0,993**	**0,992**	0,994
IG (g)	0,665	0,811	0,863	0,955	**0,975**	0,988	**0,991**	0,990	**0,992**	0,994
DF (g)	0,623	0,798	0,831	0,893	0,941	0,970	0,987	0,992	**0,992**	0,994
Acc2 (g)	0,690	0,865	**0,914**	0,944	0,967	0,985	0,988	0,990	0,991	0,994
MAX	0,709	0,871	0,914	0,956	0,975	0,990	0,991	0,993	0,992	
Micro-F	10	30	50	100	200	500	1000	1500	2000	**All**
tf-idf (l)	0,653	0,895	0,939	0,951	0,959	0,960	0,964	0,965	0,971	0,994
CHI (l)	0,638	**0,915**	**0,947**	**0,963**	**0,974**	0,981	0,987	0,989	0,990	0,994
IG (l)	0,735	0,896	0,918	0,958	0,973	**0,986**	0,989	**0,992**	**0,991**	0,994
DF (l)	0,745	0,865	0,883	0,917	0,949	0,964	0,973	0,973	0,978	0,994
Acc2 (l)	**0,787**	0,880	0,926	0,958	0,972	0,985	**0,991**	0,991	**0,991**	0,994
MAX	0,787	0,915	0,947	0,963	0,974	0,986	0,991	0,992	0,991	
Macro-F	10	30	50	100	200	500	1000	1500	2000	**All**
tf-idf (l)	0,720	0,880	0,935	0,950	0,957	0,959	0,964	0,964	0,970	0,994
CHI (l)	0,706	**0,908**	**0,945**	**0,963**	**0,974**	0,981	0,987	0,989	0,990	0,994
IG (l)	0,728	0,889	0,912	0,959	**0,974**	**0,986**	0,990	**0,992**	**0,991**	0,994
DF (l)	0,720	0,848	0,871	0,908	0,945	0,964	0,973	0,973	0,978	0,994
Acc2 (l)	**0,761**	0,867	0,923	0,958	0,972	0,985	**0,991**	0,991	**0,991**	0,994
MAX	0,761	0,908	0,945	0,963	0,974	0,986	0,991	0,992	0,991	

Table 7.1. Micro- and macro-averaged F-measures for Classic3 dataset

In homogenous datasets generally *CHI*, *IG* and *Acc2* outperform other methods in both the global and local policies. In global policy *CHI* achieves the best micro- and macro-averaged F-measure results with a high number of keywords from 100 to 2000 and in local policy it gets the highest scores with a few number of keywords from 30 to 200. In global policy *Acc2's* performance is slightly better than *CHI* and *IG* with a few number of keywords. In addition to this, both in global and in local policy IG is successful and has many highest scores with a high number of keywords, higher than 50 keywords. Figure 7.3 shows the comparison of the averages of F-measures under feature number criterion both in global and local policies for the Classic3 dataset.

Although document classification with all words achieves the highest performance of 99.4%, *CHI* in global policy achieves a very close performance of 99.3% with 1500 keywords. In global policy all the feature selection metrics perform almost similarly with 1500 and 2000 keywords except *Acc2.*



Micro-averaged F-measure	10-50	100-500	1000-2000
tf-idf (g)	0,825	0,958	0,991
CHI (g)	0,823	0,967	0,992
IG (g)	0,812	0,973	0,991
DF (g)	0,752	0,936	0,990
Acc2 (g)	0,840	0,965	0,989
tf-idf (l)	0,829	0,957	0,967
CHI (l)	0,833	0,973	0,989
IG (l)	0,850	0,972	0,991
DF (l)	0,831	0,943	0,975
Acc2 (l)	0,864	0,972	0,991

Macro-averaged F-measure	10-50	100-500	1000-2000
tf-idf (g)	0,811	0,956	0,991
CHI (g)	0,800	0,967	0,992
IG (g)	0,780	0,973	0,991
DF (g)	0,751	0,935	0,990
Acc2 (g)	0,823	0,965	0,990
tf-idf (l)	0,845	0,955	0,966
CHI (l)	0,853	0,973	0,989
IG (l)	0,843	0,973	0,991
DF (l)	0,813	0,939	0,975
Acc2 (l)	0,850	0,972	0,991

Figure 7. 2. Comparison of the averages of micro- and macro-averaged F-measures for Classic3 dataset

7.1.1.3. Analysis of Score and Rank Combinations

In the previous section we compared and discussed the results of the five well-known feature selection metrics on a homogenous dataset.

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,701	**0,873**	0,901	0,937	0,956	0,981	0,988	**0,992**	**0,992**
CHI (g)		0,732	0,848	0,890	**0,956**	0,956	**0,989**	**0,991**	**0,992**	**0,992**
IG (g)		0,702	0,848	0,886	**0,956**	**0,974**	0,988	**0,991**	0,990	**0,992**
DF (g)		0,622	0,800	0,833	0,894	0,943	0,970	0,986	**0,992**	**0,992**
Acc2 (g)		**0,736**	0,867	**0,916**	0,944	0,967	0,984	0,988	0,989	0,991
	MAX	0,736	0,873	0,916	0,956	0,974	0,989	0,991	0,992	0,992
Micro-F	Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (g)		**0,759**	0,844	0,915	0,955	0,964	0,984	0,989	**0,992**	**0,992**
|S| tf-idf&IG (g)		0,699	0,840	**0,916**	0,950	0,963	0,983	0,988	**0,992**	**0,992**
|S| tf-idf&DF (g)		0,724	0,817	0,879	0,927	0,955	0,979	0,988	0,991	**0,992**
|S| tf-idf&Acc2 (g)		**0,760**	**0,875**	0,909	0,943	0,961	0,984	0,989	0,991	**0,992**
|S| CHI&IG (g)		0,732	0,853	0,890	0,955	**0,975**	**0,989**	**0,991**	0,991	0,991
|S| CHI&DF (g)		**0,743**	0,848	0,898	0,942	0,961	0,984	0,989	0,990	**0,992**
|S| CHI&Acc2 (g)		0,732	0,870	**0,920**	0,952	0,971	**0,989**	0,989	**0,992**	**0,992**
|S| IG&DF (g)		0,730	0,845	0,898	0,943	0,957	0,984	**0,991**	0,991	**0,992**
|S| IG&Acc2 (g)		0,702	0,866	**0,921**	0,954	0,972	0,988	0,990	**0,992**	**0,993**
|S| DF&Acc2 (g)		**0,743**	0,835	0,881	0,940	0,961	0,981	0,990	0,989	**0,992**
	MAX	**0,760**	**0,875**	**0,921**	0,955	**0,975**	**0,989**	**0,991**	**0,992**	**0,993**
	AVERAGE	**0,732**	0,849	0,903	**0,946**	0,964	0,985	0,989	**0,991**	**0,992**
Micro-F	Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (g)		0,692	0,844	**0,922**	0,952	0,972	0,988	0,990	0,991	**0,992**
|R| tf-idf&IG (g)		0,694	0,844	**0,923**	0,950	0,969	0,988	0,990	0,991	**0,992**
|R| tf-idf&DF (g)		**0,736**	0,823	0,871	0,928	0,950	0,977	0,986	0,989	**0,992**
|R| tf-idf&Acc2 (g)		**0,760**	**0,879**	**0,916**	0,944	0,962	0,985	**0,991**	0,991	**0,992**
|R| CHI&IG (g)		0,732	0,853	0,890	0,954	**0,976**	**0,989**	**0,991**	0,991	**0,992**
|R| CHI&DF (g)		0,675	0,865	0,909	0,946	0,964	0,986	0,989	**0,992**	**0,992**
|R| CHI&Acc2 (g)		0,743	0,844	**0,924**	**0,956**	0,971	0,988	0,989	**0,992**	**0,993**
|R| IG&DF (g)		0,676	0,862	0,915	0,943	0,964	0,987	**0,991**	0,991	**0,993**
|R| IG&Acc2 (g)		0,743	0,836	**0,920**	0,950	0,971	**0,989**	0,989	**0,992**	**0,992**
|R| DF&Acc2 (g)		**0,764**	0,855	0,875	0,938	0,960	0,983	0,989	0,990	**0,992**
	MAX	**0,764**	**0,879**	**0,924**	**0,956**	**0,976**	**0,989**	**0,991**	**0,992**	**0,993**
	AVERAGE	0,722	**0,851**	**0,907**	**0,946**	**0,966**	**0,986**	**0,990**	**0,991**	**0,992**

Table 7.2. In global policy, micro-averaged F-measures of the score and rank combinations for Classic3 dataset

Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,665	**0,871**	0,898	0,936	0,953	0,980	0,989	0,992	**0,992**
CHI (g)		**0,709**	0,821	0,870	**0,956**	0,956	**0,990**	**0,991**	**0,993**	**0,992**
IG (g)		0,665	0,811	0,863	0,955	**0,975**	0,988	**0,991**	0,990	**0,992**
DF (g)		0,623	0,798	0,831	0,893	0,941	0,970	0,987	0,992	**0,992**
Acc2 (g)		0,690	0,865	**0,914**	0,944	0,967	0,985	0,988	0,990	0,991
	MAX	0,709	0,871	0,914	0,956	0,975	0,990	0,991	0,993	0,992
Macro-F	Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (g)		**0,710**	0,801	0,913	0,953	0,963	0,985	0,989	0,992	**0,992**
|S| tf-idf&IG (g)		0,556	0,801	**0,914**	0,949	0,962	0,983	0,988	0,992	**0,993**
|S| tf-idf&DF (g)		**0,724**	0,815	0,877	0,926	0,954	0,978	0,989	0,991	**0,992**
|S| tf-idf&Acc2 (g)		**0,759**	**0,873**	0,907	0,942	0,960	0,984	0,989	0,991	**0,992**
|S| CHI&IG (g)		**0,709**	0,825	0,870	0,954	**0,975**	**0,990**	**0,991**	0,991	0,991
|S| CHI&DF (g)		**0,745**	0,847	0,896	0,941	0,961	0,985	0,989	0,991	**0,992**
|S| CHI&Acc2 (g)		**0,709**	0,861	**0,919**	0,952	0,971	0,989	0,990	**0,993**	**0,992**
|S| IG&DF (g)		**0,732**	0,844	0,896	0,942	0,957	0,984	**0,991**	0,992	**0,993**
|S| IG&Acc2 (g)		0,665	0,860	**0,920**	0,954	0,972	0,988	0,990	0,992	**0,993**
|S| DF&Acc2 (g)		**0,745**	0,834	0,879	0,940	0,960	0,981	0,990	0,990	**0,992**
	MAX	**0,759**	**0,873**	**0,920**	0,954	**0,975**	**0,990**	**0,991**	**0,993**	**0,993**
	AVERAGE	**0,705**	**0,836**	0,899	**0,945**	0,963	0,985	**0,990**	**0,992**	**0,992**
Macro-F	Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (g)		0,550	0,801	**0,920**	0,950	0,971	0,989	0,990	0,991	**0,992**
|R| tf-idf&IG (g)		0,551	0,801	0,921	0,949	0,968	0,988	0,990	0,991	**0,993**
|R| tf-idf&DF (g)		**0,737**	0,821	0,870	0,927	0,949	0,976	0,987	0,990	**0,992**
|R| tf-idf&Acc2 (g)		**0,759**	**0,877**	**0,915**	0,943	0,960	0,985	**0,992**	0,991	**0,993**
|R| CHI&IG (g)		**0,709**	0,825	0,870	0,954	**0,976**	0,989	**0,991**	0,991	**0,992**
|R| CHI&DF (g)		0,538	0,864	0,907	0,945	0,963	0,987	0,990	0,992	**0,992**
|R| CHI&Acc2 (g)		0,693	0,808	**0,923**	**0,956**	0,970	0,989	0,989	0,992	**0,993**
|R| IG&DF (g)		0,540	0,861	0,913	0,942	0,963	0,987	**0,991**	0,992	**0,993**
|R| IG&Acc2 (g)		0,693	0,795	**0,918**	0,949	0,971	0,989	0,990	0,992	**0,992**
|R| DF&Acc2 (g)		**0,764**	0,853	0,874	0,938	0,958	0,983	0,989	0,991	**0,993**
	MAX	**0,764**	**0,877**	**0,923**	**0,956**	**0,976**	**0,989**	**0,992**	**0,992**	**0,993**
	AVERAGE	0,653	0,830	**0,903**	**0,945**	**0,965**	**0,986**	**0,990**	0,991	**0,992**

Table 7.3. In global policy, macro-averaged F-measures of the score and rank combinations for Classic3 dataset

In this section we evaluate the performance of the score and rank combinations on a homogenous dataset Classic3. We perform all possible score and rank combinations of two feature selection metrics.

Tables 7.2 and 7.3 show the micro- and macro-averaged F-measure results of the combinations in global policy for the Classic3 dataset. In global policy among the 10 possible combinations of two feature selection metrics, *tf-idf & Acc2, CHI & IG, CHI & Acc2, IG & DF* and *IG & Acc2* score and rank combinations are more successful than other combinations based on the highest micro- and macro-averaged F-measure values for each keyword number.

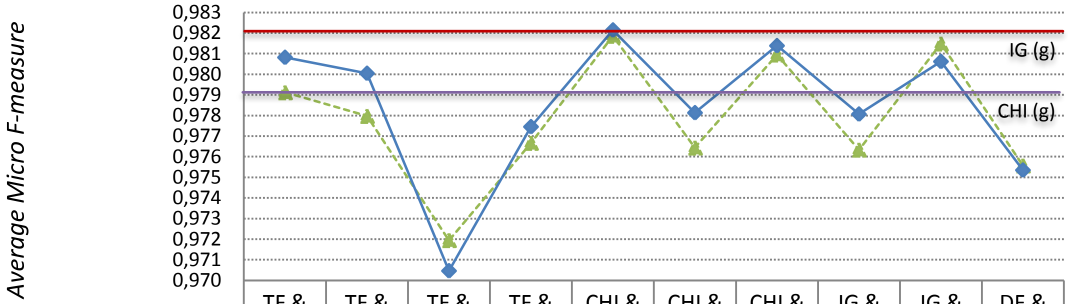
When we look at Tables 7.2 and 7.3, we can see that rank combination is slightly more successful than score combination in the homogenous dataset in global policy. For all keyword numbers from 10 to 2000, micro- and macro-averaged F-values of score and rank combinations are higher than or equal to the highest score of the individual methods. The highest values for each keyword number are indicated as MAX rows on each table.

Improvement of the F-measure values is more distinctive when the keyword number is low between 10 and 100. The main reason for this result is that the performances of the existing metrics are already high with large number of keywords because of homogeneity. The Classic3 dataset contains sufficient number of documents from three well-separated classes.

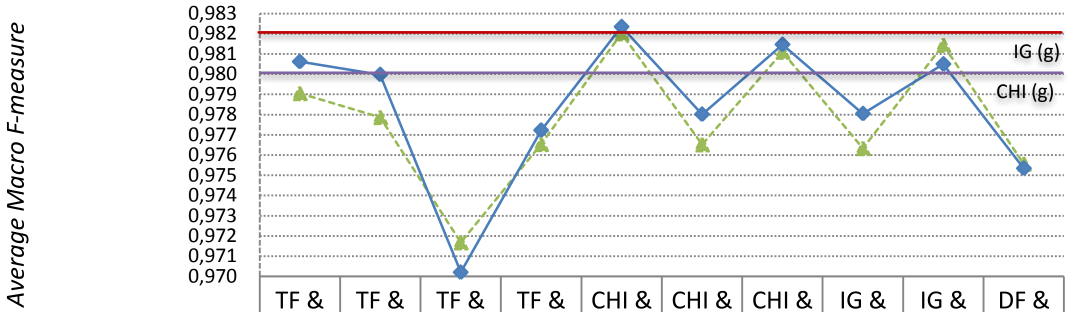
In global policy *Acc2* achieves the best F-measure results with a few number of keywords in the homogenous dataset. When we look at the related tables, we can see that both the score and rank combinations of *tf-idf & Acc2* are more successful than individual *Acc2.* Moreover the rank combination of *tf-idf & Acc2* is the most successful one among the combinations. *IG* is another successful method with many highest success rates when the keyword number is higher than 50. *CHI & IG* rank combination's performance is better than individual *IG.*

In global policy, among the individual feature selection metrics *IG* has the highest average of micro- and macro-averaged F-measure value of 98.2%. Only the score and rank combinations of *CHI & IG* achieve this success rate. Figure 3 shows the averages of micro- and macro-averaged F-measure values for the Classic3 dataset. CHI has the second highest average of micro- and macro-averaged F-measure values of 97.9% and 98.0% and many combinations are better than the performance of the CHI as seen in above figure.

As mentioned previously, although document classification with all words achieves the highest performance of 99.4%, only *CHI* shows very close performance of 99.3% with 1500 keywords. On the other hand, among the 10 possible score combinations 3 combinations and among the 10 possible rank combinations 5 combinations succeed 99.3% accuracy with 2000 keywords.



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,979	0,978	0,972	0,977	0,982	0,976	0,981	0,976	0,981	0,976
Rank Com. (g)	0,981	0,980	0,970	0,977	0,982	0,978	0,981	0,978	0,981	0,975



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,979	0,978	0,972	0,977	0,982	0,977	0,981	0,976	0,981	0,976
Rank Com. (g)	0,981	0,980	0,970	0,977	0,982	0,978	0,981	0,978	0,980	0,975

	tf-idf (g)	CHI (g)	IG (g)	DF (g)	Acc2 (g)
Micro	0,974	0,979	0,982	0,963	0,977
Macro	0,974	0,980	0,982	0,963	0,978

Figure 7.3. In global policy, comparison of score and rank combinations on the Classic3 dataset

When we perform score and rank combinations on the Classic3 dataset in local policy, score combinations are slightly better than rank combinations as seen in Tables 7.4 and 7.5. Table 7.1 shows among the individual feature selection metrics, *IG* has the highest values between 500 and 2000 keywords. In addition to this, among the score and rank combinations, score combination of *IG & Acc2* is significantly successful with high number of keyword in local policy.

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (l)		0,653	0,895	0,939	0,951	0,959	0,960	0,964	0,965	0,971
CHI (l)		0,638	**0,915**	**0,947**	**0,963**	**0,974**	0,981	0,987	0,989	0,990
IG (l)		0,735	0,896	0,918	0,958	0,973	**0,986**	0,989	**0,992**	**0,991**
DF (l)		0,745	0,865	0,883	0,917	0,949	0,964	0,973	0,973	0,978
Acc2 (l)		**0,787**	0,880	0,926	0,958	0,972	0,985	**0,991**	0,991	**0,991**
	MAX	0,787	0,915	0,947	0,963	0,974	0,986	0,991	0,992	0,991
Micro-F	Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (l)		0,648	0,907	0,944	0,957	0,965	0,974	0,978	0,982	0,985
|S| tf-idf&IG (l)		**<u>0,833</u>**	0,908	0,944	0,961	0,973	0,977	0,983	0,987	0,988
|S| tf-idf&DF (l)		**<u>0,818</u>**	0,880	0,915	0,941	0,955	0,960	0,969	0,969	0,974
|S| tf-idf&Acc2 (l)		**<u>0,837</u>**	0,910	0,926	0,962	0,973	0,976	0,984	0,987	0,987
|S| CHI&IG (l)		**<u>0,827</u>**	0,907	0,938	0,956	<u>0,977</u>	0,984	0,990	0,991	0,990
|S| CHI&DF (l)		**<u>0,829</u>**	0,893	0,915	0,948	0,970	0,972	0,981	0,985	0,986
|S| CHI&Acc2 (l)		<u>0,839</u>	0,899	0,934	0,961	0,973	0,983	0,990	**<u>0,992</u>**	**<u>0,991</u>**
|S| IG&DF (l)		**<u>0,823</u>**	0,891	0,920	0,955	0,970	0,978	0,986	0,991	0,990
|S| IG&Acc2 (l)		0,785	0,899	0,919	0,954	**<u>0,974</u>**	<u>0,986</u>	0,989	<u>0,993</u>	<u>0,993</u>
|S| DF&Acc2 (l)		**<u>0,801</u>**	0,880	0,923	0,945	**<u>0,974</u>**	0,977	0,987	0,989	0,987
	MAX	**0,839**	0,910	0,944	0,962	**0,977**	**0,986**	0,990	**0,993**	**0,993**
	AVERAGE	**0,804**	0,897	0,928	**0,954**	**0,970**	**0,977**	**0,984**	**0,987**	**0,987**
Micro-F	Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (l)		0,647	0,903	0,939	0,950	0,963	0,968	0,969	0,970	0,982
|R| tf-idf&IG (l)		0,645	0,905	0,941	0,954	0,966	0,969	0,969	0,978	0,987
|R| tf-idf&DF (l)		**<u>0,823</u>**	0,887	0,910	0,938	0,958	0,958	0,966	0,968	0,973
|R| tf-idf&Acc2 (l)		**<u>0,837</u>**	0,899	0,935	0,949	0,965	0,963	0,973	0,970	0,981
|R| CHI&IG (l)		**<u>0,819</u>**	0,907	0,941	0,953	<u>0,974</u>	0,984	<u>0,991</u>	0,991	0,990
|R| CHI&DF (l)		<u>0,844</u>	0,906	0,937	0,950	0,964	0,969	0,974	0,976	0,983
|R| CHI&Acc2 (l)		**<u>0,823</u>**	0,905	0,938	<u>0,964</u>	<u>0,974</u>	0,982	<u>0,991</u>	0,991	<u>0,991</u>
|R| IG&DF (l)		**<u>0,842</u>**	0,907	0,940	0,950	0,966	0,971	0,975	0,982	0,986
|R| IG&Acc2 (l)		0,784	0,899	0,926	0,953	0,973	0,984	0,990	<u>0,992</u>	<u>0,991</u>
|R| DF&Acc2 (l)		**<u>0,819</u>**	0,884	0,909	0,943	0,965	0,968	0,978	0,977	0,982
	MAX	**0,844**	0,907	0,941	**0,964**	**0,974**	0,984	**0,991**	**0,992**	**0,991**
	AVERAGE	0,788	**0,900**	**0,932**	0,950	0,967	0,972	0,977	0,980	0,985

Table 7.4. In local policy, micro-averaged F-measures of the score and rank combinations for Classic3 dataset

Macro-F	10	30	50	100	200	500	1000	1500	2000
tf-idf (l)	0,720	0,880	0,935	0,950	0,957	0,959	0,964	0,964	0,970
CHI (l)	0,706	**0,908**	**0,945**	**0,963**	**0,974**	0,981	0,987	0,989	0,990
IG (l)	0,728	0,889	0,912	0,959	0,974	**0,986**	0,990	**0,992**	**0,991**
DF (l)	0,720	0,848	0,871	0,908	0,945	0,964	0,973	0,973	0,978
Acc2 (l)	**0,761**	0,867	0,923	0,958	0,972	0,985	**0,991**	0,991	**0,991**
MAX	0,761	0,908	0,945	0,963	0,974	0,986	0,991	0,992	0,991
Macro-F / Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (l)	0,713	0,896	0,942	0,956	0,965	0,974	0,978	0,982	0,985
|S| tf-idf&IG (l)	**0,815**	0,898	0,942	0,960	0,973	0,978	0,983	0,988	0,988
|S| tf-idf&DF (l)	**0,796**	0,864	0,903	0,937	0,954	0,959	0,969	0,968	0,973
|S| tf-idf&Acc2 (l)	**0,824**	0,900	0,917	0,962	0,973	0,976	0,984	0,988	0,988
|S| CHI&IG (l)	**0,813**	0,901	0,935	0,956	**0,977**	0,985	0,990	0,991	0,990
|S| CHI&DF (l)	**0,813**	0,885	0,908	0,946	0,969	0,972	0,981	0,984	0,986
|S| CHI&Acc2 (l)	**0,825**	0,893	0,929	0,961	0,973	0,984	0,990	**0,992**	**0,991**
|S| IG&DF (l)	**0,808**	0,884	0,914	0,955	0,970	0,979	0,986	0,991	**0,991**
|S| IG&Acc2 (l)	**0,769**	0,890	0,917	0,953	**0,975**	**0,987**	0,990	**0,994**	**0,993**
|S| DF&Acc2 (l)	**0,781**	0,870	0,919	0,942	0,972	0,978	0,988	0,989	0,988
MAX	**0,825**	0,901	0,942	0,962	**0,977**	**0,987**	0,990	**0,994**	**0,993**
AVERAGE	**0,796**	0,888	0,923	**0,953**	**0,970**	**0,977**	**0,984**	**0,987**	**0,987**
Macro-F / Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (l)	0,712	0,894	0,936	0,948	0,963	0,968	0,968	0,969	0,981
|R| tf-idf&IG (l)	0,710	0,897	0,938	0,952	0,965	0,970	0,968	0,976	0,987
|R| tf-idf&DF (l)	**0,806**	0,878	0,899	0,935	0,957	0,957	0,965	0,967	0,972
|R| tf-idf&Acc2 (l)	**0,824**	0,890	0,932	0,947	0,965	0,962	0,972	0,969	0,981
|R| CHI&IG (l)	**0,806**	0,901	0,938	0,952	**0,974**	0,985	**0,991**	0,991	0,990
|R| CHI&DF (l)	**0,837**	0,896	0,935	0,949	0,963	0,969	0,974	0,977	0,982
|R| CHI&Acc2 (l)	**0,806**	0,899	0,933	**0,964**	**0,974**	0,982	**0,991**	0,991	**0,991**
|R| IG&DF (l)	**0,835**	0,897	0,938	0,948	0,966	0,970	0,976	0,982	0,985
|R| IG&Acc2 (l)	**0,766**	0,891	0,924	0,953	**0,974**	0,984	0,990	**0,992**	**0,991**
|R| DF&Acc2 (l)	**0,805**	0,873	0,899	0,943	0,964	0,968	0,978	0,978	0,982
MAX	**0,837**	0,901	0,938	**0,964**	**0,974**	0,985	**0,991**	**0,992**	**0,991**
AVERAGE	0,791	**0,891**	**0,927**	0,949	0,967	0,971	0,977	0,979	0,984

Table 7.5. In local policy, macro-averaged F-measures of the score and rank combinations for Classic3 dataset

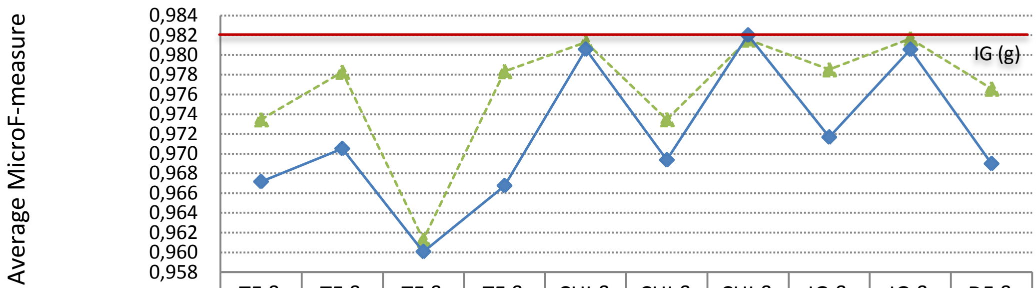
In Table 7.5, we see that the macro-averaged F-measure results of *IG & Acc2* score combination for 500, 1500 and 2000 keywords in local policy have the best results (98.7%, 99.4% and 99.3% respectively). As we analyzed previously, document classification without any feature selection achieves the highest performance of 99.4% and neither the existing metrics nor the score and rank combinations in global policy improved this result and the nearest value to this result among all methods is 99.3% in the Classic3 dataset. In the case of local policy, *IG & Acc2* score combination can reach at 99.4% macro-averaged F-measure with 1500 keywords.

In global policy *tf-idf & Acc2, CHI & IG, CHI & Acc2, IG & DF* and *IG & Acc2* combinations were more successful than others in homogenous datasets. On the other hand, *CHI & IG, CHI & Acc2* and *IG & Acc2* combinations have the best F-measure values in local policy. Rank combination of *CHI & IG* in global policy and score combination of *IG & Acc2* in local policy with large number of keywords between 200 and 2000 are notably better than the existing metrics in homogenous datasets. Moreover the rank combination of *tf-idf & Acc2* in global policy improves the performance of the existing metrics with a few number of keywords.

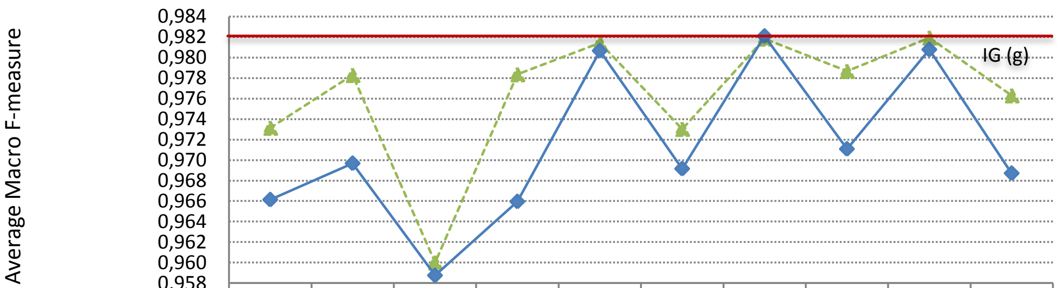
As we mentioned, the performance gap between the existing metrics and the score or rank combinations enlarges as the number of keywords decreases on homogenous datasets. In global policy, among the existing metrics Acc2 can achieve at most 73.6% micro- and 70.9% macro-averaged F-measures with 10 keywords whereas rank combination of *DF & Acc2* reaches 76.4% micro and macro-averaged F-measures. And in local policy, among the existing metrics again *Acc2* can only achieve 78.7% micro- and 76.1% macro-averaged F-measures while rank combination of CHI & *DF* reaches 84.4% micro- and 83.7% macro-averaged F-measures with 10 keywords. Beside this result, another apparent result is that the F-measure values with 10 keywords are improved by almost all score and rank combinations.

In local policy, among the individual feature selection metrics *IG* has the highest average of micro- and macro-averaged F-measure value of 98.2%. Only *CHI & Acc2* and *IG & Acc2* combinations reach this accuracy. Figure 4 shows the averages of micro- and macro-averaged F-measure values for the Classic3 dataset.

Score and rank combinations improve the accuracy in both policies. When we look at Tables 7.4 and 7.5, we see that score combination is slightly more successful than rank combination in homogenous datasets in local policy whereas the performance of rank combination is better than score combination in global policy as we observed in Table 7.2 and 3. This is an important result. As a result, rank combination outperforms score combination in case of global policy and rank combination is outperformed by score combination in case of local policy on homogenous datasets. In Figure 5, we see that score combinations in local policy is the best method for classification on homogenous datasets.



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (l)	0,973	0,978	0,961	0,978	0,981	0,973	0,982	0,979	0,982	0,977
Rank Com. (l)	0,967	0,971	0,960	0,967	0,981	0,969	0,982	0,972	0,981	0,969

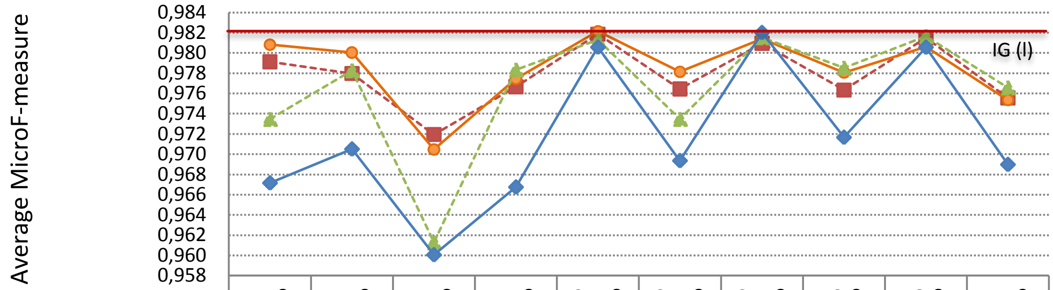


	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (l)	0,973	0,978	0,960	0,978	0,981	0,973	0,982	0,979	0,982	0,976
Rank Com. (l)	0,966	0,970	0,959	0,966	0,981	0,969	0,982	0,971	0,981	0,969

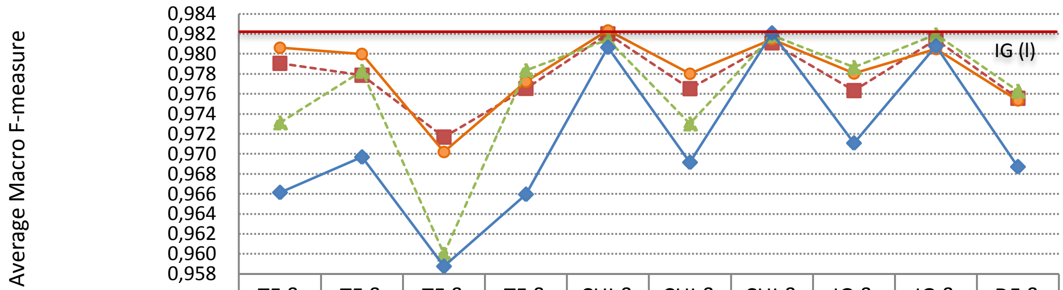
	tf-idf (l)	CHI (l)	IG (l)	DF (l)	Acc2 (l)
Micro	0,962	0,981	0,982	0,959	0,981
Macro	0,961	0,981	0,982	0,957	0,982

Figure 7.4. In local policy, comparison of score and rank combinations on the Classic3 dataset

In addition, one of the striking results in these experiments is the similar behavior of the two combinations both in global and local policies. It should be noted that if any feature selection pair is successful by combining their scores, it is also successful by combining their ranks.



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,979	0,978	0,972	0,977	0,982	0,976	0,981	0,976	0,981	0,976
Rank Com. (g)	0,981	0,980	0,970	0,977	0,982	0,978	0,981	0,978	0,981	0,975
Score Com. (l)	0,973	0,978	0,961	0,978	0,981	0,973	0,982	0,979	0,982	0,977
Rank Com. (l)	0,967	0,971	0,960	0,967	0,981	0,969	0,982	0,972	0,981	0,969

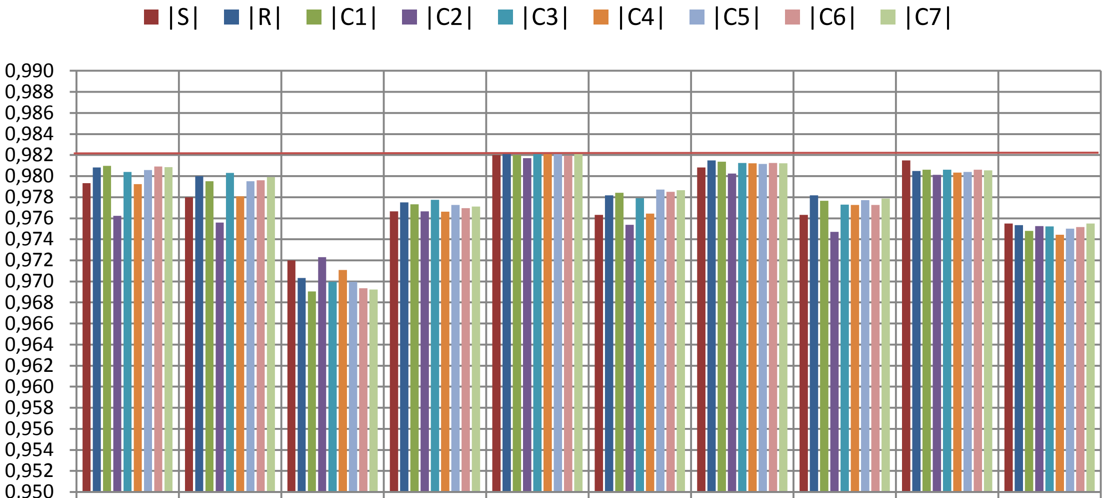


	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,979	0,978	0,972	0,977	0,982	0,977	0,981	0,976	0,981	0,976
Rank Com. (g)	0,981	0,980	0,970	0,977	0,982	0,978	0,981	0,978	0,980	0,975
Score Com. (l)	0,973	0,978	0,960	0,978	0,981	0,973	0,982	0,979	0,982	0,976
Rank Com. (l)	0,966	0,970	0,959	0,966	0,981	0,969	0,982	0,971	0,981	0,969

Figure 7.5: Comparison of score and rank combinations in global and local policy on the Classic3 dataset

7.1.1.4. Analysis of the Proposed Combinations

The success of the combinations motivated us to propose different combination methods in order to improve the performance of the classifier. In this section we analyze our seven proposed combination methods for the Classic3 dataset. In Figure 7.6 and 7.5 we can see the average of the F-measure values of the combination methods in global policy.



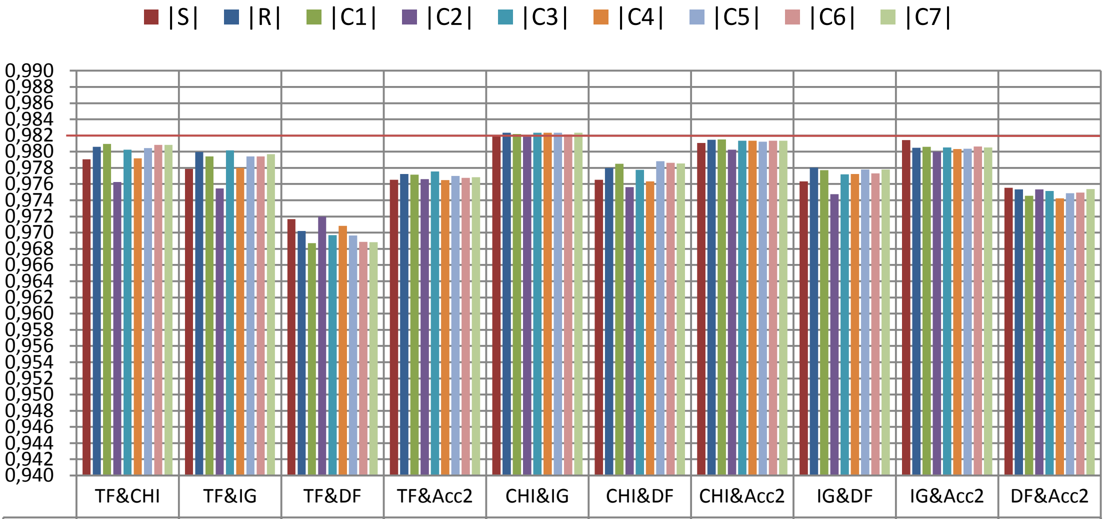
	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,979	0,978	0,972	0,977	0,982	0,976	0,981	0,976	0,982	0,976
|R|	0,981	0,980	0,970	0,978	0,982	0,978	0,982	0,978	0,981	0,975
|C1|	0,981	0,980	0,969	0,977	0,982	0,978	0,981	0,978	0,981	0,975
|C2|	0,976	0,976	0,972	0,977	0,982	0,975	0,980	0,975	0,980	0,975
|C3|	0,980	0,980	0,970	0,978	0,982	0,978	0,981	0,977	0,981	0,975
|C4|	0,979	0,978	0,971	0,977	0,982	0,976	0,981	0,977	0,980	0,974
|C5|	0,981	0,980	0,970	0,977	0,982	0,979	0,981	0,978	0,980	0,975
|C6|	0,981	0,980	0,969	0,977	0,982	0,978	0,981	0,977	0,981	0,975
|C7|	0,981	0,980	0,969	0,977	0,982	0,979	0,981	0,978	0,981	0,975

Figure 7.6. In global policy, averages of the micro-averaged F-measures of all combinations for Classic3 dataset

	10	30	50	100	200	500	1000	1500	2000
Combination |S|	0,760	0,875	0,921	0,955	0,975	0,989	0,991	0,992	0,993
Combination |R|	0,764	0,879	0,924	0,956	0,976	0,989	0,991	0,992	0,993
New Method 1	0,764	0,875	0,928	0,955	0,976	0,989	0,991	0,992	0,992
New Method 2	0,760	0,878	0,922	0,955	0,974	0,989	0,991	0,992	0,993
New Method 3	0,764	0,882	0,924	0,956	0,976	0,989	0,991	0,992	0,993
New Method 4	0,764	0,891	0,924	0,954	0,976	0,989	0,991	0,992	0,992
New Method 5	0,764	0,891	0,930	0,955	0,976	0,990	0,991	0,992	0,992
New Method 6	0,764	0,891	0,923	0,955	0,976	0,989	0,991	0,992	0,993
New Method 7	0,764	0,891	0,930	0,955	0,976	0,989	0,992	0,992	0,993

Table 7.6. In global policy, maximum micro-averaged F-measures of all combinations for Classic3 dataset

We begin with the experiments in global policy and then we analyze these combinations in the case of local policy. One of the main observations is that the successes of the proposed combinations are very close to each other in the homogenous dataset thus it is hard to say which one outperforms the others.



	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,979	0,978	0,972	0,977	0,982	0,977	0,981	0,976	0,981	0,976
|R|	0,981	0,980	0,970	0,977	0,982	0,978	0,981	0,978	0,980	0,975
|C1|	0,981	0,979	0,969	0,977	0,982	0,979	0,982	0,978	0,981	0,975
|C2|	0,976	0,975	0,972	0,977	0,982	0,976	0,980	0,975	0,980	0,975
|C3|	0,980	0,980	0,970	0,978	0,982	0,978	0,981	0,977	0,981	0,975
|C4|	0,979	0,978	0,971	0,976	0,982	0,976	0,981	0,977	0,980	0,974
|C5|	0,980	0,979	0,970	0,977	0,982	0,979	0,981	0,978	0,980	0,975
|C6|	0,981	0,979	0,969	0,977	0,982	0,979	0,981	0,977	0,981	0,975
|C7|	0,981	0,980	0,969	0,977	0,982	0,979	0,981	0,978	0,981	0,975

Figure 7.7. In global policy, averages of the macro-averaged F-measures of all combinations for Classic3 dataset

	10	30	50	100	200	500	1000	1500	2000
Combination |S|	0,759	0,873	0,920	0,954	0,975	0,990	0,991	0,993	0,993
Combination |R|	0,764	0,877	0,923	0,956	0,976	0,989	0,992	0,992	0,993
New Method 1	0,764	0,873	0,926	0,955	0,976	0,990	0,991	0,992	0,993
New Method 2	0,759	0,873	0,921	0,954	0,975	0,990	0,991	0,993	0,994
New Method 3	0,764	0,879	0,923	0,956	0,976	0,990	0,992	0,992	0,993
New Method 4	0,764	0,889	0,923	0,954	0,976	0,990	0,991	0,992	0,993
New Method 5	0,764	0,889	0,928	0,955	0,976	0,990	0,991	0,993	0,993
New Method 6	0,764	0,889	0,922	0,955	0,976	0,990	0,991	0,992	0,993
New Method 7	0,764	0,889	0,928	0,955	0,976	0,989	0,992	0,993	0,993

Table 7.7. In global policy, maximum macro-averaged F-measures of all combinations for Classic3 dataset

Tables 7.8 and 7.9 show the micro- and macro-averaged F-measures, respectively, for all the proposed methods for the Classic3 dataset. Before comparing the performance of each proposed method, we determine which combinations are the best. By all of 10 possible 2-combinations, it seems that *tf-idf & CHI, tf-idf & Acc2, CHI & IG, CHI & Acc2* and *IG & Acc2* 2-combinations are more successful than the other combinations. During the test phase, we obtain the best results from these combinations in both micro- and macro-averaged F-measure on the homogenous dataset.

In the previous section, we concluded that using rank combination for feature selection is the best method for classification in the case of global policy. Among the proposed methods, C2 improves the performance of the rank combination with a high number of keyword. As we knew that among the individual metrics the best performances (99.2% micro- and 99.3% macro-averaged F-measure) were achieved *CHI* with 1500 keywords. This performance was improved by several score and rank combinations with 99.3% and 99.3%. Furthermore that performance (99.3% micro- and 99.4% macro-averaged F-measure) is also improved by C2 of *tf-idf & CHI* with 2000 keywords. C3, C5 and C7 are also slightly better than the score and rank combinations in general when we compare their highest values and average of F-measures.

In addition to this, when we compare the F-measure values with a large number of keywords, the performance of each proposed method is almost the same. The success of classification with a few number of keywords is slightly different from each others. Therefore, this criterion helps us to determine which methods are more successful than the others. But we need to see the behavior of each combination on the other datasets in order to indicate the best one.

Table 7.8. Micro F-measure results of the proposed combinations in global policy for Classic3 dataset

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,701	**0,873**	0,901	0,937	0,956	0,981	0,988	**0,992**	**0,992**
CHI (g)		0,732	0,848	0,89	**0,956**	0,956	**0,989**	**0,991**	**0,992**	**0,992**
IG (g)		0,702	0,848	0,886	**0,956**	**0,974**	0,988	**0,991**	0,990	**0,992**
DF (g)		0,622	0,800	0,833	0,894	0,943	0,970	0,986	**0,992**	**0,992**
Acc2 (g)		**0,736**	0,867	**0,916**	0,944	0,967	0,984	0,988	0,989	0,991
	MAX	0,736	0,873	0,916	0,956	0,974	0,989	0,991	0,992	0,992
Micro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	0,760	0,875	0,921	0,955	0,975	0,989	0,991	0,992	0,993
	AVERAGE	0,732	0,849	0,903	0,946	0,964	0,985	0,989	0,991	0,992
Rank Combination	MAX	0,764	0,879	0,924	0,956	0,976	0,989	0,991	0,992	0,993
	AVERAGE	0,722	0,851	0,907	0,946	0,966	0,986	0,99	0,991	0,992
Micro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (g)		0,721	0,844	0,928	0,952	0,972	0,989	0,989	0,991	0,992
|C1| tf-idf&IG (g)		0,699	0,837	0,925	0,947	0,969	0,989	0,989	0,990	0,992
|C1| tf-idf&DF (g)		0,724	0,813	0,874	0,920	0,950	0,974	0,988	0,991	0,992
|C1| tf-idf&Acc2 (g)		0,760	0,875	0,917	0,945	0,962	0,984	0,990	0,991	0,992
|C1| CHI&IG (g)		0,732	0,853	0,890	0,955	0,976	0,989	0,991	0,991	0,991
|C1| CHI&DF (g)		0,763	0,866	0,909	0,946	0,967	0,985	0,989	0,992	0,992
|C1| CHI&Acc2 (g)		0,732	0,840	0,919	0,955	0,972	0,988	0,989	0,992	0,992
|C1| IG&DF (g)		0,658	0,868	0,912	0,944	0,964	0,985	0,990	0,991	0,992
|C1| IG&Acc2 (g)		0,702	0,840	0,922	0,950	0,971	0,988	0,990	0,992	0,992
|C1| DF&Acc2 (g)		0,764	0,859	0,879	0,937	0,960	0,982	0,989	0,989	0,992
	MAX	0,764	0,875	0,928	0,955	0,976	0,989	0,991	0,992	0,992
	AVERAGE	0,725	0,849	0,907	0,945	0,966	0,985	0,989	0,991	0,992
Micro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (g)		0,759	0,878	0,916	0,940	0,959	0,985	0,989	0,991	0,993
|C2| tf-idf&IG (g)		0,699	0,878	0,916	0,940	0,957	0,985	0,988	0,991	0,992
|C2| tf-idf&DF (g)		0,724	0,824	0,879	0,927	0,956	0,978	0,989	0,991	0,992
|C2| tf-idf&Acc2 (g)		0,760	0,875	0,909	0,944	0,961	0,984	0,989	0,991	0,992
|C2| CHI&IG (g)		0,732	0,853	0,890	0,955	0,974	0,989	0,990	0,991	0,992
|C2| CHI&DF (g)		0,723	0,838	0,875	0,939	0,960	0,981	0,990	0,991	0,992
|C2| CHI&Acc2 (g)		0,732	0,870	0,922	0,951	0,968	0,989	0,989	0,992	0,992
|C2| IG&DF (g)		0,723	0,841	0,897	0,935	0,958	0,981	0,991	0,991	0,992
|C2| IG&Acc2 (g)		0,702	0,876	0,921	0,949	0,971	0,987	0,990	0,992	0,992
|C2| DF&Acc2 (g)		0,723	0,835	0,881	0,937	0,960	0,981	0,990	0,990	0,993
	MAX	0,760	0,878	0,922	0,955	0,974	0,989	0,991	0,992	0,993
	AVERAGE	0,728	0,857	0,900	0,942	0,962	0,984	0,989	0,991	0,992
Micro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (g)		0,694	0,844	0,922	0,951	0,969	0,988	0,990	0,992	0,992
|C3| tf-idf&IG (g)		0,693	0,843	0,923	0,951	0,969	0,989	0,991	0,990	0,992
|C3| tf-idf&DF (g)		0,736	0,822	0,872	0,925	0,950	0,977	0,986	0,989	0,992
|C3| tf-idf&Acc2 (g)		0,760	0,882	0,916	0,944	0,963	0,985	0,991	0,991	0,992
|C3| CHI&IG (g)		0,732	0,853	0,890	0,954	0,976	0,989	0,991	0,991	0,992
|C3| CHI&DF (g)		0,675	0,861	0,910	0,946	0,963	0,986	0,989	0,992	0,992
|C3| CHI&Acc2 (g)		0,732	0,844	0,924	0,956	0,971	0,987	0,989	0,992	0,993
|C3| IG&DF (g)		0,676	0,862	0,914	0,940	0,962	0,988	0,991	0,991	0,992
|C3| IG&Acc2 (g)		0,732	0,836	0,920	0,950	0,971	0,989	0,989	0,992	0,992
|C3| DF&Acc2 (g)		0,764	0,853	0,876	0,938	0,960	0,983	0,988	0,990	0,992
	MAX	0,764	0,882	0,924	0,956	0,976	0,989	0,991	0,992	0,993
	AVERAGE	0,719	0,850	0,907	0,946	0,965	0,986	0,990	0,991	0,992

Micro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (g)		0,699	0,844	0,920	0,949	0,968	0,985	0,989	0,992	0,992
|C4| tf-idf&IG (g)		0,699	0,837	0,924	0,947	0,965	0,983	0,989	0,992	0,992
|C4| tf-idf&DF (g)		0,736	0,823	0,871	0,927	0,954	0,977	0,987	0,991	0,991
|C4| tf-idf&Acc2 (g)		0,760	0,891	0,917	0,944	0,960	0,984	0,989	0,991	0,992
|C4| CHI&IG (g)		0,732	0,853	0,890	0,954	0,976	0,989	0,991	0,991	0,992
|C4| CHI&DF (g)		0,676	0,866	0,907	0,939	0,965	0,984	0,990	0,990	0,991
|C4| CHI&Acc2 (g)		0,732	0,840	0,919	0,953	0,972	0,989	0,989	0,992	0,992
|C4| IG&DF (g)		0,666	0,866	0,910	0,941	0,965	0,985	0,990	0,991	0,992
|C4| IG&Acc2 (g)		0,732	0,836	0,922	0,951	0,971	0,987	0,989	0,992	0,992
|C4| DF&Acc2 (g)		0,764	0,856	0,880	0,936	0,959	0,981	0,989	0,989	0,992
	MAX	0,764	0,891	0,924	0,954	0,976	0,989	0,991	0,992	0,992
	AVERAGE	0,720	0,851	0,906	0,944	0,965	0,984	0,989	0,991	0,992
Micro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (g)		0,694	0,844	0,930	0,950	0,969	0,990	0,990	0,992	0,992
|C5| tf-idf&IG (g)		0,684	0,844	0,925	0,948	0,968	0,989	0,989	0,991	0,992
|C5| tf-idf&DF (g)		0,736	0,822	0,873	0,925	0,950	0,977	0,986	0,990	0,992
|C5| tf-idf&Acc2 (g)		0,760	0,891	0,916	0,944	0,961	0,985	0,990	0,991	0,992
|C5| CHI&IG (g)		0,732	0,853	0,890	0,954	0,976	0,989	0,991	0,991	0,992
|C5| CHI&DF (g)		0,676	0,865	0,911	0,946	0,969	0,984	0,990	0,992	0,992
|C5| CHI&Acc2 (g)		0,732	0,844	0,919	0,955	0,971	0,987	0,989	0,992	0,992
|C5| IG&DF (g)		0,676	0,859	0,914	0,942	0,966	0,986	0,990	0,991	0,992
|C5| IG&Acc2 (g)		0,732	0,836	0,922	0,950	0,971	0,988	0,989	0,992	0,992
|C5| DF&Acc2 (g)		0,764	0,854	0,882	0,936	0,960	0,983	0,989	0,989	0,992
	MAX	0,764	0,891	0,930	0,955	0,976	0,990	0,991	0,992	0,992
	AVERAGE	0,719	0,851	0,908	0,945	0,966	0,986	0,989	0,991	0,992
Micro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (g)		0,699	0,844	0,922	0,950	0,973	0,988	0,990	0,992	0,992
|C6| tf-idf&IG (g)		0,699	0,837	0,923	0,946	0,972	0,988	0,989	0,991	0,992
|C6| tf-idf&DF (g)		0,736	0,823	0,873	0,922	0,950	0,973	0,988	0,991	0,992
|C6| tf-idf&Acc2 (g)		0,760	0,891	0,917	0,943	0,962	0,984	0,990	0,991	0,992
|C6| CHI&IG (g)		0,732	0,853	0,890	0,954	0,976	0,989	0,991	0,991	0,991
|C6| CHI&DF (g)		0,676	0,865	0,915	0,946	0,969	0,984	0,989	0,991	0,992
|C6| CHI&Acc2 (g)		0,732	0,840	0,919	0,955	0,971	0,988	0,989	0,992	0,992
|C6| IG&DF (g)		0,666	0,866	0,914	0,940	0,966	0,986	0,990	0,991	0,992
|C6| IG&Acc2 (g)		0,732	0,836	0,922	0,950	0,971	0,989	0,990	0,992	0,992
|C6| DF&Acc2 (g)		0,764	0,846	0,880	0,936	0,960	0,984	0,989	0,989	0,993
	MAX	0,764	0,891	0,923	0,955	0,976	0,989	0,991	0,992	0,993
	AVERAGE	0,720	0,850	0,907	0,944	0,967	0,985	0,990	0,991	0,992
Micro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (g)		0,694	0,844	0,930	0,952	0,973	0,987	0,990	0,991	0,992
|C7| tf-idf&IG (g)		0,699	0,844	0,924	0,947	0,971	0,989	0,990	0,991	0,992
|C7| tf-idf&DF (g)		0,736	0,822	0,873	0,921	0,950	0,974	0,987	0,991	0,992
|C7| tf-idf&Acc2 (g)		0,760	0,891	0,916	0,943	0,962	0,983	0,992	0,991	0,992
|C7| CHI&IG (g)		0,732	0,853	0,890	0,954	0,976	0,989	0,991	0,992	0,992
|C7| CHI&DF (g)		0,676	0,865	0,913	0,948	0,964	0,986	0,990	0,991	0,992
|C7| CHI&Acc2 (g)		0,732	0,844	0,919	0,955	0,971	0,987	0,989	0,992	0,992
|C7| IG&DF (g)		0,676	0,859	0,915	0,943	0,964	0,986	0,991	0,990	0,993
|C7| IG&Acc2 (g)		0,732	0,836	0,922	0,950	0,972	0,989	0,989	0,992	0,992
|C7| DF&Acc2 (g)		0,764	0,854	0,882	0,938	0,960	0,983	0,990	0,990	0,992
	MAX	0,764	0,891	0,930	0,955	0,976	0,989	0,992	0,992	0,993
	AVERAGE	0,720	0,851	0,908	0,945	0,966	0,985	0,990	0,991	0,992

Table 7.9. Macro F-measure results of the proposed combinations in global policy for Classic3 dataset

Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,665	0,871	0,898	0,936	0,953	0,980	0,989	0,992	0,992
CHI (g)		0,709	0,821	0,870	0,956	0,956	0,990	0,991	0,993	0,992
IG (g)		0,665	0,811	0,863	0,955	0,975	0,988	0,991	0,990	0,992
DF (g)		0,623	0,798	0,831	0,893	0,941	0,970	0,987	0,992	0,992
Acc2 (g)		0,690	0,865	0,914	0,944	0,967	0,985	0,988	0,990	0,991
	MAX	0,709	0,871	0,914	0,956	0,975	0,99	0,991	0,993	0,992
Macro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	0,759	0,873	0,920	0,954	0,975	0,990	0,991	0,993	0,993
	AVERAGE	0,705	0,836	0,899	0,945	0,963	0,985	0,990	0,992	0,992
Rank Combination	MAX	0,764	0,877	0,923	0,956	0,976	0,989	0,992	0,992	0,993
	AVERAGE	0,653	0,830	0,903	0,945	0,965	0,986	0,990	0,991	0,992
Macro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (g)		0,551	0,801	0,926	0,951	0,972	0,989	0,990	0,991	0,992
|C1| tf-idf&IG (g)		0,556	0,795	0,923	0,945	0,969	0,989	0,990	0,991	0,993
|C1| tf-idf&DF (g)		0,724	0,811	0,872	0,919	0,949	0,973	0,988	0,991	0,992
|C1| tf-idf&Acc2 (g)		0,759	0,873	0,915	0,944	0,960	0,985	0,990	0,991	0,992
|C1| CHI&IG (g)		0,709	0,825	0,870	0,954	0,976	0,990	0,991	0,991	0,991
|C1| CHI&DF (g)		0,762	0,865	0,907	0,945	0,967	0,986	0,989	0,992	0,992
|C1| CHI&Acc2 (g)		0,709	0,805	0,915	0,955	0,972	0,988	0,989	0,992	0,992
|C1| IG&DF (g)		0,527	0,866	0,910	0,944	0,963	0,985	0,991	0,991	0,992
|C1| IG&Acc2 (g)		0,665	0,798	0,920	0,949	0,971	0,989	0,991	0,992	0,993
|C1| DF&Acc2 (g)		0,764	0,858	0,877	0,936	0,958	0,982	0,989	0,990	0,992
	MAX	0,764	0,873	0,926	0,955	0,976	0,990	0,991	0,992	0,993
	AVERAGE	0,673	0,830	0,904	0,944	0,966	0,986	0,990	0,991	0,992
Macro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (g)		0,710	0,873	0,914	0,940	0,958	0,985	0,989	0,992	0,994
|C2| tf-idf&IG (g)		0,556	0,873	0,914	0,939	0,956	0,985	0,988	0,992	0,992
|C2| tf-idf&DF (g)		0,724	0,822	0,877	0,926	0,955	0,978	0,990	0,991	0,992
|C2| tf-idf&Acc2 (g)		0,759	0,873	0,907	0,943	0,960	0,984	0,990	0,992	0,992
|C2| CHI&IG (g)		0,709	0,825	0,870	0,954	0,975	0,990	0,990	0,991	0,992
|C2| CHI&DF (g)		0,725	0,836	0,874	0,938	0,960	0,981	0,991	0,991	0,992
|C2| CHI&Acc2 (g)		0,709	0,861	0,921	0,950	0,968	0,990	0,989	0,993	0,992
|C2| IG&DF (g)		0,725	0,839	0,896	0,935	0,958	0,981	0,991	0,991	0,993
|C2| IG&Acc2 (g)		0,665	0,871	0,920	0,948	0,970	0,987	0,990	0,993	0,992
|C2| DF&Acc2 (g)		0,725	0,834	0,879	0,936	0,960	0,981	0,990	0,991	0,994
	MAX	0,759	0,873	0,921	0,954	0,975	0,990	0,991	0,993	0,994
	AVERAGE	0,701	0,851	0,897	0,941	0,962	0,984	0,990	0,992	0,992
Macro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (g)		0,551	0,801	0,920	0,949	0,969	0,989	0,990	0,992	0,992
|C3| tf-idf&IG (g)		0,550	0,801	0,921	0,949	0,968	0,990	0,991	0,991	0,992
|C3| tf-idf&DF (g)		0,737	0,821	0,870	0,924	0,949	0,976	0,987	0,990	0,992
|C3| tf-idf&Acc2 (g)		0,759	0,879	0,914	0,943	0,962	0,985	0,992	0,991	0,993
|C3| CHI&IG (g)		0,709	0,825	0,870	0,954	0,976	0,989	0,991	0,991	0,992
|C3| CHI&DF (g)		0,538	0,860	0,908	0,945	0,962	0,987	0,990	0,992	0,992
|C3| CHI&Acc2 (g)		0,709	0,808	0,923	0,956	0,970	0,988	0,989	0,992	0,993
|C3| IG&DF (g)		0,540	0,861	0,913	0,939	0,961	0,988	0,991	0,992	0,993
|C3| IG&Acc2 (g)		0,709	0,795	0,918	0,949	0,971	0,989	0,990	0,992	0,992
|C3| DF&Acc2 (g)		0,764	0,851	0,874	0,937	0,958	0,983	0,989	0,991	0,993
	MAX	0,764	0,879	0,923	0,956	0,976	0,990	0,992	0,992	0,993
	AVERAGE	0,657	0,830	0,903	0,944	0,965	0,986	0,990	0,991	0,992

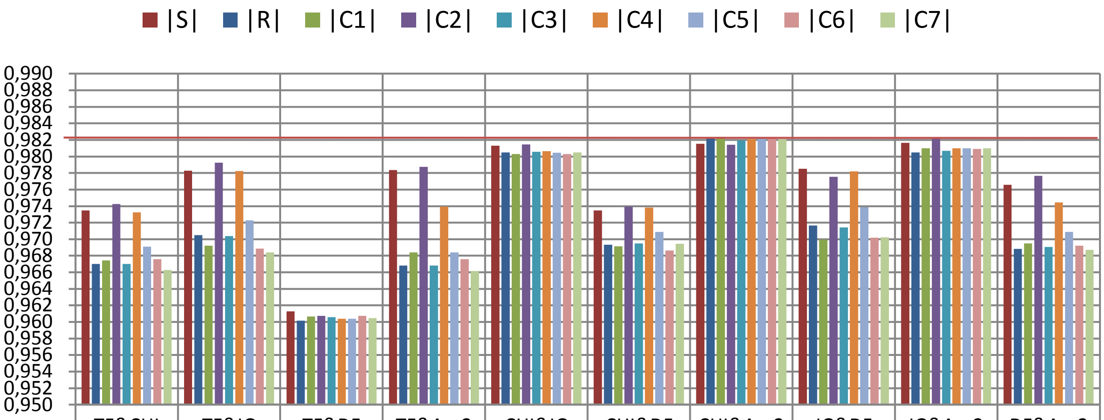
Macro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (g)		0,556	0,801	0,918	0,948	0,968	0,985	0,989	0,992	0,993
|C4| tf-idf&IG (g)		0,556	0,795	0,923	0,945	0,964	0,984	0,990	0,992	0,993
|C4| tf-idf&DF (g)		0,737	0,821	0,870	0,926	0,952	0,977	0,988	0,991	0,991
|C4| tf-idf&Acc2 (g)		0,759	0,889	0,915	0,943	0,959	0,984	0,990	0,991	0,992
|C4| CHI&IG (g)		0,709	0,825	0,870	0,954	0,976	0,989	0,991	0,991	0,992
|C4| CHI&DF (g)		0,540	0,865	0,905	0,938	0,965	0,984	0,990	0,990	0,991
|C4| CHI&Acc2 (g)		0,709	0,805	0,915	0,953	0,972	0,990	0,989	0,992	0,992
|C4| IG&DF (g)		0,531	0,865	0,908	0,941	0,964	0,985	0,990	0,991	0,992
|C4| IG&Acc2 (g)		0,709	0,795	0,920	0,950	0,971	0,988	0,990	0,992	0,992
|C4| DF&Acc2 (g)		0,764	0,855	0,878	0,935	0,958	0,981	0,989	0,990	0,992
	MAX	0,764	0,889	0,923	0,954	0,976	0,990	0,991	0,992	0,993
	AVERAGE	0,657	0,831	0,902	0,943	0,965	0,985	0,990	0,991	0,992

Macro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (g)		0,551	0,801	0,928	0,948	0,969	0,990	0,990	0,992	0,993
|C5| tf-idf&IG (g)		0,543	0,801	0,923	0,946	0,968	0,990	0,989	0,991	0,993
|C5| tf-idf&DF (g)		0,737	0,821	0,871	0,924	0,948	0,976	0,986	0,991	0,992
|C5| tf-idf&Acc2 (g)		0,759	0,889	0,914	0,943	0,959	0,986	0,990	0,991	0,993
|C5| CHI&IG (g)		0,709	0,825	0,870	0,954	0,976	0,989	0,991	0,991	0,992
|C5| CHI&DF (g)		0,540	0,863	0,910	0,945	0,968	0,985	0,990	0,992	0,992
|C5| CHI&Acc2 (g)		0,709	0,808	0,915	0,955	0,970	0,988	0,989	0,992	0,993
|C5| IG&DF (g)		0,540	0,858	0,912	0,941	0,965	0,987	0,990	0,991	0,993
|C5| IG&Acc2 (g)		0,709	0,795	0,920	0,949	0,971	0,989	0,989	0,993	0,992
|C5| DF&Acc2 (g)		0,764	0,852	0,881	0,935	0,959	0,983	0,990	0,990	0,993
	MAX	0,764	0,889	0,928	0,955	0,976	0,990	0,991	0,993	0,993
	AVERAGE	0,656	0,831	0,904	0,944	0,965	0,986	0,990	0,991	0,992

Macro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (g)		0,556	0,801	0,920	0,948	0,973	0,989	0,990	0,992	0,993
|C6| tf-idf&IG (g)		0,556	0,795	0,922	0,944	0,971	0,989	0,989	0,991	0,993
|C6| tf-idf&DF (g)		0,737	0,821	0,872	0,920	0,949	0,972	0,988	0,991	0,992
|C6| tf-idf&Acc2 (g)		0,759	0,889	0,915	0,941	0,960	0,984	0,991	0,991	0,992
|C6| CHI&IG (g)		0,709	0,825	0,870	0,954	0,976	0,990	0,991	0,991	0,991
|C6| CHI&DF (g)		0,540	0,863	0,913	0,945	0,969	0,985	0,990	0,991	0,992
|C6| CHI&Acc2 (g)		0,709	0,805	0,915	0,955	0,971	0,988	0,989	0,992	0,992
|C6| IG&DF (g)		0,531	0,865	0,912	0,939	0,965	0,986	0,990	0,991	0,992
|C6| IG&Acc2 (g)		0,709	0,795	0,920	0,949	0,971	0,989	0,990	0,992	0,992
|C6| DF&Acc2 (g)		0,764	0,845	0,878	0,935	0,959	0,983	0,989	0,990	0,993
	MAX	0,764	0,889	0,922	0,955	0,976	0,990	0,991	0,992	0,993
	AVERAGE	0,657	0,830	0,904	0,943	0,966	0,986	0,990	0,991	0,992

Macro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (g)		0,551	0,801	0,928	0,950	0,973	0,988	0,990	0,991	0,992
|C7| tf-idf&IG (g)		0,556	0,801	0,923	0,945	0,970	0,989	0,990	0,991	0,993
|C7| tf-idf&DF (g)		0,737	0,821	0,871	0,920	0,949	0,973	0,988	0,991	0,992
|C7| tf-idf&Acc2 (g)		0,759	0,889	0,914	0,942	0,960	0,983	0,992	0,991	0,993
|C7| CHI&IG (g)		0,709	0,825	0,870	0,954	0,976	0,989	0,991	0,992	0,992
|C7| CHI&DF (g)		0,540	0,863	0,911	0,947	0,964	0,987	0,990	0,992	0,992
|C7| CHI&Acc2 (g)		0,709	0,808	0,915	0,955	0,971	0,988	0,989	0,992	0,993
|C7| IG&DF (g)		0,540	0,858	0,913	0,942	0,963	0,986	0,992	0,991	0,993
|C7| IG&Acc2 (g)		0,709	0,795	0,920	0,949	0,971	0,989	0,989	0,993	0,992
|C7| DF&Acc2 (g)		0,764	0,852	0,881	0,938	0,959	0,983	0,991	0,990	0,992
	MAX	0,764	0,889	0,928	0,955	0,976	0,989	0,992	0,993	0,993
	AVERAGE	0,657	0,831	0,905	0,944	0,966	0,986	0,990	0,991	0,992

When we compared score and rank combinations in Classic3 dataset in local policy, we stated that the score combination improves the performance of the individual metrics and outperforms the rank combination of two features selection metrics. When evaluating the performance of the proposed methods, we see that none of the proposed combination outperforms the success of the score combination.

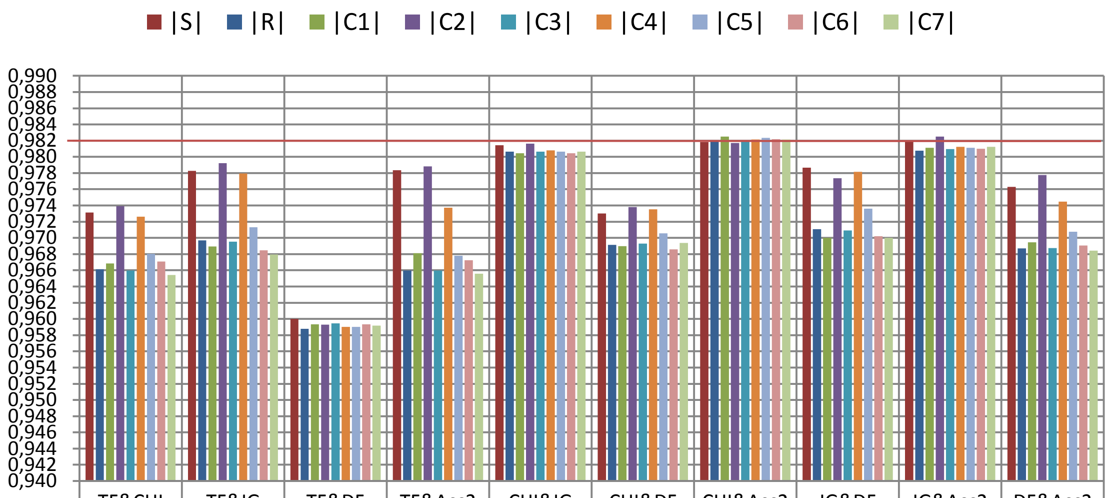


	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,973	0,978	0,961	0,978	0,981	0,973	0,982	0,979	0,982	0,977
|R|	0,967	0,971	0,960	0,967	0,981	0,969	0,982	0,972	0,981	0,969
|C1|	0,967	0,969	0,961	0,968	0,980	0,969	0,982	0,970	0,981	0,969
|C2|	0,974	0,979	0,961	0,979	0,981	0,974	0,981	0,978	0,982	0,978
|C3|	0,967	0,970	0,961	0,967	0,981	0,969	0,982	0,971	0,981	0,969
|C4|	0,973	0,978	0,960	0,974	0,981	0,974	0,982	0,978	0,981	0,974
|C5|	0,969	0,972	0,960	0,968	0,980	0,971	0,982	0,974	0,981	0,971
|C6|	0,968	0,969	0,961	0,968	0,980	0,969	0,982	0,970	0,981	0,969
|C7|	0,966	0,968	0,960	0,966	0,980	0,969	0,982	0,970	0,981	0,969

Figure 7.8. In local policy, averages of the micro-averaged F-measures of all combinations for Classic3 dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	0,839	0,910	0,944	0,962	0,977	0,986	0,990	0,993	0,993
Combination | R |	0,844	0,907	0,941	0,964	0,974	0,984	0,991	0,992	0,991
Combination |C1|	0,837	0,909	0,943	0,965	0,975	0,984	0,990	0,992	0,992
Combination |C2|	0,835	0,907	0,939	0,961	0,975	0,990	0,990	0,993	0,992
Combination |C3|	0,844	0,908	0,942	0,964	0,974	0,984	0,991	0,992	0,991
Combination |C4|	0,844	0,914	0,942	0,963	0,975	0,985	0,990	0,993	0,992
Combination |C5|	0,844	0,914	0,943	0,964	0,974	0,984	0,991	0,992	0,992
Combination |C6|	0,844	0,915	0,943	0,964	0,974	0,984	0,991	0,992	0,992
Combination |C7|	0,844	0,916	0,942	0,964	0,975	0,985	0,990	0,992	0,992

Table 7.10. In local policy, maximum micro-averaged F-measures of all combinations for Classic3 dataset



	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,973	0,978	0,960	0,978	0,981	0,973	0,982	0,979	0,982	0,976
|R|	0,966	0,970	0,959	0,966	0,981	0,969	0,982	0,971	0,981	0,969
|C1|	0,967	0,969	0,959	0,968	0,980	0,969	0,982	0,970	0,981	0,969
|C2|	0,974	0,979	0,959	0,979	0,982	0,974	0,982	0,977	0,982	0,978
|C3|	0,966	0,970	0,959	0,966	0,981	0,969	0,982	0,971	0,981	0,969
|C4|	0,973	0,978	0,959	0,974	0,981	0,974	0,982	0,978	0,981	0,974
|C5|	0,968	0,971	0,959	0,968	0,981	0,971	0,982	0,974	0,981	0,971
|C6|	0,967	0,968	0,959	0,967	0,980	0,969	0,982	0,970	0,981	0,969
|C7|	0,965	0,968	0,959	0,966	0,981	0,969	0,982	0,970	0,981	0,968

Figure 7.9. In local policy, averages of the macro-averaged F-measures of all combinations for Classic3 dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	0,825	0,901	0,942	0,962	0,977	0,987	0,990	0,994	0,993
Combination | R |	0,837	0,901	0,938	0,964	0,974	0,985	0,991	0,992	0,991
Combination |C1|	0,824	0,903	0,940	0,965	0,975	0,985	0,990	0,993	0,993
Combination |C2|	0,816	0,898	0,935	0,960	0,975	0,990	0,990	0,993	0,992
Combination |C3|	0,837	0,903	0,939	0,964	0,975	0,985	0,991	0,992	0,991
Combination |C4|	0,837	0,909	0,940	0,963	0,975	0,985	0,990	0,993	0,992
Combination |C5|	0,837	0,909	0,941	0,964	0,974	0,985	0,991	0,993	0,993
Combination |C6|	0,837	0,910	0,941	0,964	0,974	0,985	0,991	0,993	0,992
Combination |C7|	0,837	0,911	0,939	0,964	0,975	0,985	0,990	0,993	0,992

Table 7.11. In local policy, maximum macro-averaged F-measures of all combinations for Classic3 dataset

Tables 7.12 and 7.13 show the micro- and macro-averaged F-measure results, respectively, for all seven proposed combination methods in local policy for the Classic3 dataset.

Firstly, we determine which combinations are better than others. Among the 10 possible combinations of the two feature selection metrics, the best combinations are *CHI & IG, CHI & Acc2* and *IG & Acc2*.

Although none of the proposed combination outperforms the success of the score combination in general, C5, C6 and C7 are more successful than score combination with a few number of keywords, less than 200 keywords. In addition among the proposed methods, C2 is the most successful one with a high number of keyword.

Table 7.12. Micro F-measure results of the proposed combinations in local policy for Classic3 dataset

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (l)		0,653	0,895	0,939	0,951	0,959	0,960	0,964	0,965	0,971
CHI (l)		0,638	0,915	0,947	0,963	0,974	0,981	0,987	0,989	0,990
IG (l)		0,735	0,896	0,918	0,958	0,973	0,986	0,989	0,992	0,991
DF (l)		0,745	0,865	0,883	0,917	0,949	0,964	0,973	0,973	0,978
Acc2 (l)		0,787	0,880	0,926	0,958	0,972	0,985	0,991	0,991	0,991
	MAX	0,787	0,915	0,947	0,963	0,974	0,986	0,991	0,992	0,991
Micro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	0,839	0,910	0,944	0,962	0,977	0,986	0,990	0,993	0,993
	AVERAGE	0,804	0,897	0,928	0,954	0,970	0,977	0,984	0,987	0,987
Rank Combination	MAX	0,844	0,907	0,941	0,964	0,974	0,984	0,991	0,992	0,991
	AVERAGE	0,788	0,900	0,932	0,950	0,967	0,972	0,978	0,980	0,985
Micro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (l)		0,648	0,903	0,940	0,952	0,967	0,968	0,971	0,970	0,978
|C1| tf-idf&IG (l)		0,652	0,901	0,943	0,954	0,967	0,970	0,972	0,973	0,979
|C1| tf-idf&DF (l)		0,827	0,888	0,915	0,938	0,955	0,960	0,969	0,969	0,973
|C1| tf-idf&Acc2 (l)		0,837	0,903	0,933	0,952	0,963	0,968	0,973	0,977	0,977
|C1| CHI&IG (l)		0,832	0,906	0,942	0,954	0,974	0,984	0,988	0,991	0,990
|C1| CHI&DF (l)		0,832	0,909	0,938	0,951	0,968	0,969	0,977	0,974	0,976
|C1| CHI&Acc2 (l)		0,828	0,909	0,942	0,965	0,975	0,981	0,990	0,991	0,992
|C1| IG&DF (l)		0,834	0,901	0,929	0,953	0,966	0,972	0,974	0,976	0,979
|C1| IG&Acc2 (l)		0,785	0,900	0,919	0,954	0,973	0,984	0,989	0,992	0,992
|C1| DF&Acc2 (l)		0,817	0,877	0,910	0,948	0,965	0,970	0,979	0,977	0,978
	MAX	0,837	0,909	0,943	0,965	0,975	0,984	0,990	0,992	0,992
	AVERAGE	0,789	0,900	0,931	0,952	0,967	0,973	0,978	0,979	0,981
Micro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (l)		0,645	0,906	0,939	0,957	0,967	0,974	0,980	0,983	0,984
|C2| tf-idf&IG (l)		0,833	0,907	0,933	0,961	0,974	0,980	0,986	0,988	0,988
|C2| tf-idf&DF (l)		0,809	0,877	0,911	0,940	0,954	0,959	0,969	0,968	0,975
|C2| tf-idf&Acc2 (l)		0,835	0,893	0,935	0,958	0,973	0,978	0,986	0,988	0,989
|C2| CHI&IG (l)		0,807	0,904	0,934	0,958	0,975	0,984	0,990	0,991	0,991
|C2| CHI&DF (l)		0,810	0,884	0,907	0,946	0,967	0,973	0,982	0,988	0,988
|C2| CHI&Acc2 (l)		0,824	0,899	0,935	0,960	0,974	0,982	0,990	0,992	0,991
|C2| IG&DF (l)		0,798	0,887	0,920	0,950	0,967	0,981	0,986	0,991	0,990
|C2| IG&Acc2 (l)		0,785	0,896	0,920	0,955	0,974	0,990	0,990	0,993	0,992
|C2| DF&Acc2 (l)		0,773	0,878	0,927	0,951	0,970	0,977	0,988	0,990	0,990
	MAX	0,835	0,907	0,939	0,961	0,975	0,990	0,990	0,993	0,992
	AVERAGE	0,792	0,893	0,926	0,953	0,969	0,978	0,985	0,987	0,988
Micro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (l)		0,647	0,908	0,939	0,950	0,963	0,968	0,969	0,970	0,982
|C3| tf-idf&IG (l)		0,645	0,904	0,942	0,954	0,965	0,970	0,969	0,978	0,987
|C3| tf-idf&DF (l)		0,820	0,886	0,910	0,940	0,958	0,958	0,967	0,968	0,973
|C3| tf-idf&Acc2 (l)		0,837	0,899	0,935	0,950	0,965	0,962	0,973	0,970	0,981
|C3| CHI&IG (l)		0,822	0,905	0,941	0,953	0,974	0,984	0,991	0,991	0,990
|C3| CHI&DF (l)		0,842	0,907	0,937	0,950	0,965	0,969	0,974	0,976	0,983
|C3| CHI&Acc2 (l)		0,834	0,907	0,940	0,964	0,973	0,982	0,991	0,991	0,991
|C3| IG&DF (l)		0,844	0,905	0,940	0,949	0,965	0,971	0,975	0,983	0,986
|C3| IG&Acc2 (l)		0,785	0,899	0,925	0,953	0,974	0,984	0,990	0,992	0,991
|C3| DF&Acc2 (l)		0,816	0,883	0,909	0,944	0,966	0,967	0,978	0,977	0,982
	MAX	0,844	0,908	0,942	0,964	0,974	0,984	0,991	0,992	0,991
	AVERAGE	0,789	0,900	0,932	0,951	0,967	0,972	0,978	0,980	0,984

Micro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (I)		0,648	0,902	0,939	0,955	0,967	0,972	0,980	0,981	0,985
|C4| tf-idf&IG (I)		0,647	0,899	0,942	0,958	0,972	0,980	0,983	0,988	0,989
|C4| tf-idf&DF (I)		0,823	0,888	0,914	0,940	0,954	0,959	0,968	0,968	0,973
|C4| tf-idf&Acc2 (I)		0,837	0,899	0,934	0,955	0,964	0,974	0,981	0,983	0,986
|C4| CHI&IG (I)		0,824	0,906	0,942	0,955	0,975	0,984	0,989	0,991	0,991
|C4| CHI&DF (I)		0,844	0,910	0,940	0,951	0,967	0,974	0,981	0,985	0,985
|C4| CHI&Acc2 (I)		0,839	0,914	0,940	0,963	0,973	0,984	0,990	0,991	0,992
|C4| IG&DF (I)		0,844	0,901	0,927	0,955	0,973	0,976	0,986	0,989	0,991
|C4| IG&Acc2 (I)		0,785	0,900	0,919	0,954	0,973	0,985	0,990	0,993	0,992
|C4| DF&Acc2 (I)		0,816	0,877	0,911	0,949	0,967	0,975	0,982	0,986	0,989
	MAX	0,844	0,914	0,942	0,963	0,975	0,985	0,990	0,993	0,992
	AVERAGE	0,791	0,900	0,931	0,953	0,968	0,976	0,983	0,986	0,987
Micro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (I)		0,648	0,903	0,939	0,951	0,964	0,968	0,973	0,976	0,983
|C5| tf-idf&IG (I)		0,647	0,902	0,943	0,951	0,966	0,970	0,978	0,981	0,988
|C5| tf-idf&DF (I)		0,829	0,886	0,910	0,940	0,955	0,958	0,968	0,968	0,972
|C5| tf-idf&Acc2 (I)		0,837	0,898	0,936	0,948	0,963	0,966	0,973	0,978	0,982
|C5| CHI&IG (I)		0,822	0,905	0,942	0,954	0,974	0,984	0,989	0,990	0,990
|C5| CHI&DF (I)		0,844	0,910	0,938	0,949	0,966	0,971	0,975	0,983	0,982
|C5| CHI&Acc2 (I)		0,838	0,914	0,940	0,964	0,973	0,982	0,991	0,992	0,992
|C5| IG&DF (I)		0,844	0,908	0,926	0,954	0,965	0,973	0,980	0,984	0,988
|C5| IG&Acc2 (I)		0,785	0,899	0,926	0,955	0,973	0,984	0,989	0,992	0,992
|C5| DF&Acc2 (I)		0,816	0,883	0,910	0,946	0,965	0,972	0,978	0,982	0,983
	MAX	0,844	0,914	0,943	0,964	0,974	0,984	0,991	0,992	0,992
	AVERAGE	0,791	0,901	0,931	0,951	0,966	0,973	0,979	0,983	0,985
Micro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (I)		0,648	0,903	0,939	0,953	0,966	0,968	0,970	0,972	0,978
|C6| tf-idf&IG (I)		0,647	0,902	0,943	0,953	0,966	0,970	0,971	0,976	0,978
|C6| tf-idf&DF (I)		0,823	0,887	0,912	0,939	0,955	0,960	0,970	0,968	0,972
|C6| tf-idf&Acc2 (I)		0,837	0,899	0,936	0,948	0,963	0,966	0,975	0,976	0,977
|C6| CHI&IG (I)		0,822	0,912	0,942	0,954	0,974	0,984	0,988	0,991	0,991
|C6| CHI&DF (I)		0,844	0,910	0,940	0,953	0,967	0,969	0,973	0,974	0,976
|C6| CHI&Acc2 (I)		0,839	0,915	0,939	0,964	0,974	0,981	0,991	0,991	0,992
|C6| IG&DF (I)		0,844	0,907	0,926	0,953	0,965	0,973	0,976	0,976	0,979
|C6| IG&Acc2 (I)		0,785	0,900	0,919	0,954	0,973	0,984	0,989	0,992	0,992
|C6| DF&Acc2 (I)		0,817	0,884	0,908	0,946	0,965	0,970	0,980	0,976	0,979
	MAX	0,844	0,915	0,943	0,964	0,974	0,984	0,991	0,992	0,992
	AVERAGE	0,791	0,902	0,930	0,952	0,967	0,973	0,978	0,979	0,981
Micro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (I)		0,648	0,905	0,939	0,950	0,962	0,969	0,969	0,970	0,977
|C7| tf-idf&IG (I)		0,647	0,904	0,942	0,954	0,964	0,970	0,970	0,973	0,979
|C7| tf-idf&DF (I)		0,823	0,886	0,910	0,938	0,956	0,960	0,969	0,968	0,972
|C7| tf-idf&Acc2 (I)		0,837	0,898	0,937	0,947	0,964	0,966	0,972	0,971	0,977
|C7| CHI&IG (I)		0,822	0,905	0,942	0,953	0,974	0,985	0,990	0,991	0,991
|C7| CHI&DF (I)		0,844	0,908	0,938	0,952	0,967	0,969	0,975	0,976	0,977
|C7| CHI&Acc2 (I)		0,838	0,916	0,942	0,964	0,975	0,981	0,990	0,991	0,992
|C7| IG&DF (I)		0,844	0,907	0,930	0,950	0,967	0,973	0,975	0,977	0,979
|C7| IG&Acc2 (I)		0,785	0,899	0,926	0,954	0,974	0,984	0,990	0,992	0,991
|C7| DF&Acc2 (I)		0,816	0,883	0,908	0,946	0,966	0,968	0,977	0,978	0,978
	MAX	0,844	0,916	0,942	0,964	0,975	0,985	0,990	0,992	0,992
	AVERAGE	0,790	0,901	0,931	0,951	0,967	0,973	0,978	0,979	0,981

Table 7.13. Macro F-measure results of the proposed combinations in local policy for Classic3 dataset

Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (l)		0,720	0,880	0,935	0,950	0,957	0,959	0,964	0,964	0,970
CHI (l)		0,706	**0,908**	**0,945**	**0,963**	**0,974**	0,981	0,987	0,989	0,990
IG (l)		0,728	0,889	0,912	0,959	0,974	**0,986**	0,990	**0,992**	**0,991**
DF (l)		0,720	0,848	0,871	0,908	0,945	0,964	0,973	0,973	0,978
Acc2 (l)		**0,761**	0,867	0,923	0,958	0,972	0,985	**0,991**	0,991	**0,991**
	MAX	0,761	0,908	0,945	0,963	0,974	0,986	0,991	0,992	0,991
Macro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	**0,825**	0,901	0,942	0,962	**0,977**	**0,987**	0,990	**0,994**	**0,993**
	AVERAGE	0,796	0,888	0,923	0,953	0,970	0,977	0,984	0,987	0,987
Rank Combination	MAX	**0,837**	0,901	0,938	**0,964**	**0,974**	0,985	**0,991**	**0,992**	**0,991**
	AVERAGE	0,791	0,891	0,927	0,949	0,967	0,971	0,977	0,979	0,984
Macro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (l)		0,713	0,892	0,937	0,950	0,966	0,968	0,970	0,969	0,979
|C1| tf-idf&IG (l)		0,717	0,891	0,940	0,952	0,967	0,970	0,972	0,974	0,979
|C1| tf-idf&DF (l)		**0,810**	0,880	0,903	0,935	0,953	0,958	0,969	0,968	0,972
|C1| tf-idf&Acc2 (l)		0,824	0,892	0,929	0,951	0,963	0,967	0,973	0,977	0,977
|C1| CHI&IG (l)		**0,820**	0,901	0,939	0,955	**0,974**	0,984	0,988	0,991	0,990
|C1| CHI&DF (l)		**0,820**	0,899	0,936	0,950	0,968	0,968	0,977	0,974	0,976
|C1| CHI&Acc2 (l)		**0,817**	0,903	0,938	0,965	0,975	0,982	0,990	0,991	0,993
|C1| IG&DF (l)		**0,821**	0,891	0,923	0,952	0,966	0,972	0,974	0,977	0,980
|C1| IG&Acc2 (l)		**0,769**	0,891	0,918	0,954	**0,974**	0,985	0,990	0,993	**0,992**
|C1| DF&Acc2 (l)		**0,801**	0,868	0,898	0,947	0,964	0,971	0,979	0,978	0,978
	MAX	**0,824**	**0,903**	0,940	**0,965**	**0,975**	0,985	0,990	**0,993**	**0,993**
	AVERAGE	0,791	0,891	0,926	0,951	0,967	0,972	0,978	0,979	0,982
Macro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (l)		0,714	0,894	0,935	0,955	0,967	0,974	0,980	0,983	0,984
|C2| tf-idf&IG (l)		**0,815**	0,898	0,924	0,959	0,974	0,980	0,985	0,989	0,989
|C2| tf-idf&DF (l)		**0,782**	0,859	0,899	0,935	0,952	0,958	0,968	0,967	0,975
|C2| tf-idf&Acc2 (l)		0,816	0,877	0,930	0,957	0,973	0,978	0,987	0,988	0,990
|C2| CHI&IG (l)		**0,795**	0,896	0,931	0,958	0,975	0,985	0,990	0,991	0,991
|C2| CHI&DF (l)		**0,789**	0,874	0,896	0,945	0,967	0,974	0,981	0,988	0,988
|C2| CHI&Acc2 (l)		**0,807**	0,891	0,930	0,960	**0,974**	0,982	0,990	**0,992**	**0,991**
|C2| IG&DF (l)		**0,779**	0,879	0,916	0,947	0,968	0,982	0,986	0,991	0,990
|C2| IG&Acc2 (l)		**0,768**	0,888	0,919	0,955	0,975	0,990	0,990	0,993	0,992
|C2| DF&Acc2 (l)		0,749	0,867	0,923	0,949	0,970	0,978	0,988	0,991	0,990
	MAX	**0,816**	0,898	0,935	0,960	**0,975**	**0,990**	0,990	**0,993**	**0,992**
	AVERAGE	0,782	0,882	0,920	0,952	0,969	0,978	0,985	0,987	0,988
Macro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (l)		0,712	0,898	0,936	0,948	0,962	0,968	0,968	0,969	0,981
|C3| tf-idf&IG (l)		0,710	0,896	0,939	0,952	0,965	0,970	0,968	0,976	0,986
|C3| tf-idf&DF (l)		**0,804**	0,876	0,899	0,938	0,957	0,957	0,966	0,968	0,972
|C3| tf-idf&Acc2 (l)		**0,824**	0,890	0,932	0,948	0,965	0,961	0,972	0,969	0,981
|C3| CHI&IG (l)		**0,812**	0,899	0,938	0,952	**0,974**	0,985	0,991	0,991	0,990
|C3| CHI&DF (l)		**0,835**	0,897	0,935	0,949	0,965	0,969	0,974	0,977	0,982
|C3| CHI&Acc2 (l)		**0,818**	0,903	0,935	0,964	0,973	0,982	0,991	0,991	0,991
|C3| IG&DF (l)		0,837	0,896	0,938	0,947	0,964	0,971	0,976	0,982	0,985
|C3| IG&Acc2 (l)		**0,768**	0,891	0,924	0,953	0,975	0,984	0,990	0,992	0,991
|C3| DF&Acc2 (l)		**0,800**	0,872	0,899	0,943	0,965	0,967	0,978	0,978	0,982
	MAX	**0,837**	**0,903**	0,939	**0,964**	**0,975**	0,985	**0,991**	**0,992**	**0,991**
	AVERAGE	0,792	0,892	0,928	0,949	0,966	0,971	0,977	0,979	0,984

Macro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (I)		0,713	0,892	0,936	0,952	0,966	0,972	0,980	0,981	0,985
|C4| tf-idf&IG (I)		0,712	0,889	0,940	0,957	0,972	0,980	0,982	0,988	0,989
|C4| tf-idf&DF (I)		0,806	0,879	0,902	0,937	0,953	0,958	0,967	0,968	0,972
|C4| tf-idf&Acc2 (I)		0,824	0,889	0,931	0,953	0,964	0,974	0,980	0,984	0,987
|C4| CHI&IG (I)		0,813	0,901	0,939	0,955	0,975	0,984	0,989	0,991	0,991
|C4| CHI&DF (I)		0,837	0,901	0,938	0,950	0,966	0,974	0,981	0,985	0,985
|C4| CHI&Acc2 (I)		0,825	0,909	0,936	0,963	0,972	0,984	0,990	0,991	0,992
|C4| IG&DF (I)		0,837	0,892	0,920	0,954	0,973	0,976	0,986	0,989	0,991
|C4| IG&Acc2 (I)		0,769	0,891	0,918	0,954	0,973	0,985	0,990	0,993	0,992
|C4| DF&Acc2 (I)		0,800	0,868	0,900	0,948	0,965	0,976	0,982	0,986	0,989
	MAX	0,837	0,909	0,940	0,963	0,975	0,985	0,990	0,993	0,992
	AVERAGE	0,794	0,891	0,926	0,952	0,968	0,976	0,983	0,986	0,987

Macro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (I)		0,713	0,894	0,936	0,949	0,963	0,968	0,972	0,974	0,982
|C5| tf-idf&IG (I)		0,712	0,892	0,941	0,949	0,965	0,970	0,976	0,980	0,988
|C5| tf-idf&DF (I)		0,812	0,876	0,899	0,937	0,954	0,957	0,967	0,968	0,971
|C5| tf-idf&Acc2 (I)		0,824	0,888	0,933	0,947	0,963	0,966	0,972	0,977	0,982
|C5| CHI&IG (I)		0,812	0,899	0,939	0,954	0,974	0,985	0,990	0,991	0,990
|C5| CHI&DF (I)		0,837	0,901	0,937	0,948	0,965	0,971	0,975	0,982	0,982
|C5| CHI&Acc2 (I)		0,825	0,909	0,937	0,964	0,973	0,982	0,991	0,992	0,993
|C5| IG&DF (I)		0,837	0,898	0,920	0,953	0,964	0,973	0,980	0,984	0,988
|C5| IG&Acc2 (I)		0,768	0,891	0,924	0,955	0,974	0,985	0,989	0,993	0,992
|C5| DF&Acc2 (I)		0,800	0,872	0,899	0,945	0,964	0,972	0,978	0,982	0,983
	MAX	0,837	0,909	0,941	0,964	0,974	0,985	0,991	0,993	0,993
	AVERAGE	0,794	0,892	0,927	0,950	0,966	0,973	0,979	0,982	0,985

Macro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (I)		0,713	0,892	0,936	0,951	0,965	0,968	0,969	0,971	0,978
|C6| tf-idf&IG (I)		0,712	0,891	0,941	0,951	0,965	0,969	0,970	0,977	0,978
|C6| tf-idf&DF (I)		0,806	0,878	0,901	0,936	0,953	0,958	0,969	0,968	0,972
|C6| tf-idf&Acc2 (I)		0,824	0,889	0,933	0,947	0,963	0,965	0,975	0,976	0,978
|C6| CHI&IG (I)		0,812	0,905	0,939	0,954	0,974	0,984	0,988	0,991	0,991
|C6| CHI&DF (I)		0,837	0,901	0,938	0,952	0,967	0,969	0,973	0,974	0,977
|C6| CHI&Acc2 (I)		0,825	0,910	0,935	0,964	0,974	0,981	0,991	0,991	0,992
|C6| IG&DF (I)		0,837	0,898	0,920	0,952	0,965	0,973	0,976	0,976	0,979
|C6| IG&Acc2 (I)		0,769	0,891	0,918	0,954	0,974	0,985	0,989	0,993	0,992
|C6| DF&Acc2 (I)		0,801	0,873	0,898	0,945	0,965	0,970	0,980	0,976	0,979
	MAX	0,837	0,910	0,941	0,964	0,974	0,985	0,991	0,993	0,992
	AVERAGE	0,794	0,893	0,926	0,951	0,966	0,972	0,978	0,979	0,982

Macro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (I)		0,713	0,895	0,936	0,948	0,961	0,968	0,969	0,969	0,977
|C7| tf-idf&IG (I)		0,712	0,896	0,939	0,952	0,964	0,970	0,970	0,972	0,979
|C7| tf-idf&DF (I)		0,806	0,876	0,899	0,936	0,954	0,959	0,968	0,967	0,971
|C7| tf-idf&Acc2 (I)		0,824	0,888	0,934	0,945	0,964	0,966	0,971	0,970	0,977
|C7| CHI&IG (I)		0,812	0,899	0,939	0,952	0,975	0,985	0,990	0,991	0,991
|C7| CHI&DF (I)		0,837	0,899	0,937	0,950	0,967	0,970	0,976	0,976	0,978
|C7| CHI&Acc2 (I)		0,825	0,911	0,938	0,964	0,975	0,981	0,990	0,991	0,992
|C7| IG&DF (I)		0,837	0,897	0,924	0,949	0,966	0,972	0,976	0,977	0,979
|C7| IG&Acc2 (I)		0,768	0,891	0,924	0,953	0,975	0,985	0,990	0,993	0,991
|C7| DF&Acc2 (I)		0,800	0,872	0,898	0,945	0,965	0,967	0,977	0,978	0,978
	MAX	0,837	0,911	0,939	0,964	0,975	0,985	0,990	0,993	0,992
	AVERAGE	0,794	0,892	0,927	0,949	0,967	0,972	0,978	0,979	0,981

7.2. Skew Datasets

The Hitech and LA1 datasets are neither highly skew as the Wap and Reuters datasets nor homogenous as the Classic3 dataset and they were categorized as skew datasets in our study. There are two main reasons why the Hitech and LA1 datasets are labeled as a skew dataset.

In the first place, beside the Hitech and LA1 datasets have similar class distributions and consist of same number of classes, the major difference of these two datasets is dimensionality of the term space compared to other datasets. When we look at the Table 6.1, we can see that the difference between the number of document and the number of terms is more explicit than others. High dimensionality of the term space is also the main hardness for accurate classification in these datasets.

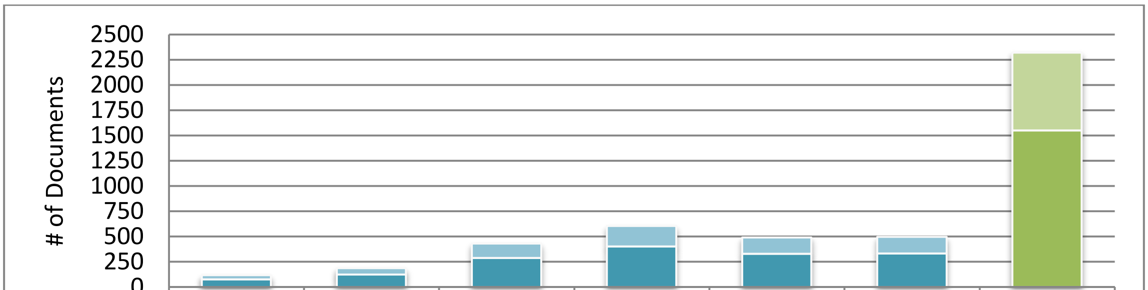
The second reason is the proportion of the shared terms in the datasets. As we discussed in Subsection 7.1, we categorized the Classic3 dataset as a homogenous dataset because about 50 percent of the terms occur in only one class and the documents that share many common terms belong to the same class in the dataset. This means that each class is disjoint from each other clearly. However in the Hitech and LA1 datasets similar topics overlap with each other. For example, the Hitech dataset contains documents about electronics, technology, medical, health, research and computers. Among these topics, the medical and health topics overlap with each other and the other three topics of the Hitech electronics, technology and computers also overlap with each other. In addition, only about 15% of terms in the Hitech dataset and 20% of terms in the LA1 dataset belong to one topic. Moreover, about 30% of the terms are shared among two topics and nearly 18% of the terms are shared among all topics in the both dataset.

Thus we categorize the Hitech and LA1 document corpora as skew datasets in our study. Firstly, we begin with the Hitech dataset and then we continue to our analysis on the LA1 dataset.

7.2.1. The Hitech Dataset

7.2.1.1. Property of the Dataset

The Hitech dataset was derived from the San Jose Mercury newspaper articles that are distributed as part of the TREC collection TIPSTER Vol. 3. It was constructed by selecting documents that are part of certain topics in which the various articles were categorized [38].



	Electronics	Technology	Medical	Health	Research	Computer	All
Test	40	62	143	201	161	163	770
Training	76	124	286	402	330	332	1550
All	116	186	429	603	491	495	2300

Figure 7.10. Property of the Hitech dataset

The Hitech dataset is one of the document corpora that many researchers used in their studies [6, 7, 8, 45, 46]. Figure 7.10 shows the property of the Hitech dataset that includes 2,300 documents about 116 Electronics, 186 Technology, 429 Medical, 603 Health, 491 Research and 495 Computers in six categories. In the study, two thirds of each class is selected for the training set and the remaining one third is used for testing.

In addition, as we can see in Figure 7.10, the last four categories constitute almost 90 percent of the Hitech dataset. The other two topics electronics and technology constitute only about 10 percent of all documents. Thus the performance of the classifier significantly decreases while classifying the documents in these categories.

7.2.1.2. Analysis of the Existing Metrics

In this section we compare both the local and global version of the feature selection metrics on the Hitech dataset. Table 7.14 shows the micro- and macro-averaged F-measure results for the global and local policies of 5 well known feature selection metrics on the Hitech dataset as a function of the number of keywords. The highest score among the metrics for each keyword number is shown in red font as we mentioned.

Micro-F	10	30	50	100	200	500	1000	1500	2000	All
tf-idf (g)	0,372	0,518	0,538	0,603	0,606	0,623	0,643	0,647	0,659	0,649
CHI (g)	0,485	0,559	**0,597**	**0,621**	0,637	0,633	**0,651**	**0,670**	**0,667**	0,649
IG (g)	0,430	0,523	0,559	**0,621**	**0,641**	0,645	0,649	0,658	0,666	0,649
DF (g)	0,214	0,546	0,538	0,583	0,616	0,609	0,624	0,624	0,629	0,649
Acc2 (g)	**0,521**	**0,581**	0,575	0,606	0,607	**0,657**	0,642	0,637	0,661	0,649
MAX	0,521	0,581	0,597	0,621	0,641	0,657	0,651	0,670	0,667	
Macro-F	10	30	50	100	200	500	1000	1500	2000	All
tf-idf (g)	0,228	0,371	0,465	0,507	0,505	0,530	0,538	0,582	0,598	0,558
CHI (g)	0,340	0,437	**0,509**	0,528	0,550	0,570	**0,610**	**0,611**	**0,605**	0,558
IG (g)	0,301	0,433	0,461	**0,538**	**0,558**	**0,572**	0,597	0,601	0,602	0,558
DF (g)	0,141	0,389	0,383	0,461	0,527	0,510	0,524	0,526	0,532	0,558
Acc2 (g)	**0,433**	**0,507**	0,496	0,521	0,522	0,567	0,582	0,567	0,603	0,558
MAX	0,433	0,507	0,509	0,538	0,558	0,572	0,610	0,611	0,605	
Micro-F	10	30	50	100	200	500	1000	1500	2000	All
tf-idf (l)	0,551	0,596	0,613	0,627	0,624	0,644	0,621	0,618	0,627	0,649
CHI (l)	0,557	0,590	0,620	0,631	0,636	0,636	0,619	0,630	0,632	0,649
IG (l)	0,510	0,610	0,617	0,638	0,630	**0,654**	0,644	0,634	0,638	0,649
DF (l)	0,501	0,550	0,578	0,624	0,613	0,622	0,644	**0,664**	**0,661**	0,649
Acc2 (l)	**0,558**	**0,612**	**0,636**	**0,649**	**0,637**	0,651	**0,659**	0,647	0,646	0,649
MAX	0,558	0,612	0,636	0,649	0,637	0,654	0,659	0,664	0,661	
Macro-F	10	30	50	100	200	500	1000	1500	2000	All
tf-idf (l)	**0,486**	**0,555**	**0,571**	0,571	0,564	0,589	0,567	0,549	0,561	0,558
CHI (l)	0,477	0,495	0,536	0,572	0,567	0,551	0,545	0,552	0,567	0,558
IG (l)	0,456	0,529	0,539	**0,577**	**0,571**	0,591	0,573	0,555	0,557	0,558
DF (l)	0,397	0,485	0,507	0,549	0,540	0,549	0,592	**0,611**	**0,603**	0,558
Acc2 (l)	0,459	0,522	0,550	0,571	0,564	**0,596**	**0,600**	0,583	0,593	0,558
MAX	0,486	0,555	0,571	0,577	0,571	0,596	0,600	0,611	0,603	

Table 7.14. Micro- and macro-averaged F-measures for Hitech dataset

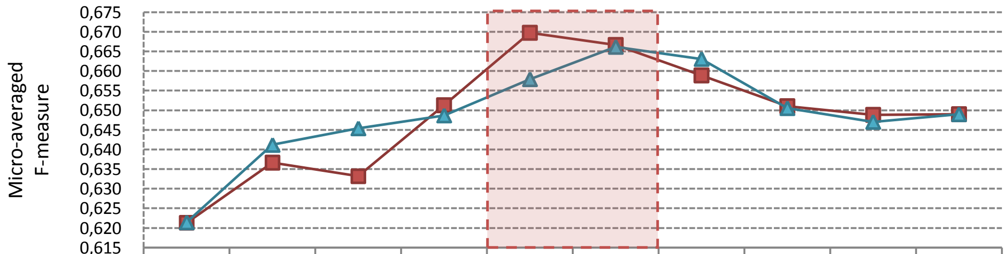
Micro- and macro-averaged F-measure results of the Classic3 dataset are similar to each other. The average of the difference between micro and macro F-measure is less than %1 (about 0.4%). However macro-averaged F-measure results are noticeably less than micro-averaged F-measure results for the Hitech dataset. In this dataset, the average of the difference between F-measures increases by about 7-8%. As we mentioned above, the Hitech dataset consists of six categories that four of them constitute almost 90 percent of the dataset and the other two categories constitute only about 10 percent of all documents. Moreover, we know that micro-averaged F-measure gives equal weight to each document whereas macro-averaged F-measure gives equal weight to each category regardless of its frequency. Thus the former is influenced by the performance of the classifier on common categories and the latter tends to be influenced by the performance of the classifier on rare categories. This is the main reason why the results of the micro- and macro-averaged F-measures are not close to each other like the Classic3 dataset.

Other observation over the Hitech dataset is the local policy achieves better micro- and macro-averaged F-measure performance than the global policy when the keyword number is low (for micro less than 500 keywords and for macro less than 1000 keywords). On the other hand, when the keyword number is high, the global policy performs better than the local policy in the Hitech dataset. As we mentioned, the global policy is not successful finding keywords that carry critical information for all classes when the keyword number is low although the accuracy increases gradually as the number of keywords increases. The highest scores are achieved by the global policy with 1500 keywords, 67% micro- and 61.1% macro-averaged F-measure performance.

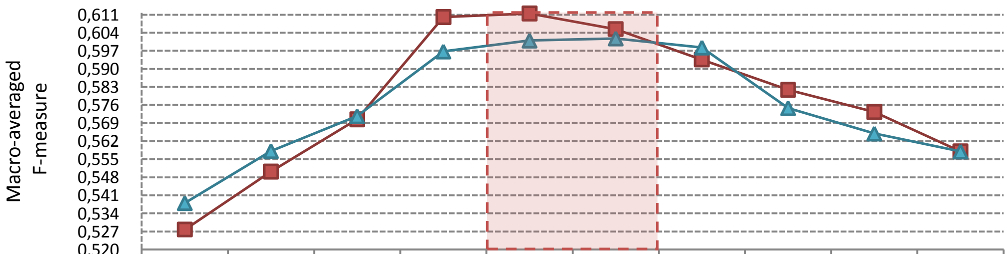
When the test documents are classified without applying any feature selection method, the classifier only achieves 64.9% micro- and 55.8% macro-averaged F-measures. On the other hand, both the global and local policies improve this performance with feature selection. For instance, the global policy with 500 and more keywords achieves higher results than the all word approach in micro-averaged F-measure and with 200 and more keywords achieves higher results than the all word approach in macro-averaged F-measure. In addition, this range drops to 100 and 50 keywords in the local policy with micro- and macro-averaged F-measure respectively.

In Table 7.14, we see that micro- and macro-averaged F-measures reach their peak values of 67% and 61.1% at 1500 keywords in the Hitech dataset. Furthermore we knew that the classifier only achieves 64.9% micro- and 55.8% macro-averaged F-measures by all words approach. This implies that the performance of the classifier drops after the peak values. In Figure 7.11 we tested the two metrics (*CHI* and *IG*), that have the highest values with a high number of keywords (1000 to 2000 keywords), also with 3000, 5000 and 10000 keywords in global policy, in order to verify this assumption.

As seen in Figure 7.11, the success rates drop after the number of keywords 1500 and 2000. This means selecting too many keywords lead to overfitting. The rationale for this descent is the training set may not be enough for extracting many relevant features for classification. Since we know that the high dimensionality of the term space is one of the main hardness for classification and in the Hitech dataset the difference between the number of document and the number of terms is very explicit.



	100	200	500	1000	1500	2000	3000	5000	10000	All
CHI (g)	0,621	0,637	0,633	0,651	0,670	0,667	0,659	0,651	0,649	0,649
IG (g)	0,621	0,641	0,645	0,649	0,658	0,666	0,663	0,651	0,647	0,649



	100	200	500	1000	1500	2000	3000	5000	10000	All
CHI (g)	0,528	0,550	0,570	0,610	0,611	0,605	0,594	0,582	0,573	0,558
IG (g)	0,538	0,558	0,572	0,597	0,601	0,602	0,598	0,575	0,565	0,558

Figure 7.11. Analysis of micro- and macro-averaged F-measures with a high number of keywords in global policy for Hitech dataset

In Hitech dataset generally *CHI*, *IG* and *Acc2* outperform other methods in global policy. When we compare to average of the F-measure results, *CHI* is the most successful metric among them. *CHI* achieves the best micro- and macro-averaged F-measure results especially with a high number of keywords from 1000 to 2000 while the performance of Acc2 is better than others with a few number of keywords in global policy.

On the other hand *Acc2* is significantly outperforms other metrics in local policy, especially with a few number of keywords. *Tf-idf* and *DF* also get successful results in local policy. DF metric has the highest results when the number of keywords is high and *tf-idf* performs much better than all metrics in term of macro-averaged results when the number of keywords is low. One of the exceptional observations is that *CHI* is the worst metric in local policy although it is the best metric in global policy. The following figure demonstrates the success rates of the existing metrics under feature number criterion.

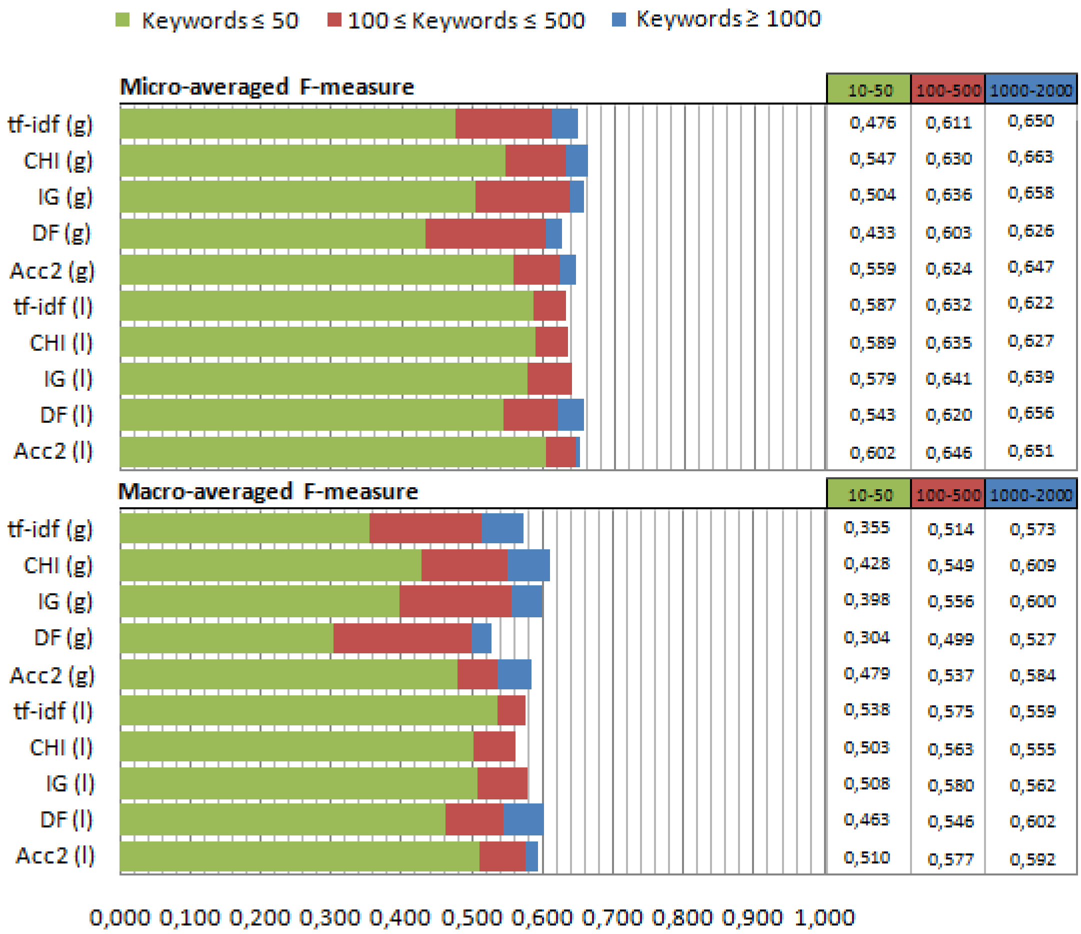


Figure 7.12. Comparison of the averages of F-measures for Hitech dataset

7.2.1.3.Analysis of Score and Rank Combinations

Tables 7.15 and 7.16 show, respectively, the micro- and macro-averaged F-measure results of the score combination and rank combination in global policy for the Hitech dataset.

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,372	0,518	0,538	0,603	0,606	0,623	0,643	0,647	0,659
CHI (g)		0,485	0,559	**0,597**	**0,621**	0,637	0,633	**0,651**	**0,670**	**0,667**
IG (g)		0,430	0,523	0,559	**0,621**	**0,641**	0,645	0,649	0,658	0,666
DF (g)		0,214	0,546	0,538	0,583	0,616	0,609	0,624	0,624	0,629
Acc2 (g)		**0,521**	**0,581**	0,575	0,606	0,607	**0,657**	0,642	0,637	0,661
	MAX	0,521	0,581	0,597	0,621	0,641	0,657	0,651	0,670	0,667
Micro-F	Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (g)		0,494	0,557	0,581	0,597	0,622	0,639	0,647	0,656	0,647
|S| tf-idf&IG (g)		0,453	0,567	0,556	0,599	0,618	0,632	**0,657**	0,654	0,650
|S| tf-idf&DF (g)		0,498	0,521	0,562	0,577	0,606	0,618	0,626	0,646	0,655
|S| tf-idf&Acc2 (g)		**0,535**	0,565	0,575	0,605	0,619	0,634	0,644	0,651	0,658
|S| CHI&IG (g)		0,499	0,561	**0,597**	0,618	0,631	0,635	**0,664**	0,655	0,658
|S| CHI&DF (g)		0,504	0,563	**0,604**	0,597	0,606	0,628	0,625	0,633	0,641
|S| CHI&Acc2 (g)		0,511	0,580	**0,611**	**0,623**	0,625	0,642	0,642	0,642	0,666
|S| IG&DF (g)		**0,527**	0,571	**0,608**	0,591	0,615	0,626	0,627	0,641	0,645
|S| IG&Acc2 (g)		**0,521**	0,563	0,577	0,614	0,626	0,635	0,650	0,655	**0,673**
|S| DF&Acc2 (g)		0,517	**0,592**	**0,605**	0,608	0,615	0,635	0,619	0,631	0,637
	MAX	**0,535**	**0,592**	**0,611**	0,623	0,631	0,642	**0,664**	0,656	**0,673**
	AVERAGE	0,506	**0,564**	0,588	0,603	0,618	0,632	0,640	0,646	0,653
Micro-F	Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (g)		**0,526**	0,559	0,594	0,605	0,625	0,655	**0,662**	0,660	0,660
|R| tf-idf&IG (g)		0,454	0,561	**0,603**	0,615	0,628	0,642	**0,655**	0,666	0,662
|R| tf-idf&DF (g)		0,501	0,543	0,563	0,590	0,601	0,625	0,627	0,642	0,648
|R| tf-idf&Acc2 (g)		**0,527**	0,570	0,580	0,608	0,629	0,634	0,645	0,660	0,663
|R| CHI&IG (g)		0,499	0,560	**0,604**	0,616	0,623	0,646	**0,661**	0,665	0,655
|R| CHI&DF (g)		**0,535**	0,564	0,588	0,610	0,628	0,656	0,646	0,669	**0,669**
|R| CHI&Acc2 (g)		0,511	0,566	**0,607**	0,620	0,628	0,645	**0,663**	0,662	0,663
|R| IG&DF (g)		**0,534**	0,557	0,588	0,611	0,622	0,636	0,640	0,652	0,665
|R| IG&Acc2 (g)		**0,521**	0,561	0,578	**0,626**	0,626	0,638	**0,657**	0,665	**0,677**
|R| DF&Acc2 (g)		0,515	0,566	0,593	0,605	0,614	0,627	0,642	0,634	0,640
	MAX	**0,535**	0,570	**0,607**	**0,626**	0,629	0,656	**0,663**	0,669	**0,677**
	AVERAGE	**0,512**	0,561	**0,590**	**0,611**	**0,622**	**0,640**	**0,650**	**0,657**	**0,660**

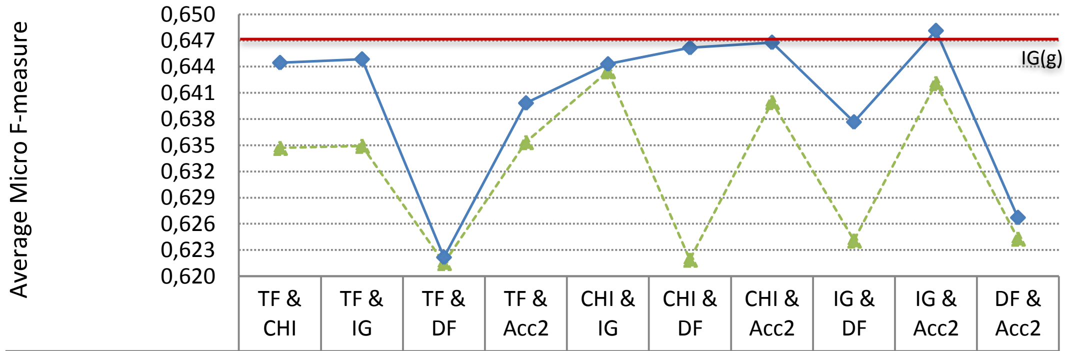
Table 7.15. In global policy, micro-averaged F-measures of the score and rank combinations for Hitech dataset

Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,228	0,371	0,465	0,507	0,505	0,530	0,538	0,582	0,598
CHI (g)		0,340	0,437	**0,509**	0,528	0,550	0,570	**0,610**	**0,611**	**0,605**
IG (g)		0,301	0,433	0,461	**0,538**	**0,558**	**0,572**	0,597	0,601	0,602
DF (g)		0,141	0,389	0,383	0,461	0,527	0,510	0,524	0,526	0,532
Acc2 (g)		**0,433**	**0,507**	0,496	0,521	0,522	0,567	0,582	0,567	0,603
	MAX	0,433	0,507	0,509	0,538	0,558	0,572	0,610	0,611	0,605
Macro-F	Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (g)		0,348	0,460	0,478	0,505	0,528	0,543	0,576	0,605	0,585
|S| tf-idf&IG (g)		0,309	0,467	0,460	0,510	0,520	0,537	0,562	0,600	0,587
|S| tf-idf&DF (g)		0,350	0,369	0,444	0,491	0,517	0,516	0,524	0,541	0,592
|S| tf-idf&Acc2 (g)		0,424	0,470	0,475	0,529	0,523	0,543	0,545	0,595	0,597
|S| CHI&IG (g)		0,351	0,469	0,503	0,527	0,541	0,579	**0,610**	0,596	0,596
|S| CHI&DF (g)		0,355	0,402	**0,513**	0,511	0,506	0,529	0,519	0,537	0,570
|S| CHI&Acc2 (g)		0,403	0,506	0,518	0,541	0,537	0,552	0,580	0,584	0,604
|S| IG&DF (g)		0,373	0,438	**0,513**	0,508	0,514	0,533	0,519	0,535	0,568
|S| IG&Acc2 (g)		**0,433**	0,468	0,497	0,525	0,538	0,549	0,585	0,597	0,610
|S| DF&Acc2 (g)		0,405	0,491	**0,514**	0,516	0,528	0,546	0,525	0,541	0,569
	MAX	**0,433**	0,506	**0,518**	0,541	0,541	**0,579**	**0,610**	0,605	**0,610**
	AVERAGE	0,375	0,454	0,491	0,516	0,525	0,543	0,555	0,573	0,588
Macro-F	Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (g)		0,372	0,471	0,495	0,505	0,535	0,600	0,599	0,601	0,594
|R| tf-idf&IG (g)		0,311	0,467	0,503	0,517	0,542	0,545	0,601	**0,611**	0,599
|R| tf-idf&DF (g)		0,352	0,388	0,446	0,503	0,507	0,523	0,520	0,537	0,545
|R| tf-idf&Acc2 (g)		0,371	0,473	0,491	0,515	0,536	0,543	0,565	0,609	0,594
|R| CHI&IG (g)		0,351	0,468	**0,513**	0,533	0,537	**0,599**	0,602	0,606	0,590
|R| CHI&DF (g)		0,379	0,470	0,495	0,511	0,528	0,554	0,573	0,606	**0,609**
|R| CHI&Acc2 (g)		0,403	0,494	0,524	0,532	0,549	**0,594**	0,608	0,601	0,598
|R| IG&DF (g)		0,377	0,427	0,489	0,512	0,531	0,534	0,542	0,567	**0,608**
|R| IG&Acc2 (g)		0,433	0,462	0,498	0,531	0,544	0,551	0,605	0,608	0,615
|R| DF&Acc2 (g)		0,362	0,432	0,503	0,518	0,530	0,544	0,540	0,529	0,537
	MAX	**0,433**	0,494	**0,524**	**0,533**	0,549	**0,600**	**0,608**	**0,611**	**0,615**
	AVERAGE	0,371	0,455	0,496	0,518	0,534	0,559	0,575	0,588	0,589

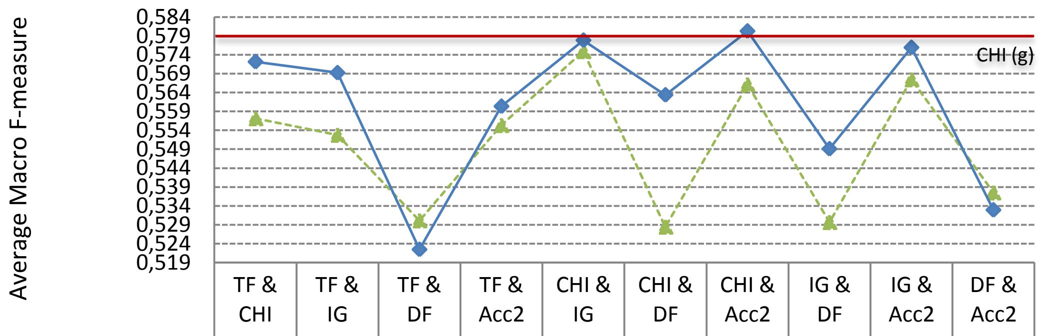
Table 7.16. In global policy, macro-averaged F-measures of the score and rank combinations for Hitech dataset

In addition, Tables 7.17 and 7.18 show the micro- and macro-averaged F-measure results of the combination experiments in local policy for the Hitech dataset. In the first place, the F-measure results of the score and rank combinations are compared with the existing metrics in the case of global policy. When we look at Tables 7.15 and 7.16, we can see that rank combination is more successful than score combination in global policy. This difference is clearer in the average of the each keyword number as seen "AVERAGE" rows at tables. Improvement of the F-measure values by combinations is more apparent when the keyword number is low between 10 and 100.

In global policy among the 10 possible combinations of two feature selection metrics, *CHI & IG*, *CHI & Acc2* and *IG & Acc2* score and rank combinations are more successful than other combinations based on the highest micro- and macro-averaged F-measure values for each keyword number.



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,635	0,635	0,622	0,635	0,643	0,622	0,640	0,624	0,642	0,624
Rank Com. (g)	0,644	0,645	0,622	0,640	0,644	0,646	0,647	0,638	0,648	0,627



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,557	0,553	0,530	0,555	0,575	0,528	0,566	0,530	0,567	0,537
Rank Com. (g)	0,572	0,569	0,522	0,560	0,578	0,563	0,580	0,549	0,576	0,533

	tf-idf (g)	CHI (g)	IG (g)	DF (g)	Acc2 (g)
Micro	0,630	0,646	**0,647**	0,614	0,635
Macro	0,543	**0,579**	0,578	0,513	0,560

Figure 7.13. In global policy, comparison of score and rank combinations on the Hitech dataset

In global policy, among the individual feature selection metrics *CHI* has the highest average of micro- and macro-averaged F-measure values of 0.639 and 0.569. The success of the CHI metric is improved by |*R*| *CHI & Acc2* and |*R*| *CHI & IG* combinations. Figure 7.13 shows the averages of micro- and macro-averaged F-measure values from 50 to 2000 keywords for the Hitech dataset to compare between score combination and rank combination of feature selection metrics.

In global policy the performance of Acc2 is better than others with a few number of keywords. When we look at the related tables, we can see that both the score and rank combinations of *CHI & Acc2, IG & DF* and *DF & Acc2* are more successful than individual *Acc2* when the keyword number is low. On the other hand the other individual metric *CHI* achieves the best micro- and macro-averaged F-measure results especially with a high number of keywords from 1000 to 2000. Only the rank combination of *IG & Acc2* outperforms its success. Although the success of individual metrics cannot be improved at some number of keywords, we can say that the rank combination of *IG & Acc2* is the most successful combination among the score and rank combinations in global policy when the number of keywords is high.

Finally, among the existing metrics the best performance of 67% micro-averaged and 61.1% macro-averaged F-measure is achieved by CHI with 1500 keywords in global policy. This performance is outperformed by the rank combination of *IG & Acc2* with value of 67.7% micro-averaged and 61.5% macro-averaged F-measure in global policy.

We also compare the results of the score and rank combinations with the existing metrics in the case of local policy. When we perform these two combinations on the Hitech dataset in local policy, score combination is significantly better than the rank combination when the number of keywords is high but rank combination is more successful than score combination when the number of keywords is low from 10 to 50 as seen in Tables 7.17 and 7.18 and Figure 7.13. Furthermore the success of the combinations compare to the individual metrics is more explicit in local policy and among the 10 possible combinations of two feature selection metrics *tf-idf & IG*, *tf-idf & Acc2, CHI & Acc2, IG & DF* and *IG & Acc2* score and rank combinations are more successful than other combinations in the experiments.

In Table 7.14, we summarized that among the individual feature selection metrics, *Acc2* is significantly outperforms other metrics in the case of local policy. One of the interesting results of the experiments is all possible binary combination of *Acc2* with other metrics (*tf-idf & Acc2, CHI & Acc2, IG & Acc2 and DF & Acc2)* has successful F-measure results in local policy.

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (l)		0,551	0,596	0,613	0,627	0,624	0,644	0,621	0,618	0,627
CHI (l)		0,557	0,590	0,620	0,631	0,636	0,636	0,619	0,630	0,632
IG (l)		0,510	0,610	0,617	0,638	0,630	**0,654**	0,644	0,634	0,638
DF (l)		0,501	0,550	0,578	0,624	0,613	0,622	0,644	**0,664**	**0,661**
Acc2 (l)		**0,558**	**0,612**	**0,636**	**0,649**	**0,637**	0,651	**0,659**	0,647	0,646
	MAX	0,558	0,612	0,636	0,649	0,637	0,654	0,659	0,664	0,661
Micro-F	Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (l)		0,542	0,604	0,607	0,646	0,635	0,642	0,638	0,627	0,634
|S| tf-idf&IG (l)		0,540	0,611	0,624	0,647	0,654	0,648	0,650	0,625	0,637
|S| tf-idf&DF (l)		0,537	0,576	0,595	0,629	0,632	0,646	0,647	0,649	0,635
|S| tf-idf&Acc2 (l)		0,530	0,608	0,640	0,644	0,660	0,662	0,657	0,639	0,639
|S| CHI&IG (l)		0,539	0,610	0,609	0,641	0,637	0,649	0,629	0,639	0,635
|S| CHI&DF (l)		0,550	0,584	0,601	0,633	0,648	0,649	0,661	0,661	0,651
|S| CHI&Acc2 (l)		0,560	0,621	0,639	0,647	0,644	0,656	0,659	0,646	0,634
|S| IG&DF (l)		0,533	0,605	0,618	0,643	0,632	0,652	0,660	0,659	0,646
|S| IG&Acc2 (l)		0,553	0,609	0,637	0,639	0,647	0,653	0,665	0,652	0,645
|S| DF&Acc2 (l)		0,534	0,605	0,642	0,640	0,633	0,641	0,652	0,661	0,661
	MAX	0,560	0,621	0,642	0,647	0,660	0,662	0,665	0,661	0,661
	AVERAGE	0,542	0,603	0,621	0,641	0,642	0,650	0,652	0,646	0,642
Micro-F	Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (l)		0,554	0,622	0,622	0,623	0,625	0,620	0,622	0,608	0,619
|R| tf-idf&IG (l)		0,547	0,617	0,616	0,628	0,641	0,628	0,636	0,634	0,618
|R| tf-idf&DF (l)		0,517	0,567	0,599	0,624	0,638	0,654	0,651	0,646	0,638
|R| tf-idf&Acc2 (l)		0,541	0,592	0,612	0,631	0,654	0,655	0,658	0,633	0,644
|R| CHI&IG (l)		0,557	0,608	0,623	0,646	0,639	0,643	0,632	0,637	0,631
|R| CHI&DF (l)		0,529	0,614	0,636	0,637	0,633	0,651	0,643	0,637	0,635
|R| CHI&Acc2 (l)		0,568	0,617	0,647	0,637	0,629	0,659	0,643	0,637	0,643
|R| IG&DF (l)		0,529	0,620	0,637	0,625	0,647	0,661	0,656	0,646	0,632
|R| IG&Acc2 (l)		0,553	0,605	0,638	0,639	0,641	0,655	0,652	0,644	0,640
|R| DF&Acc2 (l)		0,522	0,610	0,631	0,641	0,639	0,654	0,661	0,656	0,658
	MAX	0,568	0,622	0,647	0,646	0,654	0,661	0,661	0,656	0,658
	AVERAGE	0,542	0,607	0,626	0,633	0,639	0,648	0,645	0,638	0,636

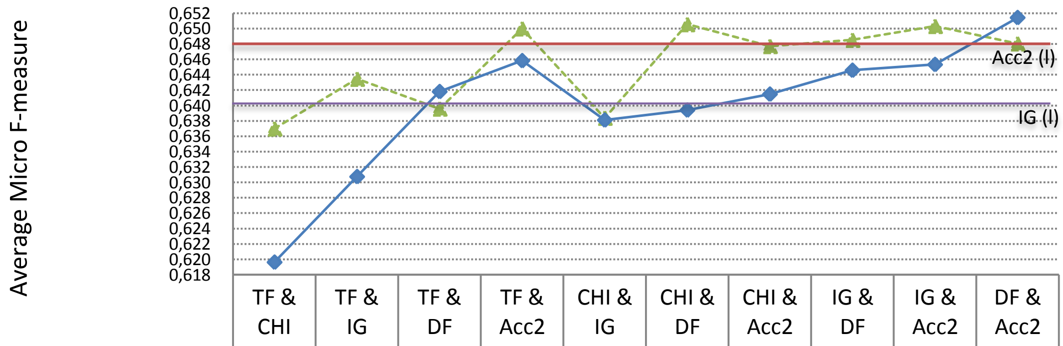
Table 7.17. In local policy, micro-averaged F-measures of the score and rank combinations for Hitech dataset

Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (l)		**0,486**	**0,555**	**0,571**	0,571	0,564	0,589	0,567	0,549	0,561
CHI (l)		0,477	0,495	0,536	0,572	0,567	0,551	0,545	0,552	0,567
IG (l)		0,456	0,529	0,539	**0,577**	**0,571**	0,591	0,573	0,555	0,557
DF (l)		0,397	0,485	0,507	0,549	0,540	0,549	0,592	**0,611**	**0,603**
Acc2 (l)		0,459	0,522	0,550	0,571	0,564	**0,596**	**0,600**	0,583	0,593
	MAX	0,486	0,555	0,571	0,577	0,571	0,596	0,600	0,611	0,603
Macro-F	Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (l)		0,473	0,550	0,552	**0,595**	0,569	0,565	0,576	0,580	0,579
|S| tf-idf&IG (l)		0,457	**0,565**	**0,582**	**0,593**	**0,594**	0,585	0,595	0,571	0,572
|S| tf-idf&DF (l)		0,456	0,502	0,530	**0,578**	**0,573**	0,590	0,592	0,589	0,573
|S| tf-idf&Acc2 (l)		0,456	0,542	**0,579**	**0,591**	**0,610**	**0,601**	**0,603**	0,583	0,573
|S| CHI&IG (l)		0,438	0,523	0,529	0,567	**0,575**	0,570	0,549	0,563	0,555
|S| CHI&DF (l)		0,475	0,512	0,525	0,559	**0,578**	0,595	0,591	0,595	0,583
|S| CHI&Acc2 (l)		0,472	0,541	0,560	**0,594**	**0,583**	**0,597**	0,589	0,579	0,565
|S| IG&DF (l)		0,450	0,528	0,535	0,564	0,553	**0,596**	0,599	0,593	0,579
|S| IG&Acc2 (l)		0,452	0,521	0,548	0,568	**0,593**	**0,596**	**0,604**	0,581	0,590
|S| DF&Acc2 (l)		0,453	0,521	**0,572**	0,570	0,563	0,560	0,594	0,592	0,598
	MAX	0,475	**0,565**	**0,582**	**0,595**	**0,610**	**0,601**	**0,604**	0,595	0,598
	AVERAGE	0,458	0,530	0,551	**0,578**	**0,579**	**0,586**	**0,589**	**0,583**	**0,577**
Macro-F	Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (l)		0,479	**0,566**	0,559	0,556	0,540	0,552	0,554	0,552	0,556
|R| tf-idf&IG (l)		0,476	**0,565**	0,569	0,570	**0,589**	0,558	0,566	0,575	0,562
|R| tf-idf&DF (l)		0,447	0,499	0,528	0,561	**0,573**	**0,605**	**0,603**	0,582	0,575
|R| tf-idf&Acc2 (l)		0,454	0,519	0,540	0,565	**0,606**	**0,596**	0,597	0,576	0,586
|R| CHI&IG (l)		0,478	0,520	**0,571**	**0,585**	**0,579**	0,564	0,553	0,560	0,564
|R| CHI&DF (l)		0,441	0,534	0,555	**0,594**	**0,579**	0,589	0,571	0,560	0,573
|R| CHI&Acc2 (l)		0,484	0,546	**0,588**	**0,590**	0,565	**0,599**	0,576	0,554	0,584
|R| IG&DF (l)		0,433	0,541	0,558	0,540	**0,597**	**0,597**	0,585	0,576	0,573
|R| IG&Acc2 (l)		0,452	0,514	0,556	0,572	**0,584**	0,593	0,588	0,567	0,562
|R| DF&Acc2 (l)		0,444	0,525	0,548	0,565	**0,572**	**0,596**	**0,602**	0,601	0,597
	MAX	0,484	**0,566**	**0,588**	**0,594**	**0,606**	**0,605**	**0,603**	0,601	0,597
	AVERAGE	**0,459**	**0,533**	**0,557**	0,570	0,578	0,585	0,579	0,570	0,573

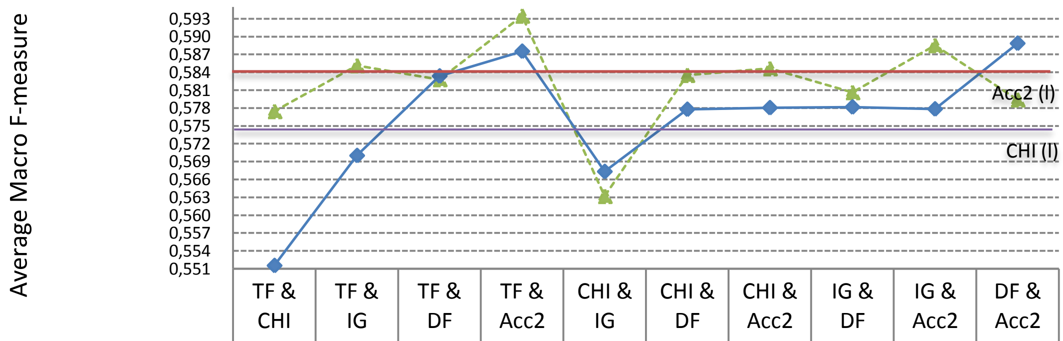
Table 7.18. In local policy, macro-averaged F-measures of the score and rank combinations for Hitech dataset

In local policy, *DF* has the highest 66.4% micro- and 61.1% macro-averaged F-measure value with 1500 keywords. The score combination of *IG & Acc2* increases the micro-averaged F-measure to 66.5% with 1000 keywords and the score combination of *tf-idf & Acc2* approaches the highest macro-averaged F-measure with the 61% success rate by using only 200 keywords.

As we mentioned, among the individual feature selection metrics, *Acc2* is the best metric in local policy when the keyword numbers are lower than 1500. After testing the combinations of the individual metrics, score and rank combinations of *CHI & Acc2 and tf-idf & Acc2* excel the success of the individual *Acc2*. Thus we can say that their score combinations are better than all individual metrics when the keyword numbers are lower than 1500.



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (I)	0,637	0,643	0,640	0,650	0,638	0,651	0,648	0,649	0,650	0,648
Rank Com. (I)	0,620	0,631	0,642	0,646	0,638	0,639	0,641	0,645	0,645	0,651



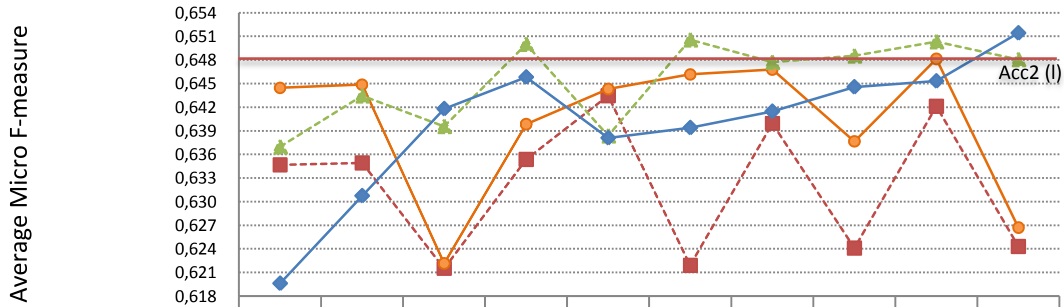
	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (I)	0,577	0,585	0,583	0,593	0,563	0,584	0,585	0,581	0,589	0,579
Rank Com. (I)	0,552	0,570	0,583	0,588	0,567	0,578	0,578	0,578	0,578	0,589

	tf-idf (I)	CHI (I)	IG (I)	DF (I)	Acc2 (I)
Micro	0,627	0,631	0,640	0,638	**0,648**
Macro	0,567	0,559	0,571	0,574	**0,584**

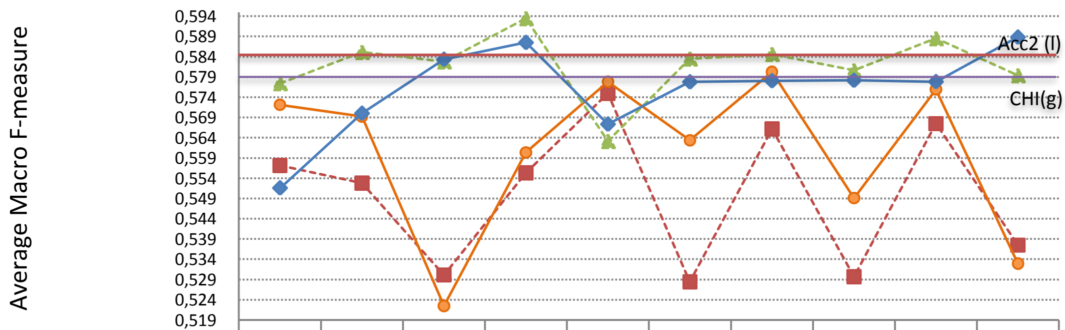
Figure 7.14. In global policy, comparison of score and rank combinations on the Hitech dataset

In Figure 7.14, we can see the success of the combination approach more clearly. We knew that *Acc2* has the highest average of F-measure in local policy and *CHI* is the second best metric that has the second highest average of F-measure in global policy. Almost all score and rank combinations in local policy are better than the second best average. Thus we can say that combining individual metrics is a good approach for the Hitech dataset.

Finally, the performance of rank combination is better than score combination in global policy whereas score combination is more successful than rank combination in local policy in the Hitech dataset as seen in Figure 7.15. Score combination in local policy is the best method for classification on this dataset like homogenous datasets.



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,635	0,635	0,622	0,635	0,643	0,622	0,640	0,624	0,642	0,624
Rank Com. (g)	0,644	0,645	0,622	0,640	0,644	0,646	0,647	0,638	0,648	0,627
Score Com. (l)	0,637	0,643	0,640	0,650	0,638	0,651	0,648	0,649	0,650	0,648
Rank Com. (l)	0,620	0,631	0,642	0,646	0,638	0,639	0,641	0,645	0,645	0,651



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,557	0,553	0,530	0,555	0,575	0,528	0,566	0,530	0,567	0,537
Rank Com. (g)	0,572	0,569	0,522	0,560	0,578	0,563	0,580	0,549	0,576	0,533
Score Com. (l)	0,577	0,585	0,583	0,593	0,563	0,584	0,585	0,581	0,589	0,579
Rank Com. (l)	0,552	0,570	0,583	0,588	0,567	0,578	0,578	0,578	0,578	0,589

Figure 7.15. Comparison of score and rank combinations in global and local policy on the Hitech dataset

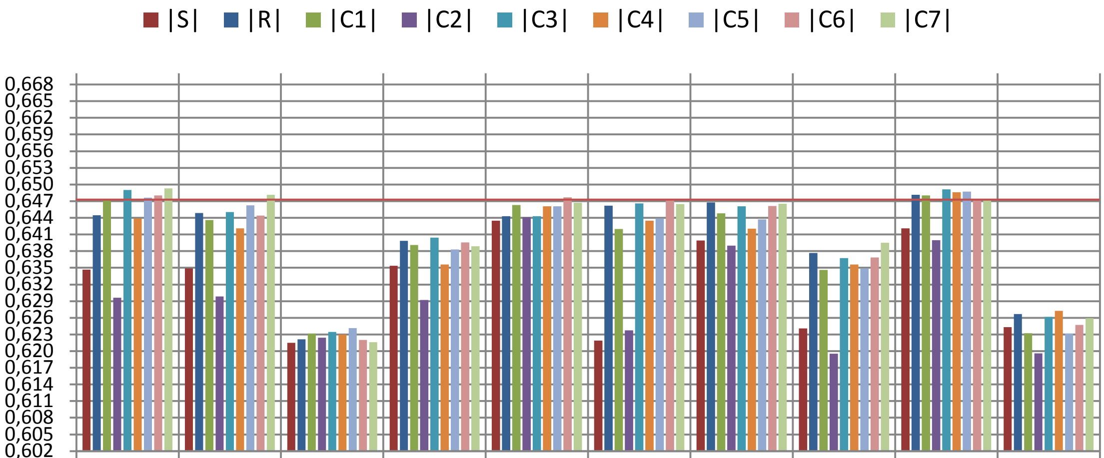
7.2.1.4. Analysis of the Proposed Combinations

In the previous sections firstly we analyzed the existing metrics for the Hitech dataset. According to that analysis, the global policy was more successful than the local policy when the keyword number was high but it was outperformed by the local policy when the keyword number was low and we concluded that *CHI* was the best metric with a high number of keywords in global policy while *Acc2* was significantly outperforms other metrics with a few number of keywords in local policy. Then we evaluated the performance of the score and rank combinations on this dataset. The performance of rank combination was better than score combination in global policy whereas score combination was more successful than rank combination in local policy. In global policy among the 10 possible combinations of two feature selection metrics, *CHI & IG*, *CHI & Acc2* and *IG & Acc2* rank combinations were more successful than other combinations and in local policy *tf-idf & Acc2, CHI & Acc2* and *IG & Acc2* score combinations were more successful than others in these experiments. As a result we can say that combining individual metrics is a good approach for the Hitech dataset.

After testing the score and rank combinations, we now test the proposed methods on the Hitech dataset. We begin with the experiments in global policy and then we analyze these combinations in the case of local policy. Tables 7.21 and 7.22 show the micro- and macro-averaged F-measure results, respectively, for all seven proposed combination methods in global policy for the Hitech dataset.

Before comparing the performance of each proposed method, we determine which combinations of two feature selection metrics are the best. In Figures 7.16 and 7.17, we can see that *IG & Acc2, tf-idf & CHI, CHI & IG* and *CHI & Acc2* combinations are more successful than the other combinations but we can also see that *tf-idf &DF, DF & Acc2, tf-idf & Acc2 and IG & DF* combinations are unsuccessful. Especially *tf-idf &DF* and *DF & Acc2* are the worst combinations among the 10 possible combinations. In these figures we can see that all possible combinations of the feature selection metric *CHI* have better performance than other combinations. We should not forget that *CHI* is the best metric with a high number of keywords in global policy. We also knew, the performance of the individual metrics *CHI*, *IG* and *Acc2* outperformed other metrics in global policy.

In Figures 7.16 and 7.17, we can also see that combinations of these successful individual metrics show the best performance during test phase.

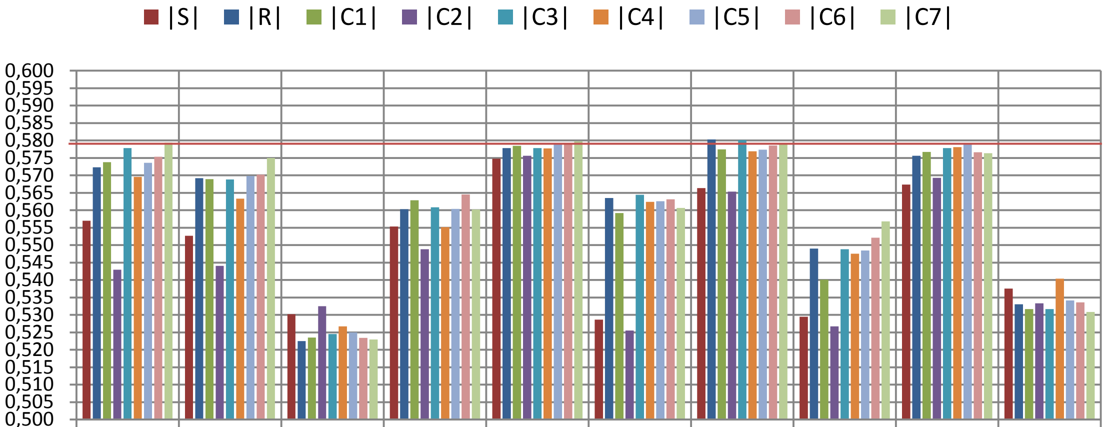


	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,635	0,635	0,622	0,635	0,643	0,622	0,640	0,624	0,642	0,624
|R|	0,644	0,645	0,622	0,640	0,644	0,646	0,647	0,638	0,648	0,627
|C1|	0,647	0,644	0,623	0,639	0,646	0,642	0,645	0,635	0,648	0,623
|C2|	0,630	0,630	0,622	0,629	0,644	0,624	0,639	0,620	0,640	0,620
|C3|	0,649	0,645	0,623	0,640	0,644	0,647	0,646	0,637	0,649	0,626
|C4|	0,644	0,642	0,623	0,636	0,646	0,643	0,642	0,636	0,649	0,627
|C5|	0,648	0,646	0,624	0,638	0,646	0,644	0,644	0,635	0,649	0,623
|C6|	0,648	0,644	0,622	0,640	0,648	0,647	0,646	0,637	0,647	0,625
|C7|	0,649	0,648	0,622	0,639	0,647	0,647	0,647	0,640	0,647	0,626

Figure 7.16. In global policy, averages of the micro-averaged F-measures of all combinations for Hitech dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	**0,535**	**0,592**	**0,611**	0,623	**0,631**	0,642	**0,664**	0,656	0,673
Combination | R |	**0,535**	0,570	0,607	**0,626**	0,629	**0,656**	0,663	**0,669**	0,677
Combination |C1|	0,538	0,593	0,602	0,629	0,637	0,649	0,659	0,668	0,673
Combination |C2|	0,525	0,580	0,603	0,619	0,632	0,641	0,662	0,663	0,673
Combination |C3|	0,535	0,571	0,607	0,626	0,632	0,654	0,663	0,668	0,677
Combination |C4|	0,538	0,589	0,606	0,622	0,637	0,644	0,661	0,671	0,677
Combination |C5|	0,543	0,578	0,607	0,629	0,632	0,658	0,663	0,667	0,675
Combination |C6|	0,543	0,587	0,609	0,633	0,634	0,657	0,664	0,669	0,675
Combination |C7|	0,543	0,582	0,609	0,634	0,638	0,654	0,665	0,667	0,677

Table 7.19. In global policy, maximum micro-averaged F-measure of all combinations for Hitech dataset



	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,557	0,553	0,530	0,555	0,575	0,529	0,566	0,530	0,567	0,538
|R|	0,572	0,569	0,523	0,560	0,578	0,564	0,580	0,549	0,576	0,533
|C1|	0,574	0,569	0,524	0,563	0,578	0,559	0,577	0,540	0,577	0,532
|C2|	0,543	0,544	0,533	0,549	0,576	0,526	0,565	0,527	0,569	0,533
|C3|	0,578	0,569	0,524	0,561	0,578	0,564	0,580	0,549	0,578	0,532
|C4|	0,570	0,563	0,527	0,555	0,578	0,562	0,577	0,548	0,578	0,540
|C5|	0,574	0,570	0,525	0,560	0,579	0,563	0,577	0,548	0,579	0,534
|C6|	0,575	0,570	0,523	0,565	0,579	0,563	0,579	0,552	0,577	0,534
|C7|	0,579	0,575	0,523	0,560	0,580	0,561	0,579	0,557	0,576	0,531

Figure 7.17. In global policy, averages of the macro-averaged F-measures of all combinations for Hitech dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	**0,433**	**0,506**	0,518	**0,541**	0,541	0,579	**0,610**	0,605	0,610
Combination | R |	**0,433**	0,494	**0,524**	0,533	**0,549**	**0,600**	0,608	**0,611**	**0,615**
Combination |C1|	0,433	0,493	0,519	0,542	0,547	0,604	0,605	0,612	0,614
Combination |C2|	0,433	0,503	0,510	0,536	0,547	0,579	0,609	0,603	0,613
Combination |C3|	0,433	0,488	0,524	0,533	0,553	0,599	0,608	0,611	0,615
Combination |C4|	0,433	0,493	0,514	0,547	0,547	0,594	0,610	0,613	0,617
Combination |C5|	0,433	0,493	0,524	0,541	0,554	0,598	0,607	0,612	0,616
Combination |C6|	0,433	0,492	0,523	0,542	0,548	0,599	0,606	0,613	0,616
Combination |C7|	0,433	0,492	0,523	0,549	0,548	0,603	0,608	0,611	0,616

Table 7.20. In global policy, maximum macro-averaged F-measure of all combinations for Hitech dataset

In the case of global policy rank combination was more successful than score combination when we compared two combinations in the previous section. After testing proposed combination methods, we can say that the performances of the all proposed methods except C2 are better than score combination. Among these proposed methods *C4, C5, C6* and *C7* also outperform the success of the rank combination when we look at the micro- and macro-averaged F-measure values. The most successful combinations are obtained from these proposed combination methods. C4, C6 and C7 apparently improve the performance of the rank combination from 10 to 2000 keywords and the worst proposed method is C2. Its performance is even worse than the performance of the score combination.

As we said before, the performance of the combinations *IG & Acc2, tf-idf & CHI, CHI & IG* and *CHI & Acc2* outperforms others. Among them *tf-idf & CHI* and *CHI & Acc2, CHI & IG* are especially successful when the keyword numbers less than 1000 and combination of *IG & Acc2* has the highest results when the keyword number is higher than 50.

One of the important results of this experiment is that all proposed combination methods improve the highest F-measure values of the individual metrics as seen in Tables 7.19 and 7.20. From the previous experiments we found out that the highest scores were achieved by *CHI* in global policy with 1500 keywords, 67% micro- and 61.1% macro-averaged F-measure performance. This performance was outperformed by the rank combination of *IG & Acc2* with value of 67.7% micro-averaged and 61.5% macro-averaged F-measure in global policy when we performed rank combination on Hitech dataset. Furthermore the success of the rank combination is achieved by C3 of *IG & Acc2* with value of 67.7% micro-averaged and 61.5% macro-averaged F-measure and is improved by C7 of *IG & Acc2* with value of 67.7% and 61.6% and by C4 of *IG & Acc2* with value of 67.7% micro-averaged and 61.7% macro-averaged F-measure. The performance of the C4 with 2000 keywords is the highest results that obtained from the experiments in Hitech dataset.

Finally, the success of the combinations in proposed methods is related with the success of the combinations in score and rank combinations. If the rank or score combination of two feature selection metric is successful, they are also generally successful in the proposed combination methods. For instance performance of the combinations *IG & Acc2, CHI & IG* and *CHI & Acc2* are more successful than the others in the case of rank combination and these three combinations also give the best results with our proposed methods but it should be noted that the performance of the some combinations is improved by only the proposed methods. For example the performance of the combinations *tf-idf & CHI, CHI & IG* and *tf-idf & IG* is significantly enhanced by the methods C6 and C7.

Table 7.21. Micro F-measure results of the proposed combinations in global policy for Hitech dataset

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,372	0,518	0,538	0,603	0,606	0,623	0,643	0,647	0,659
CHI (g)		0,485	0,559	**0,597**	**0,621**	0,637	0,633	**0,651**	**0,670**	**0,667**
IG (g)		0,430	0,523	0,559	**0,621**	**0,641**	0,645	0,649	0,658	0,666
DF (g)		0,214	0,546	0,538	0,583	0,616	0,609	0,624	0,624	0,629
Acc2 (g)		**0,521**	**0,581**	0,575	0,606	0,607	**0,657**	0,642	0,637	0,661
	MAX	0,521	0,581	0,597	0,621	0,641	0,657	0,651	0,670	0,667
Micro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	0,535	0,592	0,611	0,623	0,631	0,642	0,664	0,656	0,673
	AVERAGE	0,506	0,564	0,588	0,603	0,618	0,632	0,640	0,646	0,653
Rank Combination	MAX	0,535	0,570	0,607	0,626	0,629	0,656	0,663	0,669	0,677
	AVERAGE	0,512	0,561	0,590	0,611	0,622	0,640	0,650	0,657	0,660
Micro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (g)		0,490	0,558	**0,601**	**0,629**	**0,637**	0,646	0,649	0,663	0,659
|C1| tf-idf&IG (g)		0,453	0,567	0,592	**0,624**	0,625	0,635	**0,657**	0,658	0,662
|C1| tf-idf&DF (g)		0,498	0,550	0,565	0,587	0,606	0,627	0,628	0,644	0,646
|C1| tf-idf&Acc2 (g)		**0,534**	0,564	0,580	0,618	0,625	0,631	0,640	0,658	0,661
|C1| CHI&IG (g)		0,499	0,560	0,590	0,618	0,629	0,649	**0,659**	0,658	0,666
|C1| CHI&DF (g)		**0,538**	0,577	0,573	**0,623**	**0,634**	0,641	0,641	0,646	**0,667**
|C1| CHI&Acc2 (g)		0,512	0,570	**0,602**	0,616	0,630	0,636	**0,658**	0,660	**0,669**
|C1| IG&DF (g)		0,465	0,576	0,587	0,611	0,622	0,638	0,639	0,642	0,656
|C1| IG&Acc2 (g)		**0,521**	0,561	0,578	0,620	**0,633**	0,636	**0,659**	0,668	**0,673**
|C1| DF&Acc2 (g)		**0,532**	**0,593**	0,595	0,601	0,612	0,625	0,633	0,636	0,633
	MAX	**0,538**	**0,593**	**0,602**	**0,629**	**0,637**	0,649	**0,659**	0,668	**0,673**
	AVERAGE	0,504	0,568	0,586	0,615	0,625	0,636	0,646	0,653	0,659
Micro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (g)		0,403	0,535	0,564	0,596	0,606	0,618	0,646	0,654	0,656
|C2| tf-idf&IG (g)		0,411	0,556	0,559	0,595	0,610	0,620	0,648	0,653	0,654
|C2| tf-idf&DF (g)		0,429	0,528	0,567	0,578	0,604	0,621	0,638	0,638	0,656
|C2| tf-idf&Acc2 (g)		**0,525**	0,567	0,571	0,592	0,618	0,619	0,644	0,645	0,658
|C2| CHI&IG (g)		0,499	0,561	0,594	0,614	**0,632**	0,640	**0,662**	0,663	0,655
|C2| CHI&DF (g)		0,452	0,568	0,570	0,593	0,619	0,627	0,634	0,637	0,633
|C2| CHI&Acc2 (g)		0,511	0,568	0,575	0,610	0,622	0,641	0,646	0,642	**0,673**
|C2| IG&DF (g)		0,500	0,564	0,587	0,591	0,611	0,621	0,626	0,635	0,634
|C2| IG&Acc2 (g)		**0,521**	0,580	0,576	0,619	0,622	0,639	0,649	0,644	0,666
|C2| DF&Acc2 (g)		0,514	0,578	**0,603**	0,593	0,611	0,627	0,621	0,629	0,636
	MAX	**0,525**	0,580	**0,603**	0,619	**0,632**	0,641	**0,662**	0,663	**0,673**
	AVERAGE	0,476	0,561	0,577	0,598	0,616	0,627	0,641	0,644	0,652
Micro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (g)		**0,526**	0,568	0,594	**0,621**	**0,632**	0,653	**0,661**	0,665	0,661
|C3| tf-idf&IG (g)		0,453	0,554	**0,601**	0,618	0,627	0,642	**0,655**	0,666	0,663
|C3| tf-idf&DF (g)		0,504	0,543	0,567	0,595	0,605	0,627	0,624	0,641	0,648
|C3| tf-idf&Acc2 (g)		**0,527**	0,571	0,578	0,608	**0,632**	0,634	0,645	0,660	0,663
|C3| CHI&IG (g)		0,499	0,560	**0,604**	0,616	0,623	0,646	**0,661**	0,665	0,655
|C3| CHI&DF (g)		**0,526**	0,565	0,595	0,614	0,628	0,654	0,647	0,668	**0,669**
|C3| CHI&Acc2 (g)		0,511	0,565	**0,607**	0,620	0,628	0,644	**0,663**	0,662	0,660
|C3| IG&DF (g)		**0,532**	0,557	0,584	0,607	0,623	0,635	0,640	0,654	0,662
|C3| IG&Acc2 (g)		**0,521**	0,561	0,578	**0,626**	**0,631**	0,638	**0,657**	0,666	**0,677**
|C3| DF&Acc2 (g)		**0,535**	0,565	0,589	0,603	0,611	0,629	0,641	0,633	0,640
	MAX	**0,535**	0,571	**0,607**	**0,626**	**0,632**	0,654	**0,663**	0,668	**0,677**
	AVERAGE	0,513	0,561	0,590	0,613	0,624	0,640	0,649	0,658	0,660

Micro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (g)		0,501	0,557	0,591	0,613	0,637	0,641	0,652	0,671	0,649
|C4| tf-idf&IG (g)		0,453	0,562	0,571	0,619	0,621	0,636	0,661	0,656	0,659
|C4| tf-idf&DF (g)		0,513	0,547	0,562	0,583	0,602	0,620	0,631	0,652	0,649
|C4| tf-idf&Acc2 (g)		0,534	0,565	0,577	0,617	0,619	0,628	0,638	0,659	0,652
|C4| CHI&IG (g)		0,499	0,560	0,590	0,618	0,631	0,638	0,661	0,666	0,662
|C4| CHI&DF (g)		0,538	0,573	0,602	0,619	0,624	0,644	0,648	0,655	0,671
|C4| CHI&Acc2 (g)		0,512	0,570	0,606	0,622	0,631	0,636	0,649	0,657	0,656
|C4| IG&DF (g)		0,538	0,572	0,586	0,606	0,626	0,628	0,642	0,647	0,665
|C4| IG&Acc2 (g)		0,521	0,562	0,578	0,619	0,633	0,639	0,657	0,667	0,677
|C4| DF&Acc2 (g)		0,535	0,589	0,588	0,602	0,621	0,635	0,628	0,635	0,642
	MAX	0,538	0,589	0,606	0,622	0,637	0,644	0,661	0,671	0,677
	AVERAGE	0,514	0,566	0,585	0,612	0,625	0,635	0,647	0,657	0,658

Micro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (g)		0,543	0,568	0,592	0,613	0,630	0,658	0,655	0,667	0,664
|C5| tf-idf&IG (g)		0,453	0,562	0,592	0,619	0,628	0,636	0,662	0,666	0,667
|C5| tf-idf&DF (g)		0,513	0,557	0,563	0,590	0,597	0,629	0,632	0,641	0,656
|C5| tf-idf&Acc2 (g)		0,527	0,573	0,581	0,611	0,626	0,635	0,640	0,657	0,661
|C5| CHI&IG (g)		0,499	0,560	0,604	0,617	0,625	0,642	0,663	0,665	0,664
|C5| CHI&DF (g)		0,526	0,574	0,592	0,611	0,622	0,655	0,638	0,664	0,674
|C5| CHI&Acc2 (g)		0,511	0,570	0,607	0,617	0,628	0,638	0,657	0,664	0,659
|C5| IG&DF (g)		0,535	0,566	0,565	0,601	0,621	0,632	0,642	0,648	0,665
|C5| IG&Acc2 (g)		0,521	0,561	0,578	0,629	0,632	0,633	0,657	0,666	0,675
|C5| DF&Acc2 (g)		0,534	0,578	0,588	0,595	0,611	0,628	0,635	0,628	0,642
	MAX	0,543	0,578	0,607	0,629	0,632	0,658	0,663	0,667	0,675
	AVERAGE	0,516	0,567	0,586	0,610	0,622	0,639	0,648	0,657	0,663

Micro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (g)		0,543	0,558	0,597	0,624	0,634	0,654	0,653	0,664	0,660
|C6| tf-idf&IG (g)		0,453	0,558	0,592	0,624	0,619	0,642	0,658	0,663	0,660
|C6| tf-idf&DF (g)		0,513	0,541	0,562	0,585	0,608	0,621	0,629	0,642	0,647
|C6| tf-idf&Acc2 (g)		0,534	0,565	0,580	0,618	0,627	0,632	0,639	0,661	0,660
|C6| CHI&IG (g)		0,499	0,560	0,601	0,621	0,631	0,641	0,664	0,662	0,668
|C6| CHI&DF (g)		0,539	0,581	0,561	0,633	0,622	0,657	0,641	0,667	0,664
|C6| CHI&Acc2 (g)		0,512	0,569	0,609	0,621	0,630	0,642	0,657	0,662	0,665
|C6| IG&DF (g)		0,465	0,567	0,578	0,610	0,627	0,630	0,641	0,648	0,665
|C6| IG&Acc2 (g)		0,521	0,561	0,578	0,618	0,629	0,637	0,656	0,669	0,675
|C6| DF&Acc2 (g)		0,535	0,587	0,592	0,593	0,628	0,628	0,635	0,635	0,629
	MAX	0,543	0,587	0,609	0,633	0,634	0,657	0,664	0,669	0,675
	AVERAGE	0,511	0,565	0,585	0,615	0,625	0,639	0,647	0,657	0,659

Micro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (g)		0,543	0,568	0,591	0,628	0,638	0,654	0,661	0,657	0,657
|C7| tf-idf&IG (g)		0,453	0,562	0,604	0,634	0,622	0,643	0,655	0,665	0,669
|C7| tf-idf&DF (g)		0,513	0,557	0,563	0,590	0,600	0,625	0,628	0,639	0,647
|C7| tf-idf&Acc2 (g)		0,527	0,573	0,581	0,606	0,627	0,632	0,646	0,657	0,665
|C7| CHI&IG (g)		0,499	0,560	0,606	0,617	0,623	0,650	0,665	0,665	0,662
|C7| CHI&DF (g)		0,526	0,571	0,566	0,632	0,632	0,632	0,648	0,667	0,667
|C7| CHI&Acc2 (g)		0,511	0,569	0,609	0,621	0,630	0,642	0,662	0,663	0,663
|C7| IG&DF (g)		0,535	0,565	0,578	0,606	0,632	0,639	0,644	0,647	0,668
|C7| IG&Acc2 (g)		0,521	0,561	0,578	0,621	0,629	0,638	0,656	0,664	0,677
|C7| DF&Acc2 (g)		0,535	0,582	0,588	0,606	0,617	0,628	0,638	0,632	0,634
	MAX	0,543	0,582	0,609	0,634	0,638	0,654	0,665	0,667	0,677
	AVERAGE	0,516	0,567	0,586	0,616	0,625	0,638	0,650	0,656	0,661

Table 7.22. Macro F-measure results of the proposed combinations in global policy for Hitech dataset

Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,228	0,371	0,465	0,507	0,505	0,530	0,538	0,582	0,598
CHI (g)		0,340	0,437	**0,509**	0,528	0,550	0,570	**0,610**	**0,611**	**0,605**
IG (g)		0,301	0,433	0,461	**0,538**	**0,558**	**0,572**	0,597	0,601	0,602
DF (g)		0,141	0,389	0,383	0,461	0,527	0,510	0,524	0,526	0,532
Acc2 (g)		**0,433**	**0,507**	0,496	0,521	0,522	0,567	0,582	0,567	0,603
	MAX	0,433	0,507	0,509	0,538	0,558	0,572	0,610	0,611	0,605
Macro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	0,433	0,506	0,518	0,541	0,541	0,579	0,610	0,605	0,610
	AVERAGE	0,375	0,454	0,492	0,516	0,525	0,543	0,555	0,573	0,588
Rank Combination	MAX	0,433	0,494	0,524	0,533	0,549	0,600	0,608	0,611	0,615
	AVERAGE	0,371	0,455	0,496	0,518	0,534	0,559	0,576	0,588	0,589
Macro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (g)		0,345	0,466	0,499	0,535	0,545	**0,574**	0,598	0,609	0,582
|C1| tf-idf&IG (g)		0,309	0,467	0,496	0,537	0,536	0,542	0,605	0,602	0,591
|C1| tf-idf&DF (g)		0,350	0,394	0,448	0,499	0,509	0,527	0,519	0,541	0,546
|C1| tf-idf&Acc2 (g)		0,422	0,469	0,483	**0,542**	0,531	0,539	0,562	0,605	0,599
|C1| CHI&IG (g)		0,351	0,468	0,501	0,526	0,540	**0,604**	0,603	0,595	0,603
|C1| CHI&DF (g)		0,381	0,480	0,474	0,526	0,546	0,548	0,554	0,574	**0,607**
|C1| CHI&Acc2 (g)		0,405	0,493	**0,519**	0,534	0,541	**0,572**	0,602	0,605	**0,611**
|C1| IG&DF (g)		0,344	0,488	0,491	0,514	0,534	0,534	0,535	0,548	0,575
|C1| IG&Acc2 (g)		**0,433**	0,462	0,498	0,533	0,547	0,551	0,603	**0,612**	**0,614**
|C1| DF&Acc2 (g)		0,404	0,492	**0,510**	0,509	0,533	0,542	0,530	0,543	0,532
	MAX	0,433	0,493	**0,519**	0,542	0,547	0,604	0,605	0,612	**0,614**
	AVERAGE	0,375	0,468	0,492	0,526	0,536	0,553	0,571	0,583	0,586
Macro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (g)		0,256	0,422	0,482	0,498	0,514	0,525	0,535	0,594	0,591
|C2| tf-idf&IG (g)		0,261	0,467	0,474	0,497	0,510	0,528	0,544	0,593	0,593
|C2| tf-idf&DF (g)		0,268	0,373	0,407	0,492	0,515	0,518	0,532	0,546	0,592
|C2| tf-idf&Acc2 (g)		0,418	0,468	0,476	0,510	0,522	0,530	0,549	0,586	0,597
|C2| CHI&IG (g)		0,351	0,469	0,501	0,523	0,547	**0,579**	0,609	0,603	0,593
|C2| CHI&DF (g)		0,302	0,407	0,408	0,503	0,520	0,527	0,531	0,536	0,537
|C2| CHI&Acc2 (g)		0,403	0,492	0,492	0,522	0,543	0,561	0,582	0,570	**0,613**
|C2| IG&DF (g)		0,351	0,403	0,465	0,502	0,522	0,523	0,524	0,530	0,559
|C2| IG&Acc2 (g)		**0,433**	0,503	0,494	0,536	0,541	0,551	0,586	0,594	**0,607**
|C2| DF&Acc2 (g)		0,403	0,486	**0,510**	0,511	0,529	0,532	0,525	0,536	0,566
	MAX	0,433	0,503	**0,510**	0,536	0,547	**0,579**	0,609	0,603	**0,613**
	AVERAGE	0,345	0,449	0,471	0,509	0,526	0,537	0,552	0,569	0,585
Macro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (g)		0,372	0,480	0,506	0,532	0,543	**0,598**	0,599	0,600	0,595
|C3| tf-idf&IG (g)		0,309	0,460	0,501	0,518	0,541	0,542	0,601	**0,611**	0,599
|C3| tf-idf&DF (g)		0,354	0,388	0,448	0,507	0,510	0,528	0,518	0,539	0,545
|C3| tf-idf&Acc2 (g)		0,371	0,477	0,481	0,515	0,539	0,543	0,565	0,609	0,594
|C3| CHI&IG (g)		0,351	0,468	**0,513**	0,533	0,537	**0,599**	0,602	0,606	0,590
|C3| CHI&DF (g)		0,372	0,471	0,501	0,516	0,528	0,552	0,576	0,606	**0,609**
|C3| CHI&Acc2 (g)		0,403	0,488	**0,524**	0,532	0,549	**0,593**	0,608	0,601	0,597
|C3| IG&DF (g)		0,404	0,427	0,486	0,503	0,531	0,534	0,542	0,580	0,603
|C3| IG&Acc2 (g)		**0,433**	0,462	0,498	0,531	**0,553**	0,554	0,605	0,609	**0,615**
|C3| DF&Acc2 (g)		0,379	0,431	0,506	0,517	0,520	0,545	0,541	0,529	0,537
	MAX	**0,433**	0,488	**0,524**	0,533	**0,553**	**0,599**	0,608	**0,611**	**0,615**
	AVERAGE	0,375	0,455	0,496	0,520	0,535	0,559	0,576	0,589	0,588

Macro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (g)		0,352	0,462	0,497	0,527	0,538	0,560	0,602	0,613	0,577
|C4| tf-idf&IG (g)		0,309	0,462	0,479	0,513	0,533	0,537	0,610	0,595	0,592
|C4| tf-idf&DF (g)		0,361	0,392	0,445	0,497	0,512	0,520	0,527	0,544	0,560
|C4| tf-idf&Acc2 (g)		0,422	0,470	0,481	0,533	0,526	0,536	0,541	0,608	0,587
|C4| CHI&IG (g)		0,351	0,468	0,501	0,526	0,541	0,594	0,604	0,605	0,596
|C4| CHI&DF (g)		0,381	0,486	0,498	0,528	0,532	0,547	0,561	0,597	0,609
|C4| CHI&Acc2 (g)		0,405	0,493	0,514	0,547	0,543	0,578	0,598	0,601	0,594
|C4| IG&DF (g)		0,410	0,482	0,496	0,509	0,529	0,529	0,542	0,571	0,606
|C4| IG&Acc2 (g)		0,433	0,463	0,498	0,533	0,547	0,553	0,608	0,610	0,617
|C4| DF&Acc2 (g)		0,379	0,489	0,501	0,511	0,539	0,550	0,534	0,538	0,570
	MAX	0,433	0,493	0,514	0,547	0,547	0,594	0,610	0,613	0,617
	AVERAGE	0,380	0,467	0,491	0,522	0,534	0,550	0,573	0,588	0,591

Macro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (g)		0,385	0,480	0,493	0,524	0,531	0,587	0,594	0,608	0,598
|C5| tf-idf&IG (g)		0,309	0,463	0,496	0,524	0,542	0,541	0,607	0,603	0,602
|C5| tf-idf&DF (g)		0,361	0,400	0,446	0,503	0,503	0,529	0,521	0,536	0,557
|C5| tf-idf&Acc2 (g)		0,371	0,477	0,483	0,522	0,534	0,542	0,562	0,604	0,599
|C5| CHI&IG (g)		0,351	0,468	0,513	0,537	0,537	0,598	0,605	0,600	0,597
|C5| CHI&DF (g)		0,372	0,482	0,499	0,495	0,520	0,558	0,586	0,601	0,616
|C5| CHI&Acc2 (g)		0,403	0,493	0,524	0,531	0,546	0,581	0,601	0,605	0,601
|C5| IG&DF (g)		0,408	0,469	0,475	0,499	0,529	0,535	0,545	0,574	0,608
|C5| IG&Acc2 (g)		0,433	0,462	0,498	0,541	0,554	0,549	0,606	0,612	0,614
|C5| DF&Acc2 (g)		0,377	0,440	0,506	0,508	0,524	0,545	0,530	0,530	0,568
	MAX	0,433	0,493	0,524	0,541	0,554	0,598	0,607	0,612	0,616
	AVERAGE	0,377	0,463	0,493	0,518	0,532	0,556	0,576	0,587	0,596

Macro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (g)		0,385	0,466	0,496	0,531	0,538	0,592	0,591	0,608	0,592
|C6| tf-idf&IG (g)		0,309	0,458	0,496	0,537	0,535	0,548	0,604	0,604	0,593
|C6| tf-idf&DF (g)		0,361	0,387	0,445	0,499	0,511	0,524	0,520	0,540	0,547
|C6| tf-idf&Acc2 (g)		0,422	0,470	0,483	0,542	0,534	0,543	0,561	0,609	0,598
|C6| CHI&IG (g)		0,351	0,468	0,507	0,528	0,541	0,599	0,606	0,602	0,600
|C6| CHI&DF (g)		0,381	0,489	0,466	0,533	0,524	0,559	0,549	0,608	0,606
|C6| CHI&Acc2 (g)		0,405	0,492	0,523	0,537	0,541	0,584	0,600	0,604	0,605
|C6| IG&DF (g)		0,344	0,472	0,483	0,515	0,538	0,539	0,541	0,570	0,609
|C6| IG&Acc2 (g)		0,433	0,462	0,498	0,532	0,544	0,551	0,604	0,613	0,616
|C6| DF&Acc2 (g)		0,379	0,487	0,508	0,508	0,548	0,545	0,535	0,538	0,528
	MAX	0,433	0,492	0,523	0,542	0,548	0,599	0,606	0,613	0,616
	AVERAGE	0,377	0,465	0,490	0,526	0,535	0,558	0,571	0,589	0,589

Macro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (g)		0,385	0,480	0,492	0,535	0,542	0,603	0,602	0,602	0,591
|C7| tf-idf&IG (g)		0,309	0,463	0,506	0,549	0,539	0,547	0,601	0,609	0,604
|C7| tf-idf&DF (g)		0,361	0,400	0,446	0,503	0,506	0,527	0,522	0,535	0,545
|C7| tf-idf&Acc2 (g)		0,371	0,477	0,483	0,513	0,535	0,542	0,566	0,607	0,597
|C7| CHI&IG (g)		0,351	0,468	0,517	0,537	0,537	0,603	0,606	0,600	0,595
|C7| CHI&DF (g)		0,372	0,481	0,473	0,533	0,534	0,534	0,550	0,606	0,607
|C7| CHI&Acc2 (g)		0,403	0,492	0,523	0,537	0,538	0,585	0,608	0,604	0,601
|C7| IG&DF (g)		0,408	0,468	0,489	0,511	0,548	0,544	0,552	0,571	0,615
|C7| IG&Acc2 (g)		0,433	0,462	0,498	0,534	0,545	0,546	0,606	0,611	0,616
|C7| DF&Acc2 (g)		0,379	0,484	0,506	0,520	0,527	0,544	0,534	0,528	0,531
	MAX	0,433	0,492	0,523	0,549	0,548	0,603	0,608	0,611	0,616
	AVERAGE	0,377	0,467	0,493	0,527	0,535	0,558	0,575	0,587	0,590

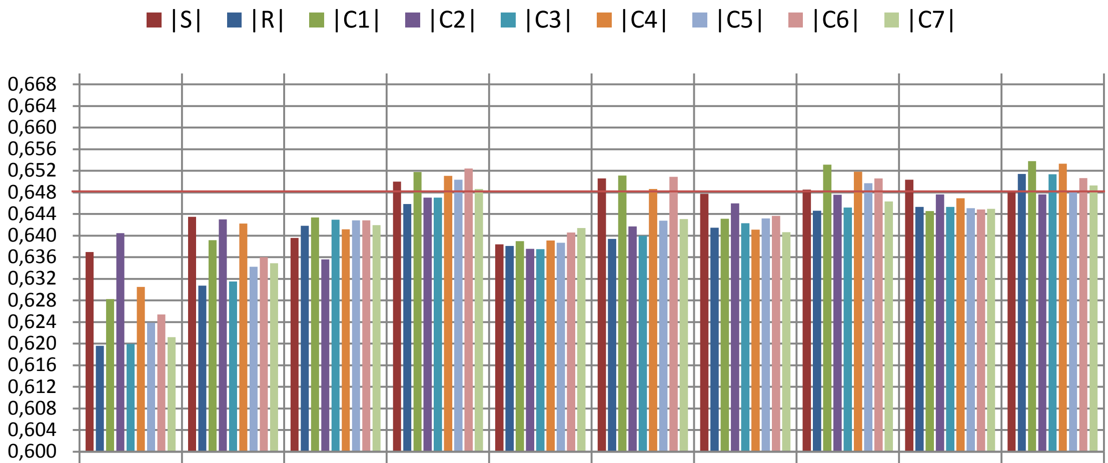
Before perform our proposed methods in local policy on the Hitech dataset, we summarize what we learnt from the previous experiments.

When we evaluated the performance of the individual feature selection metrics in local policy, *Acc2* and *DF* outperforms other metrics. Acc2 achieved the highest results when keyword numbers are lower than 1500 and DF achieved the highest results when the keyword numbers are higher than 1000. Although CHI was one of the most successful metrics in global policy, it was the worst metric in local policy.

When we performed score and rank combinations on the Hitech dataset in local policy, score combination was significantly better than the rank combination when the number of keywords was high but rank combination was more successful than score combination when the number of keywords was between 10 and 50. Furthermore among the 10 possible combinations of two feature selection metrics *tf-idf & IG, tf-idf & Acc2, CHI & Acc2, IG & DF* and *IG & Acc2* combinations were more successful than other combinations in the previous experiments. Finally we concluded that *CHI & Acc2 and tf-idf & Acc2* excel the success of the best individual feature selection metric *Acc2.*

Tables 7.25 and 7.26 show the micro- and macro-averaged F-measure results, respectively, for all seven proposed combination methods in local policy for the Hitech dataset. Firstly, we determine which combinations are better and which are worse than others. Among the 10 possible combinations of two feature selection metrics, the best combinations are *DF & Acc2, tf-idf & Acc2, IG & DF* and *CHI & DF.* On the other hand the worst combinations are *tf-idf & CHI, tf-idf & IG* and *CHI & IG* as seen in Figures 7.18 and 7.19. One of the observations from these experiments is that the success of the combination depends on the performance of the individual metrics which means the combination of two feature selection metrics can achieve better performance only if the performance of each individual feature selection metric is successful.

Secondly, we discuss the experimental results of the proposed combination methods. In the case of local policy rank combination was outperformed by score combination when we compared score and rank combinations in the previous section. The experiments of the proposed methods show that all proposed methods outperforms the rank combination and among these methods *C2, C4, C5* and *C6* improve the success of the rank combination when we look at the micro-averaged F-measure values and *C1, C3, C4, C5, C6* and *C7* improve the success of the rank combination when we look at the macro-averaged F-measure values. In related tables, we can see that improvement of the macro-average F-measures is more explicit than the improvement of the micro-averaged F-measures. From this observation we can conclude that the proposed methods more apparently improve the performance of the classifier on rare categories in the Hitech dataset.

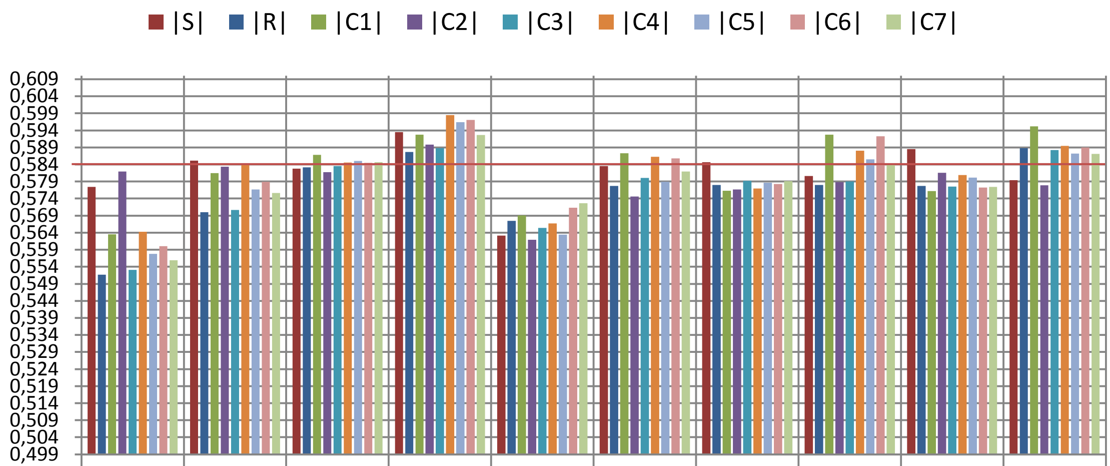


	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,637	0,643	0,640	0,650	0,638	0,651	0,648	0,649	0,650	0,648
|R|	0,620	0,631	0,642	0,646	0,638	0,639	0,641	0,645	0,645	0,651
|C1|	0,628	0,639	0,643	0,652	0,639	0,651	0,643	0,653	0,645	0,654
|C2|	0,640	0,643	0,636	0,647	0,638	0,642	0,646	0,648	0,648	0,648
|C3|	0,620	0,631	0,643	0,647	0,637	0,640	0,642	0,645	0,645	0,651
|C4|	0,630	0,642	0,641	0,651	0,639	0,649	0,641	0,652	0,647	0,653
|C5|	0,624	0,634	0,643	0,650	0,639	0,643	0,643	0,650	0,645	0,648
|C6|	0,625	0,636	0,643	0,652	0,641	0,651	0,644	0,651	0,645	0,651
|C7|	0,621	0,635	0,642	0,649	0,641	0,643	0,641	0,646	0,645	0,649

Figure 7.18. In local policy, averages of the micro-averaged F-measures of all combinations for Hitech dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	0,560	0,621	0,642	**0,647**	**0,660**	**0,662**	**0,665**	**0,661**	**0,661**
Combination | R |	**0,568**	**0,622**	**0,647**	0,646	0,654	0,661	0,661	0,656	0,658
Combination |C1|	0,560	0,625	0,639	**0,656**	**0,660**	0,661	0,659	0,664	**0,656**
Combination |C2|	**0,566**	0,621	**0,647**	0,648	**0,660**	0,657	**0,670**	**0,669**	**0,656**
Combination |C3|	**0,568**	0,622	**0,653**	0,648	**0,659**	0,661	0,662	0,656	**0,656**
Combination |C4|	0,561	0,621	0,642	**0,659**	0,656	**0,668**	**0,669**	0,662	**0,656**
Combination |C5|	0,561	**0,629**	0,640	0,653	**0,659**	**0,669**	0,657	0,662	0,652
Combination |C6|	0,561	0,624	0,641	0,650	0,654	0,663	0,662	**0,665**	0,655
Combination |C7|	0,561	**0,627**	0,644	0,646	0,651	0,665	0,663	0,657	**0,659**

Table 7.23. In local policy, maximum micro-averaged F-measure of all combinations for Hitech dataset



	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,577	0,585	0,583	0,593	0,563	0,584	0,585	0,581	0,589	0,579
|R|	0,552	0,570	0,583	0,588	0,568	0,578	0,578	0,578	0,578	0,589
|C1|	0,564	0,581	0,587	0,593	0,569	0,587	0,576	0,593	0,576	0,595
|C2|	0,582	0,583	0,582	0,590	0,562	0,575	0,577	0,579	0,582	0,578
|C3|	0,553	0,571	0,584	0,589	0,565	0,580	0,579	0,579	0,578	0,588
|C4|	0,564	0,584	0,585	0,598	0,567	0,586	0,577	0,588	0,581	0,589
|C5|	0,558	0,577	0,585	0,596	0,563	0,579	0,579	0,586	0,580	0,587
|C6|	0,560	0,579	0,584	0,597	0,571	0,586	0,578	0,592	0,577	0,589
|C7|	0,556	0,576	0,585	0,593	0,573	0,582	0,579	0,584	0,577	0,587

Figure 7.19. In local policy, averages of the macro-averaged F-measures of all combinations for Hitech dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	0,475	0,565	0,582	**0,595**	**0,610**	0,601	**0,604**	0,595	**0,598**
Combination | R |	**0,484**	**0,566**	**0,588**	0,594	0,606	**0,605**	0,603	**0,601**	0,597
Combination |C1|	0,501	0,572	0,591	0,601	0,610	0,607	0,604	0,614	0,596
Combination |C2|	0,478	0,563	0,574	0,593	0,603	0,603	0,612	0,600	0,595
Combination |C3|	0,484	0,566	0,598	0,600	0,610	0,606	0,604	0,600	0,596
Combination |C4|	0,485	0,565	0,582	0,610	0,610	0,620	0,611	0,599	0,593
Combination |C5|	0,479	0,576	0,578	0,603	0,609	0,613	0,601	0,607	0,593
Combination |C6|	0,480	0,567	0,580	0,600	0,609	0,613	0,600	0,610	0,596
Combination |C7|	0,478	0,573	0,591	0,593	0,605	0,615	0,600	0,602	0,600

Table 7.24. In local policy, maximum macro-averaged F-measure of all combinations for Hitech dataset

Among the proposed methods C4 shows the most notable performance both in the case of micro- and macro-averaged F-measure results. C2 is also very successful in the case of micro-average measure whereas it cannot achieve good result in the case of macro-averaged F-measure. C3 is especially very successful when the keyword numbers are less than 100.

A comparison of the proposed methods with the score and rank combinations also indicates that all proposed methods apparently improve the highest micro- and macro-averaged F-measure results of the individual metrics. One of the conclusions from the previous experiments was that the highest scores were achieved by *DF* with value of 66.4% micro- and 61.1% macro-averaged F-measure value with 1500 keywords and the score combination of *IG & Acc2* increased the micro-averaged F-measure to 66.5% with 1000 keywords and the score combination of *tf-idf & Acc2* approached the highest macro-averaged F-measure with the 61% by using only 200 keywords. Almost all proposed combination (C2, C4, C5 and C7) improves this final highest micro-averaged F-measure and all proposed combinations improves macro-averaged F-measure as seen in Tables 7.23 and 7.24. Furthermore the most successful performances are achieved by C4 of *tf-idf & Acc2* with value of 66.8% micro- and 62% macro-averaged F-measure value with 500 keywords. This deduction is one of the good results of the experiments since we get more accurate classification with less number of keywords.

Table 7.25. Micro F-measure results of the proposed combinations in local policy for Hitech dataset

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (l)		0,551	0,596	0,613	0,627	0,624	0,644	0,621	0,618	0,627
CHI (l)		0,557	0,590	0,620	0,631	0,636	0,636	0,619	0,630	0,632
IG (l)		0,510	0,610	0,617	0,638	0,630	**0,654**	0,644	0,634	0,638
DF (l)		0,501	0,550	0,578	0,624	0,613	0,622	0,644	**0,664**	**0,661**
Acc2 (l)		**0,558**	**0,612**	**0,636**	**0,649**	**0,637**	0,651	**0,659**	0,647	0,646
	MAX	0,558	0,612	0,636	0,649	0,637	0,654	0,659	0,664	0,661
Macro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	**0,560**	**0,621**	**0,642**	0,647	**0,660**	**0,662**	0,665	**0,661**	**0,661**
	AVERAGE	0,542	0,603	0,621	0,641	0,642	0,650	0,652	0,646	0,642
Rank Combination	MAX	**0,568**	**0,622**	**0,647**	0,646	**0,654**	**0,661**	0,661	**0,656**	**0,658**
	AVERAGE	0,542	0,607	0,626	0,633	0,639	0,648	0,645	0,638	0,636
Micro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (l)		0,554	**0,625**	0,605	0,638	0,632	0,623	0,627	0,627	0,622
|C1| tf-idf&IG (l)		0,540	0,609	0,632	**0,650**	**0,646**	0,641	0,638	0,632	0,627
|C1| tf-idf&DF (l)		0,538	0,585	0,594	0,632	0,630	**0,659**	0,648	0,651	0,641
|C1| tf-idf&Acc2 (l)		0,529	0,608	0,622	0,645	**0,660**	**0,661**	**0,659**	0,641	0,645
|C1| CHI&IG (l)		0,557	0,606	0,625	0,641	0,635	0,651	0,631	0,638	0,639
|C1| CHI&DF (l)		0,543	0,608	0,633	**0,656**	**0,654**	**0,658**	0,656	0,644	0,638
|C1| CHI&Acc2 (l)		**0,560**	**0,617**	0,634	**0,649**	0,626	**0,657**	0,648	0,641	0,637
|C1| IG&DF (l)		0,540	**0,622**	**0,638**	0,647	**0,659**	**0,661**	0,657	0,647	0,649
|C1| IG&Acc2 (l)		0,557	0,603	**0,639**	0,629	**0,644**	0,653	0,653	0,649	0,641
|C1| DF&Acc2 (l)		0,528	0,612	0,635	**0,647**	**0,648**	0,649	0,658	**0,664**	0,656
	MAX	**0,560**	**0,625**	**0,639**	**0,656**	**0,660**	**0,661**	**0,659**	**0,664**	0,656
	AVERAGE	0,545	0,610	0,626	0,643	0,644	0,651	0,647	0,643	0,639
Micro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (l)		0,540	0,596	0,614	0,642	**0,640**	0,650	0,643	0,627	0,641
|C2| tf-idf&IG (l)		0,539	0,593	0,609	0,647	**0,653**	**0,656**	0,645	0,622	0,635
|C2| tf-idf&DF (l)		0,505	0,575	0,603	0,628	0,623	0,635	0,645	0,638	0,644
|C2| tf-idf&Acc2 (l)		0,534	0,610	0,626	0,643	**0,660**	**0,657**	0,650	0,640	0,633
|C2| CHI&IG (l)		0,544	0,602	0,610	**0,648**	0,635	0,652	0,629	0,635	0,627
|C2| CHI&DF (l)		0,532	0,581	0,603	0,619	0,636	0,640	0,647	**0,664**	0,644
|C2| CHI&Acc2 (l)		**0,566**	**0,621**	**0,647**	0,638	**0,644**	0,653	0,657	0,649	0,633
|C2| IG&DF (l)		0,528	0,569	0,603	0,631	0,637	0,645	0,656	**0,663**	0,653
|C2| IG&Acc2 (l)		0,553	**0,619**	**0,642**	0,640	0,636	0,648	**0,670**	0,649	0,643
|C2| DF&Acc2 (l)		0,528	**0,613**	0,627	0,625	0,631	0,644	**0,661**	**0,669**	0,656
	MAX	**0,566**	**0,621**	**0,647**	**0,648**	**0,660**	**0,657**	**0,670**	**0,669**	0,656
	AVERAGE	0,537	0,598	0,618	0,636	0,640	0,648	0,650	0,646	0,641
Micro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (l)		0,554	**0,622**	0,623	0,629	0,621	0,621	0,621	0,607	0,619
|C3| tf-idf&IG (l)		0,550	**0,617**	0,621	0,629	**0,640**	0,631	0,637	0,635	0,618
|C3| tf-idf&DF (l)		0,518	0,570	0,597	0,631	**0,640**	**0,655**	0,649	0,645	0,638
|C3| tf-idf&Acc2 (l)		0,542	0,592	0,613	0,631	**0,659**	**0,655**	**0,661**	0,634	0,642
|C3| CHI&IG (l)		0,546	**0,613**	0,623	**0,648**	**0,638**	0,640	0,630	0,638	0,631
|C3| CHI&DF (l)		0,531	0,610	**0,639**	0,639	0,636	0,649	0,644	0,636	0,636
|C3| CHI&Acc2 (l)		**0,568**	**0,617**	**0,653**	0,643	0,629	**0,656**	0,643	0,639	0,643
|C3| IG&DF (l)		0,532	**0,613**	**0,637**	0,632	**0,645**	**0,661**	0,656	0,646	0,632
|C3| IG&Acc2 (l)		0,553	0,606	**0,636**	0,638	**0,641**	**0,655**	0,652	0,645	0,640
|C3| DF&Acc2 (l)		0,522	0,607	0,621	0,640	0,636	**0,658**	**0,662**	0,656	0,656
	MAX	**0,568**	**0,622**	**0,653**	**0,648**	**0,659**	**0,661**	**0,662**	0,656	0,656
	AVERAGE	0,541	0,607	0,626	0,636	0,638	0,648	0,646	0,638	0,636

Micro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (I)		0,556	0,618	0,594	0,647	0,628	0,611	0,636	0,633	0,629
|C4| tf-idf&IG (I)		0,545	0,605	0,620	0,659	0,641	0,640	0,634	0,639	0,640
|C4| tf-idf&DF (I)		0,519	0,583	0,597	0,631	0,631	0,651	0,647	0,647	0,642
|C4| tf-idf&Acc2 (I)		0,540	0,607	0,612	0,648	0,655	0,668	0,662	0,636	0,638
|C4| CHI&IG (I)		0,560	0,602	0,614	0,642	0,637	0,651	0,627	0,641	0,636
|C4| CHI&DF (I)		0,544	0,609	0,639	0,650	0,648	0,656	0,649	0,646	0,642
|C4| CHI&Acc2 (I)		0,561	0,621	0,640	0,639	0,627	0,661	0,640	0,636	0,644
|C4| IG&DF (I)		0,541	0,620	0,642	0,648	0,656	0,651	0,669	0,649	0,638
|C4| IG&Acc2 (I)		0,553	0,600	0,638	0,637	0,645	0,654	0,659	0,650	0,637
|C4| DF&Acc2 (I)		0,528	0,611	0,636	0,647	0,647	0,650	0,658	0,662	0,656
	MAX	0,561	0,621	0,642	0,659	0,656	0,668	0,669	0,662	0,656
	AVERAGE	0,545	0,608	0,623	0,645	0,641	0,649	0,648	0,644	0,640

Micro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (I)		0,555	0,629	0,618	0,630	0,635	0,617	0,627	0,619	0,615
|C5| tf-idf&IG (I)		0,545	0,613	0,623	0,627	0,644	0,633	0,639	0,637	0,625
|C5| tf-idf&DF (I)		0,516	0,582	0,598	0,625	0,637	0,660	0,648	0,645	0,641
|C5| tf-idf&Acc2 (I)		0,540	0,587	0,602	0,640	0,659	0,659	0,656	0,643	0,646
|C5| CHI&IG (I)		0,557	0,608	0,624	0,646	0,642	0,649	0,632	0,635	0,629
|C5| CHI&DF (I)		0,542	0,618	0,637	0,653	0,644	0,648	0,643	0,637	0,631
|C5| CHI&Acc2 (I)		0,561	0,614	0,639	0,638	0,629	0,660	0,644	0,642	0,647
|C5| IG&DF (I)		0,533	0,624	0,632	0,635	0,657	0,669	0,657	0,646	0,635
|C5| IG&Acc2 (I)		0,553	0,604	0,640	0,630	0,643	0,653	0,648	0,648	0,648
|C5| DF&Acc2 (I)		0,522	0,610	0,634	0,634	0,634	0,650	0,657	0,662	0,652
	MAX	0,561	0,629	0,640	0,653	0,659	0,669	0,657	0,662	0,652
	AVERAGE	0,542	0,609	0,625	0,636	0,642	0,650	0,645	0,641	0,637

Micro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (I)		0,549	0,624	0,615	0,631	0,627	0,629	0,626	0,625	0,614
|C6| tf-idf&IG (I)		0,545	0,615	0,623	0,643	0,636	0,637	0,635	0,636	0,630
|C6| tf-idf&DF (I)		0,521	0,578	0,593	0,628	0,632	0,659	0,649	0,650	0,641
|C6| tf-idf&Acc2 (I)		0,540	0,605	0,605	0,650	0,654	0,655	0,662	0,643	0,650
|C6| CHI&IG (I)		0,560	0,607	0,625	0,641	0,638	0,654	0,630	0,641	0,639
|C6| CHI&DF (I)		0,544	0,605	0,633	0,648	0,652	0,663	0,651	0,650	0,642
|C6| CHI&Acc2 (I)		0,561	0,618	0,641	0,644	0,626	0,658	0,648	0,643	0,642
|C6| IG&DF (I)		0,538	0,621	0,631	0,645	0,651	0,659	0,656	0,656	0,636
|C6| IG&Acc2 (I)		0,553	0,600	0,631	0,626	0,643	0,653	0,652	0,648	0,646
|C6| DF&Acc2 (I)		0,528	0,610	0,632	0,636	0,640	0,650	0,657	0,665	0,655
	MAX	0,561	0,624	0,641	0,650	0,654	0,663	0,662	0,665	0,655
	AVERAGE	0,544	0,608	0,623	0,639	0,640	0,652	0,647	0,646	0,639

Micro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (I)		0,549	0,627	0,616	0,629	0,633	0,619	0,626	0,610	0,610
|C7| tf-idf&IG (I)		0,551	0,618	0,628	0,634	0,639	0,635	0,644	0,636	0,622
|C7| tf-idf&DF (I)		0,516	0,581	0,599	0,626	0,638	0,648	0,650	0,647	0,641
|C7| tf-idf&Acc2 (I)		0,540	0,587	0,603	0,638	0,651	0,661	0,663	0,636	0,644
|C7| CHI&IG (I)		0,546	0,610	0,625	0,646	0,641	0,651	0,635	0,639	0,636
|C7| CHI&DF (I)		0,545	0,616	0,633	0,641	0,636	0,654	0,650	0,643	0,634
|C7| CHI&Acc2 (I)		0,561	0,617	0,644	0,632	0,627	0,655	0,645	0,642	0,641
|C7| IG&DF (I)		0,533	0,619	0,631	0,637	0,646	0,665	0,656	0,643	0,631
|C7| IG&Acc2 (I)		0,553	0,600	0,631	0,632	0,642	0,656	0,651	0,646	0,643
|C7| DF&Acc2 (I)		0,522	0,616	0,633	0,633	0,642	0,648	0,657	0,657	0,659
	MAX	0,561	0,627	0,644	0,646	0,651	0,665	0,663	0,657	0,659
	AVERAGE	0,542	0,609	0,624	0,635	0,639	0,649	0,648	0,640	0,636

Table 7.26. Macro F-measure results of the proposed combinations in local policy for Hitech dataset

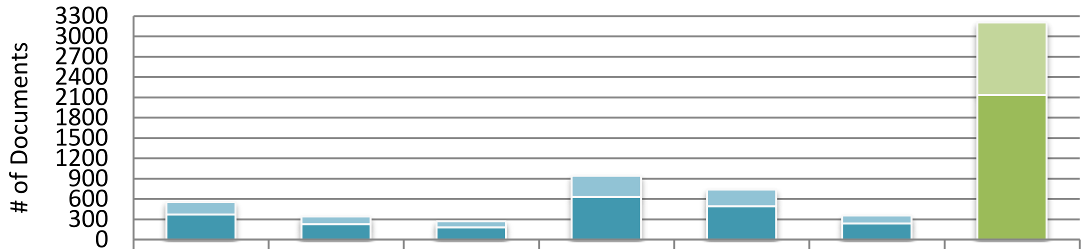
Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (l)		**0,486**	**0,555**	**0,571**	0,571	0,564	0,589	0,567	0,549	0,561
CHI (l)		0,477	0,495	0,536	0,572	0,567	0,551	0,545	0,552	0,567
IG (l)		0,456	0,529	0,539	**0,577**	**0,571**	0,591	0,573	0,555	0,557
DF (l)		0,397	0,485	0,507	0,549	0,540	0,549	0,592	**0,611**	**0,603**
Acc2 (l)		0,459	0,522	0,550	0,571	0,564	**0,596**	**0,600**	0,583	0,593
	MAX	**0,486**	0,555	0,571	0,577	0,571	0,596	0,600	**0,611**	**0,603**
Macro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	0,475	**0,565**	**0,582**	**0,595**	**0,610**	**0,601**	**0,604**	0,595	0,598
	AVERAGE	0,458	0,530	0,551	0,578	0,579	0,586	0,589	0,583	0,577
Rank Combination	MAX	0,484	**0,566**	**0,588**	**0,594**	**0,606**	**0,605**	**0,603**	0,601	0,597
	AVERAGE	0,459	0,533	0,557	0,570	0,578	0,585	0,580	0,570	0,573
Macro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (l)		**0,501**	**0,572**	0,550	**0,582**	0,560	0,547	0,557	0,574	0,561
|C1| tf-idf&IG (l)		0,458	0,551	**0,591**	**0,601**	**0,587**	0,588	0,577	0,568	0,569
|C1| tf-idf&DF (l)		0,461	0,515	0,520	**0,579**	**0,577**	**0,607**	0,593	0,588	0,577
|C1| tf-idf&Acc2 (l)		0,448	0,542	**0,571**	**0,597**	**0,610**	**0,602**	0,596	0,572	0,579
|C1| CHI&IG (l)		0,478	0,510	**0,574**	0,570	**0,573**	0,575	0,557	0,563	0,577
|C1| CHI&DF (l)		0,455	0,538	0,552	**0,599**	**0,585**	**0,601**	0,592	0,575	0,571
|C1| CHI&Acc2 (l)		0,473	0,540	0,553	**0,592**	0,563	0,593	0,573	0,561	0,576
|C1| IG&DF (l)		0,454	0,543	0,552	0,563	**0,608**	**0,605**	0,597	0,587	0,596
|C1| IG&Acc2 (l)		0,453	0,508	0,564	0,561	**0,582**	0,589	0,582	0,573	0,570
|C1| DF&Acc2 (l)		0,449	0,531	0,558	**0,579**	**0,584**	0,595	**0,604**	**0,614**	0,596
	MAX	**0,501**	**0,572**	**0,591**	**0,601**	**0,610**	**0,607**	**0,604**	**0,614**	0,596
	AVERAGE	0,463	0,535	0,558	0,582	0,583	0,590	0,583	0,577	0,577
Macro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (l)		0,464	0,539	0,554	**0,590**	**0,577**	0,578	0,580	0,581	0,587
|C2| tf-idf&IG (l)		0,457	0,540	0,559	**0,593**	**0,587**	0,593	0,593	0,564	0,570
|C2| tf-idf&DF (l)		0,407	0,504	0,526	0,577	0,571	0,581	0,595	0,582	0,585
|C2| tf-idf&Acc2 (l)		0,458	**0,563**	**0,574**	**0,592**	**0,603**	**0,603**	0,592	0,580	0,569
|C2| CHI&IG (l)		0,453	0,508	0,528	**0,577**	**0,571**	0,569	0,552	0,555	0,548
|C2| CHI&DF (l)		0,458	0,505	0,522	0,543	0,565	0,583	0,584	0,598	0,576
|C2| CHI&Acc2 (l)		0,478	0,546	0,569	0,565	**0,576**	0,589	0,590	0,576	0,563
|C2| IG&DF (l)		0,450	0,498	0,524	0,552	0,566	0,575	0,592	0,599	0,589
|C2| IG&Acc2 (l)		0,452	0,533	0,550	0,568	0,568	0,584	**0,612**	0,578	0,579
|C2| DF&Acc2 (l)		0,446	0,544	0,541	0,556	0,556	0,569	0,593	0,600	0,595
	MAX	0,478	0,563	0,574	0,593	0,603	0,603	**0,612**	0,600	0,595
	AVERAGE	0,452	0,528	0,545	0,571	0,574	0,582	0,588	0,581	0,576
Macro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (l)		0,479	**0,566**	0,560	0,565	0,539	0,555	0,553	0,550	0,556
|C3| tf-idf&IG (l)		0,477	**0,565**	**0,576**	0,572	**0,588**	0,560	0,567	0,576	0,561
|C3| tf-idf&DF (l)		0,447	0,502	0,528	0,565	0,570	**0,606**	**0,601**	0,583	0,576
|C3| tf-idf&Acc2 (l)		0,464	0,519	0,544	**0,565**	**0,610**	**0,598**	0,599	0,576	0,584
|C3| CHI&IG (l)		0,469	0,524	**0,572**	**0,586**	**0,574**	0,562	0,546	0,561	0,564
|C3| CHI&DF (l)		0,443	0,532	0,565	**0,600**	**0,587**	0,587	0,573	0,560	0,574
|C3| CHI&Acc2 (l)		0,484	0,546	**0,598**	**0,599**	0,564	0,594	0,573	0,561	0,584
|C3| IG&DF (l)		0,435	0,531	0,557	0,550	**0,595**	**0,600**	0,585	0,571	0,574
|C3| IG&Acc2 (l)		0,452	0,524	0,552	0,571	**0,584**	0,593	0,588	0,569	0,560
|C3| DF&Acc2 (l)		0,444	0,523	0,540	0,562	0,567	**0,599**	**0,604**	0,600	0,596
	MAX	0,484	0,566	**0,598**	**0,600**	0,610	**0,606**	**0,604**	0,600	0,596
	AVERAGE	0,459	0,533	0,559	0,573	0,578	0,586	0,579	0,571	0,573

Macro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (I)		0,485	0,565	0,540	0,590	0,550	0,538	0,565	0,575	0,569
|C4| tf-idf&IG (I)		0,472	0,545	0,582	0,610	0,582	0,576	0,573	0,581	0,582
|C4| tf-idf&DF (I)		0,455	0,513	0,527	0,579	0,574	0,598	0,594	0,584	0,579
|C4| tf-idf&Acc2 (I)		0,457	0,544	0,563	0,597	0,610	0,620	0,611	0,584	0,570
|C4| CHI&IG (I)		0,480	0,503	0,535	0,575	0,579	0,577	0,548	0,567	0,554
|C4| CHI&DF (I)		0,463	0,534	0,563	0,579	0,590	0,592	0,588	0,580	0,588
|C4| CHI&Acc2 (I)		0,477	0,543	0,562	0,587	0,567	0,593	0,567	0,561	0,587
|C4| IG&DF (I)		0,450	0,539	0,558	0,568	0,603	0,598	0,597	0,588	0,575
|C4| IG&Acc2 (I)		0,452	0,506	0,564	0,567	0,583	0,594	0,591	0,584	0,567
|C4| DF&Acc2 (I)		0,449	0,530	0,551	0,578	0,579	0,593	0,596	0,599	0,593
	MAX	0,485	0,565	0,582	0,610	0,610	0,620	0,611	0,599	0,593
	AVERAGE	0,464	0,532	0,554	0,583	0,582	0,588	0,583	0,580	0,576
Macro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (I)		0,479	0,576	0,561	0,561	0,563	0,546	0,559	0,562	0,555
|C5| tf-idf&IG (I)		0,472	0,556	0,578	0,584	0,586	0,578	0,567	0,578	0,567
|C5| tf-idf&DF (I)		0,446	0,511	0,529	0,555	0,593	0,613	0,595	0,573	0,581
|C5| tf-idf&Acc2 (I)		0,453	0,516	0,536	0,589	0,609	0,610	0,601	0,585	0,584
|C5| CHI&IG (I)		0,478	0,520	0,568	0,579	0,582	0,566	0,547	0,559	0,547
|C5| CHI&DF (I)		0,453	0,538	0,551	0,603	0,584	0,590	0,570	0,557	0,569
|C5| CHI&Acc2 (I)		0,477	0,539	0,565	0,587	0,566	0,595	0,572	0,566	0,587
|C5| IG&DF (I)		0,444	0,543	0,549	0,567	0,604	0,599	0,588	0,574	0,580
|C5| IG&Acc2 (I)		0,452	0,522	0,565	0,565	0,587	0,591	0,576	0,582	0,579
|C5| DF&Acc2 (I)		0,444	0,526	0,554	0,559	0,568	0,597	0,599	0,607	0,593
	MAX	0,479	0,576	0,578	0,603	0,609	0,613	0,601	0,607	0,593
	AVERAGE	0,460	0,535	0,555	0,575	0,584	0,589	0,577	0,574	0,574
Macro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (I)		0,474	0,567	0,561	0,567	0,557	0,556	0,558	0,572	0,549
|C6| tf-idf&IG (I)		0,472	0,552	0,580	0,600	0,578	0,576	0,571	0,580	0,569
|C6| tf-idf&DF (I)		0,457	0,506	0,512	0,555	0,580	0,613	0,591	0,588	0,581
|C6| tf-idf&Acc2 (I)		0,457	0,542	0,538	0,600	0,609	0,603	0,600	0,583	0,587
|C6| CHI&IG (I)		0,480	0,505	0,573	0,575	0,580	0,580	0,552	0,564	0,577
|C6| CHI&DF (I)		0,463	0,531	0,558	0,593	0,586	0,598	0,587	0,576	0,574
|C6| CHI&Acc2 (I)		0,477	0,541	0,568	0,591	0,565	0,594	0,578	0,562	0,579
|C6| IG&DF (I)		0,453	0,541	0,547	0,575	0,603	0,597	0,591	0,601	0,587
|C6| IG&Acc2 (I)		0,452	0,506	0,559	0,562	0,584	0,588	0,580	0,574	0,575
|C6| DF&Acc2 (I)		0,449	0,525	0,549	0,565	0,570	0,597	0,597	0,610	0,596
	MAX	0,480	0,567	0,580	0,600	0,609	0,613	0,600	0,610	0,596
	AVERAGE	0,463	0,531	0,554	0,578	0,581	0,590	0,580	0,581	0,577
Macro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (I)		0,474	0,573	0,557	0,563	0,564	0,543	0,559	0,558	0,549
|C7| tf-idf&IG (I)		0,478	0,559	0,581	0,584	0,587	0,567	0,572	0,582	0,562
|C7| tf-idf&DF (I)		0,446	0,509	0,528	0,557	0,584	0,597	0,598	0,590	0,581
|C7| tf-idf&Acc2 (I)		0,453	0,516	0,540	0,588	0,605	0,615	0,600	0,571	0,577
|C7| CHI&IG (I)		0,469	0,506	0,573	0,579	0,586	0,575	0,560	0,562	0,575
|C7| CHI&DF (I)		0,467	0,536	0,554	0,593	0,574	0,595	0,583	0,571	0,575
|C7| CHI&Acc2 (I)		0,477	0,540	0,591	0,589	0,565	0,594	0,579	0,564	0,583
|C7| IG&DF (I)		0,443	0,543	0,553	0,564	0,599	0,599	0,588	0,578	0,577
|C7| IG&Acc2 (I)		0,452	0,518	0,560	0,561	0,583	0,595	0,584	0,573	0,568
|C7| DF&Acc2 (I)		0,444	0,532	0,551	0,558	0,573	0,590	0,600	0,602	0,600
	MAX	0,478	0,573	0,591	0,593	0,605	0,615	0,600	0,602	0,600
	AVERAGE	0,460	0,533	0,559	0,574	0,582	0,587	0,582	0,575	0,575

7.2.2. The LA1 Dataset

7.2.2.1. Property of the Dataset

In the study, we try to choose datasets that are popular and commonly used in the literature. Our third dataset LA1 is one of the well known collection of documents [6, 7, 8, 45, 46, 47] that is part of the TREC-5 collection [30] and includes news articles from the Los Angeles Times.



	Financial	Foreign	National	Metro	Sports	Entertainment	All
Test	185	115	91	315	246	118	1070
Training	370	226	182	628	492	236	2134
All	555	341	273	943	738	354	3204

Figure 7.20. Property of the LA1 dataset

The categories correspond to the desk of the paper that each article appeared and contains 3,204 documents from the Financial (555), Foreign (341), National (273), Metro (943), Sports (738) and Entertainment (354) desks as shown Figure 7.20. Two thirds of each class is selected for the training set and the remaining is used for testing like the Hitech dataset. There are two dominant topics in the LA1 dataset; these two topics Metro and Sport cover more than half of the dataset as seen in Figure 7.20.

7.2.2.2. Analysis of the Existing Metrics

We continue our analysis of the skew dataset by testing the existing feature selection metrics on the LA1 dataset. Table 19 shows the micro- and macro-averaged F-measure results for the global and local policies of the existing feature selection metrics on the LA1 dataset. In addition, we can see the success rates of the existing metrics under feature number criterion in Figure 7.21.

Micro-F	10	30	50	100	200	500	1000	1500	2000	All
tf-idf (g)	0,465	0,648	0,722	0,767	0,793	0,816	0,817	0,825	0,833	0,841
CHI (g)	0,340	0,635	0,663	0,745	0,789	0,822	0,828	0,824	**0,838**	0,841
IG (g)	0,388	0,664	0,724	0,769	0,804	0,828	**0,829**	**0,833**	**0,838**	0,841
DF (g)	0,103	0,397	0,642	0,709	0,762	0,799	0,821	0,832	0,827	0,841
Acc2 (g)	**0,478**	**0,687**	**0,758**	**0,790**	**0,812**	**0,831**	**0,829**	0,830	0,829	0,841
MAX	0,478	0,687	0,758	0,790	0,812	0,831	0,829	0,833	0,838	
Macro-F	10	30	50	100	200	500	1000	1500	2000	All
tf-idf (g)	0,284	0,528	0,628	0,692	0,715	0,752	0,748	0,753	0,765	0,777
CHI (g)	0,318	0,523	0,549	0,651	0,722	0,762	**0,765**	0,757	**0,775**	0,777
IG (g)	0,301	0,510	0,603	0,658	0,745	**0,771**	0,764	0,762	0,772	0,777
DF (g)	0,117	0,227	0,515	0,588	0,688	0,724	0,756	0,766	0,760	0,777
Acc2 (g)	**0,387**	**0,584**	**0,677**	**0,732**	**0,754**	0,769	0,758	**0,768**	0,759	0,777
MAX	0,387	0,584	0,677	0,732	0,754	0,771	0,765	0,768	0,775	
Micro-F	10	30	50	100	200	500	1000	1500	2000	All
tf-idf (l)	0,631	0,731	0,761	0,785	0,789	0,807	0,814	0,812	0,815	0,841
CHI (l)	**0,671**	0,736	0,761	0,788	0,813	0,823	0,833	**0,840**	0,838	0,841
IG (l)	0,660	0,739	**0,765**	0,793	0,807	**0,830**	0,831	0,831	0,833	0,841
DF (l)	0,318	0,688	0,740	0,766	0,782	0,814	0,815	0,827	0,826	0,841
Acc2 (l)	0,659	**0,742**	0,764	**0,802**	**0,817**	0,829	**0,835**	**0,840**	**0,840**	0,841
MAX	0,671	0,742	0,765	0,802	0,817	0,830	0,835	0,840	0,840	
Macro-F	10	30	50	100	200	500	1000	1500	2000	All
tf-idf (l)	0,552	0,674	0,706	0,728	0,735	0,755	0,762	0,756	0,764	0,777
CHI (l)	**0,607**	0,686	**0,715**	0,741	0,766	0,772	0,778	**0,788**	**0,785**	0,777
IG (l)	0,578	**0,688**	0,714	0,743	0,756	**0,781**	**0,779**	0,775	0,777	0,777
DF (l)	0,376	0,582	0,665	0,702	0,715	0,755	0,758	0,767	0,768	0,777
Acc2 (l)	0,546	0,682	0,712	**0,759**	**0,773**	0,770	0,774	0,776	0,781	0,777
MAX	0,607	0,688	0,715	0,759	0,773	0,781	0,779	0,788	0,785	

Table 7.27. Micro- and macro-averaged F-measures for LA1 dataset

Over the LA1 dataset, the first observation is macro-averaged F-measure results are noticeably less than micro-averaged F-measure results like the Hitech dataset. In this dataset, the average of the difference between F-measures is about 5-7%. The gap between the two F-measure results is especially higher (about 9.2%) when the number of keywords is low (less than 100 keywords) but it gets closed with the high number of keywords. This difference is not an unexpected result since we knew that the LA1 dataset consists of six categories and it is dominated by the two topics that cover more than half of the dataset as seen in Figure 7.20.

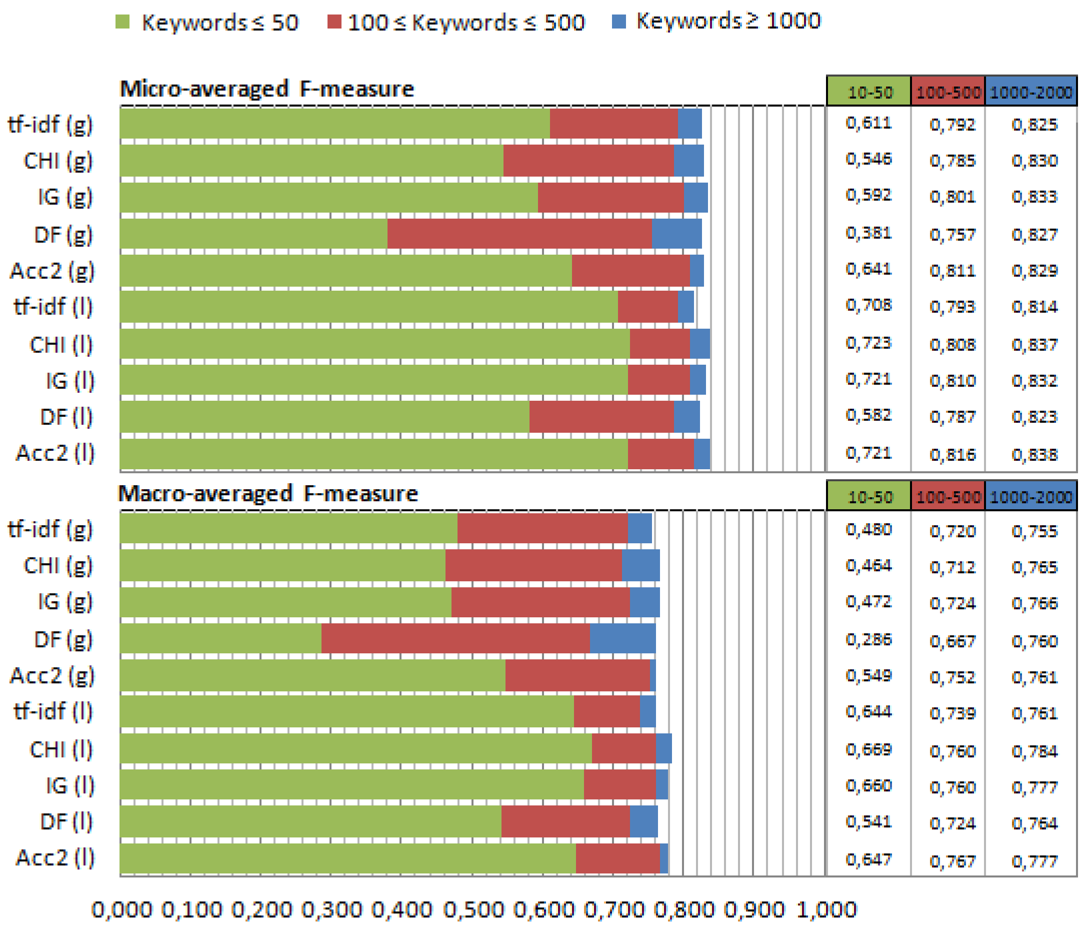


Figure 7.21. Comparison of the averages of micro- and macro-averaged F-measures for LA1 dataset

Although we observed that the local policy only achieves better F-measure performance than the global policy when the keyword number is low in the Classic3 and Hitech datasets, in LA1 dataset the local policy is remarkably better than the global policy with all number of keywords from 10 to 2000.

Thus we can say that the superiority of the local policy over the global policy become more apparent in the LA1 dataset. Especially it performs is about 9-12% better than the global policy when the keyword number is low.

The classifier achieves the best performance of 84.1% micro- and 77.7% macro-averaged F-measure when the test documents are classified without applying any feature selection method in the LA1 dataset. The closest performance to these results is achieved by the local policy with 1500 keywords, 84% micro- and 78.8% macro-averaged F-measure. As a result only macro-averaged F-measure is improved by the local policy but it should be noted that the improvement of the macro-averaged F-measure starts with 500 keywords. The local policy with 500 and more keywords achieves higher results than the all word approach in macro-averaged F-measure.

In LA1 dataset generally *CHI*, *IG* and *Acc2* are better than other metrics. The superiority of *Acc2* over the other metrics is emphasized more clearly in this dataset. *Acc2* outperforms other metrics both in global and local policies. On the other hand, we see that *IG* is better than Acc2 in global policy when the number of keywords is high from 1000 to 2000 keyword but this success rates are improved by Acc2 in local policy. Moreover *CHI* is better than other metrics when we compare to average of the macro-averaged F-measure results in local policy. Figure 7.21 shows the success rates of the existing metrics under feature number criterion.

7.2.2.3. Analysis of Score and Rank Combinations

First of all, we analyze the results of the score and rank combinations in the case of global policy. Tables 7.28 and 7.29 show the micro- and macro-averaged F-measure results of the experiments in global policy for the LA1 dataset.

In the previous datasets Classic3 and Hitech, we concluded that rank combination is significantly better than score combination in the case of global policy. This conclusion is also valid for the LA1 dataset. Although the results of the combinations are very close to each other, rank combination is still better than score combination as seen Tables 7.28 and 7.29.

In global policy among the 10 possible combinations of the three feature selection metrics, *tf-idf & CHI, tf-idf & Acc2* and *IG & Acc2* score and rank combinations are more successful than other combinations based on the highest micro- and macro-averaged F-measure values for each keyword number.

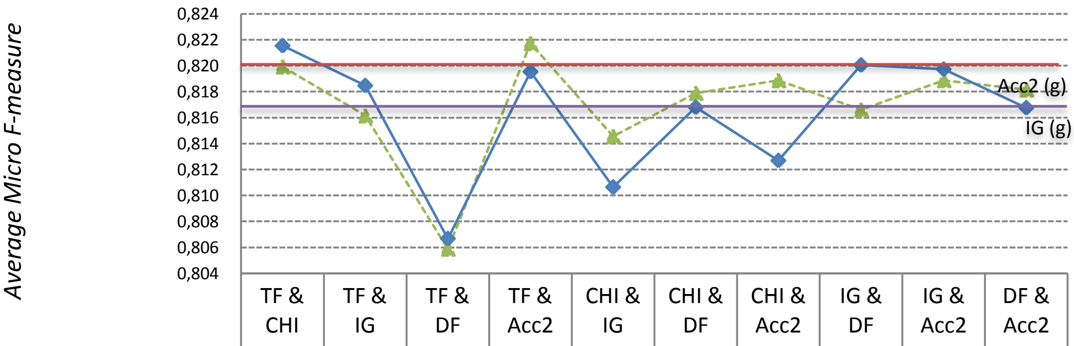
Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,465	0,648	0,722	0,767	0,793	0,816	0,817	0,825	0,833
CHI (g)		0,340	0,635	0,663	0,745	0,789	0,822	0,828	0,824	**0,838**
IG (g)		0,388	0,664	0,724	0,769	0,804	0,828	**0,829**	**0,833**	**0,838**
DF (g)		0,103	0,397	0,642	0,709	0,762	0,799	0,821	0,832	0,827
Acc2 (g)		**0,478**	**0,687**	**0,758**	**0,790**	**0,812**	**0,831**	**0,829**	0,830	0,829
	MAX	0,478	0,687	0,758	0,790	0,812	0,831	0,829	0,833	0,838
Micro-F	Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (g)		**0,541**	0,632	0,715	**0,791**	**0,812**	0,821	0,825	0,831	**0,840**
|S| tf-idf&IG (g)		0,461	0,653	0,732	0,784	0,811	0,817	0,824	0,825	0,836
|S| tf-idf&DF (g)		0,455	0,617	0,688	0,749	0,790	0,815	0,826	0,826	0,830
|S| tf-idf&Acc2 (g)		**0,549**	**0,689**	0,749	0,788	**0,817**	0,830	**0,833**	0,828	0,836
|S| CHI&IG (g)		0,388	0,628	0,683	0,756	0,809	0,825	0,828	**0,833**	0,836
|S| CHI&DF (g)		0,292	0,618	0,699	0,783	0,809	0,823	**0,830**	0,829	0,834
|S| CHI&Acc2 (g)		0,344	0,629	0,705	0,782	0,810	0,829	**0,832**	0,829	0,832
|S| IG&DF (g)		0,283	0,619	0,719	0,781	0,811	0,819	**0,831**	0,828	0,829
|S| IG&Acc2 (g)		0,476	0,686	0,755	0,788	0,806	0,825	**0,833**	0,832	0,829
|S| DF&Acc2 (g)		0,277	0,638	0,737	0,788	0,809	0,827	**0,829**	0,830	0,825
	MAX	**0,549**	**0,689**	0,755	**0,791**	**0,817**	0,830	**0,833**	**0,833**	**0,840**
	AVERAGE	0,407	0,641	0,718	0,779	0,808	0,823	0,829	0,829	0,833
Micro-F	Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (g)		**0,541**	0,652	0,702	0,784	**0,813**	0,829	0,827	**0,837**	**0,840**
|R| tf-idf&IG (g)		**0,481**	0,672	0,723	0,783	**0,814**	0,819	0,828	**0,834**	0,833
|R| tf-idf&DF (g)		0,475	0,665	0,683	0,745	0,793	0,810	**0,834**	0,827	0,833
|R| tf-idf&Acc2 (g)		**0,548**	**0,708**	0,752	0,784	**0,821**	0,822	**0,831**	0,824	0,835
|R| CHI&IG (g)		0,466	0,622	0,683	0,745	0,803	0,826	**0,829**	0,827	0,834
|R| CHI&DF (g)		**0,534**	0,643	0,688	0,775	0,800	**0,833**	0,824	0,831	**0,838**
|R| CHI&Acc2 (g)		**0,482**	0,660	0,713	0,751	0,809	0,825	**0,831**	0,827	0,833
|R| IG&DF (g)		0,457	0,660	0,725	0,789	0,808	0,826	**0,832**	**0,833**	0,831
|R| IG&Acc2 (g)		0,476	0,686	0,725	0,788	0,808	0,823	**0,829**	0,835	0,836
|R| DF&Acc2 (g)		0,447	0,667	0,716	0,787	0,804	0,824	**0,829**	0,827	0,829
	MAX	**0,548**	**0,708**	0,752	0,789	**0,821**	**0,833**	**0,834**	**0,837**	**0,840**
	AVERAGE	0,491	0,663	0,711	0,773	0,808	0,824	0,829	0,830	0,834

Table 7.28. In global policy, micro-averaged F-measures of the score and sank combinations for LA1 dataset

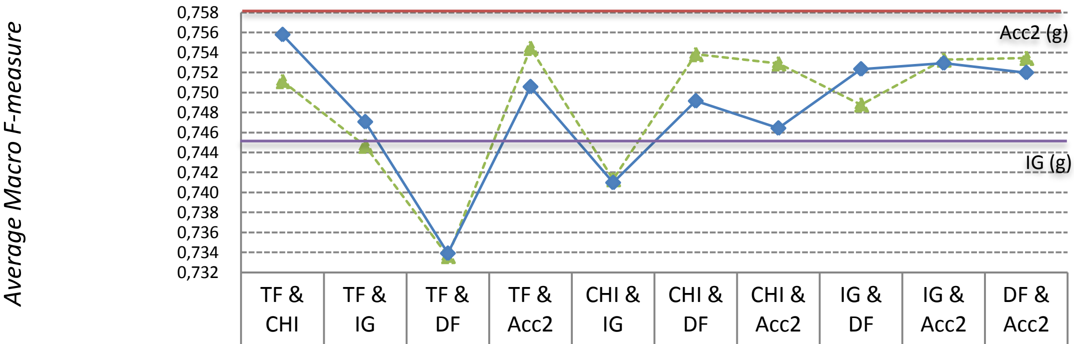
Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,284	0,528	0,628	0,692	0,715	0,752	0,748	0,753	0,765
CHI (g)		0,318	0,523	0,549	0,651	0,722	0,762	**0,765**	0,757	**0,775**
IG (g)		0,301	0,510	0,603	0,658	0,745	**0,771**	0,764	0,762	0,772
DF (g)		0,117	0,227	0,515	0,588	0,688	0,724	0,756	0,766	0,760
Acc2 (g)		**0,387**	**0,584**	**0,677**	**0,732**	**0,754**	0,769	0,758	**0,768**	0,759
	MAX	0,387	0,584	0,677	0,732	0,754	0,771	0,765	0,768	0,775
Macro-F	Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (g)		0,383	0,470	0,603	0,719	0,747	0,753	0,755	0,761	0,772
|S| tf-idf&IG (g)		0,250	0,481	0,644	0,695	0,749	0,748	0,751	0,757	0,767
|S| tf-idf&DF (g)		0,233	0,456	0,570	0,660	0,713	0,748	0,760	0,756	0,764
|S| tf-idf&Acc2 (g)		**0,397**	0,570	0,674	0,723	0,752	0,761	**0,766**	0,757	0,768
|S| CHI&IG (g)		0,301	0,470	0,562	0,643	0,746	0,769	0,762	0,763	0,764
|S| CHI&DF (g)		0,140	0,465	0,553	0,714	0,746	0,760	**0,765**	**0,770**	0,769
|S| CHI&Acc2 (g)		0,217	0,471	0,605	0,708	0,751	**0,771**	0,762	0,763	0,762
|S| IG&DF (g)		0,149	0,441	0,603	0,704	0,749	0,753	**0,766**	0,760	0,760
|S| IG&Acc2 (g)		0,298	0,549	**0,678**	0,716	0,748	0,762	**0,767**	0,765	0,761
|S| DF&Acc2 (g)		0,134	0,508	0,644	0,724	0,745	0,766	**0,768**	0,761	0,757
	MAX	**0,397**	0,570	**0,678**	0,724	0,752	**0,771**	**0,768**	**0,770**	0,772
	AVERAGE	0,250	0,488	**0,614**	**0,701**	**0,745**	**0,759**	0,762	**0,761**	**0,765**
Macro-F	Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (g)		0,383	0,527	0,594	0,717	0,751	0,769	0,758	0,767	0,773
|R| tf-idf&IG (g)		0,393	0,552	0,604	0,707	0,747	0,746	0,754	0,764	0,764
|R| tf-idf&DF (g)		0,261	0,532	0,553	0,652	0,728	0,736	0,765	0,759	0,763
|R| tf-idf&Acc2 (g)		**0,395**	**0,591**	0,645	0,711	**0,754**	0,753	0,763	0,755	0,768
|R| CHI&IG (g)		0,256	0,463	0,561	0,650	0,740	0,769	**0,767**	0,757	0,763
|R| CHI&DF (g)		0,356	0,520	0,573	0,704	0,736	0,769	0,758	0,761	0,767
|R| CHI&Acc2 (g)		0,297	0,533	0,606	0,671	0,749	0,767	**0,772**	0,756	0,763
|R| IG&DF (g)		0,267	0,530	0,606	0,716	0,742	0,763	**0,768**	0,765	0,761
|R| IG&Acc2 (g)		0,299	0,549	0,615	0,719	0,743	0,761	0,760	0,767	0,767
|R| DF&Acc2 (g)		0,232	0,536	0,611	0,721	0,749	0,759	**0,763**	0,760	0,760
	MAX	**0,395**	**0,591**	0,645	0,721	**0,754**	0,769	**0,772**	0,767	0,773
	AVERAGE	**0,314**	**0,533**	0,597	0,697	0,744	**0,759**	**0,763**	**0,761**	**0,765**

Table 7.29. In global policy, macro-averaged F-measures of the score and rank combinations for LA1 dataset

According to test results at Table 7.28, among the existing metrics *Acc2* is the best metric with a few number of keywords, less than 1500 keywords, and *IG* is the best metric with a high number of keywords, higher than 500 keywords, in global policy. After the combination experiments, we can say that the success of the *Acc2* is improved by the score and rank combinations of tf-*idf & Acc2* and the success of the *IG* metric is improved by the rank combination of *tf-idf & CHI.*



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,820	0,816	0,806	0,822	0,815	0,818	0,819	0,817	0,819	0,818
Rank Com. (g)	0,822	0,818	0,807	0,820	0,811	0,817	0,813	0,820	0,820	0,817



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,751	0,745	0,734	0,754	0,741	0,754	0,753	0,749	0,753	0,753
Rank Com. (g)	0,756	0,747	0,734	0,751	0,741	0,749	0,746	0,752	0,753	0,752

	tf-idf (g)	CHI (g)	IG (g)	DF (g)	Acc2 (g)
Micro	0,809	0,808	0,817	0,792	0,820
Macro	0,737	0,738	0,745	0,714	0,757

Figure 7.22. In global policy, comparison of score and rank combinations on the LA1 dataset

In Figure 7.22, we knew that *Acc2* is the first and *CHI* is the second best metrics that has the highest averages of F-measure. More than half of the combinations are better than the second best average in this figure. As a result we can say that combining individual metrics is a good approach for the LA1 dataset. As we said before rank combination is slightly better than score combination in global policy but score combination is as successful as rank combination on comparison of the average of the F-measures when we look at the Figure 7.22.

Success of the combinations is more apparent when the keyword number is low in global policy. As known each method is tested with keyword numbers from 10 to 2000 in all experiments. The most significant improvement is achieved by the combinations with the lowest keyword 10 compare to individual measures. *Acc2* has the highest 47.8% micro-averaged and 38.7% macro-averaged F-measure with 10 keywords. These results are outperformed by the score combination of with tf-*idf & Acc2* with 54.9% micro-averaged and 39.7% macro-averaged F-measure.

Among the existing metrics the best performance of 83.8% micro-averaged and 77.5% macro-averaged F-measure is achieved by CHI with 2000 keywords. This performance is improved by the rank combination of *tf-idf & CHI* with value of 84% micro-averaged and 77.3% macro-averaged F-measure in global policy.

Other observation over the LA1 dataset in global policy, the improvement of the micro-averaged F-measure is higher than the macro-averaged F-measure by the combinations.

In global policy there are only two combinations *tf-idf & CHI* and *tf-idf & Acc2* improve the performance of the best individual metric *Acc2* among the 10 possible combinations. On the other hand the success of the combination is more apparent in the case of local policy. Tables 7.30 and 7.31 show, respectively, the micro- and macro-averaged F-measure results of the score and rank combinations in local policy for the LA1 dataset. Score and rank combinations of *tf-idf & IG, tf-idf & Acc2, CHI & IG, CHI & Acc2* and *IG & Acc2* are better than other combinations and among them *CHI & IG, CHI & Acc2* and *IG & Acc2* combinations achieve the best results in local policy in the LA1 dataset.

In local policy the superiority of *Acc2* over the other metrics is emphasized more clearly in this dataset on comparisons of the micro-averaged F-measure results but *CHI* is slightly better than other metrics on comparisons of the macro-averaged F-measure results.

After testing combinations, score combination of *CHI & Acc2* is apparently the best combination and outperforms the existing metrics Acc2 and CHI when compared to the micro- and macro averaged F-measure results.

Micro-F	10	30	50	100	200	500	1000	1500	2000
tf-idf (I)	0,631	0,731	0,761	0,785	0,789	0,807	0,814	0,812	0,815
CHI (I)	**0,671**	0,736	0,761	0,788	0,813	0,823	0,833	**0,840**	0,838
IG (I)	0,660	0,739	**0,765**	0,793	0,807	**0,830**	0,831	0,831	0,833
DF (I)	0,318	0,688	0,740	0,766	0,782	0,814	0,815	0,827	0,826
Acc2 (I)	0,659	**0,742**	0,764	**0,802**	**0,817**	0,829	**0,835**	**0,840**	**0,840**
MAX	0,671	0,742	0,765	0,802	0,817	0,830	0,835	0,840	0,840
Micro-F Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (I)	0,647	0,740	0,758	0,783	0,803	0,828	0,832	0,828	0,829
|S| tf-idf&IG (I)	0,650	0,738	0,771	0,783	0,809	0,837	0,837	0,833	0,832
|S| tf-idf&DF (I)	0,608	0,716	0,754	0,777	0,788	0,813	0,825	0,827	0,826
|S| tf-idf&Acc2 (I)	0,658	0,755	0,767	0,794	0,815	0,830	0,837	0,835	0,838
|S| CHI&IG (I)	0,658	0,740	0,764	0,785	0,811	0,821	0,841	0,838	0,841
|S| CHI&DF (I)	0,610	0,720	0,754	0,791	0,811	0,829	0,836	0,830	0,831
|S| CHI&Acc2 (I)	0,665	0,738	0,768	0,788	0,819	0,838	0,837	0,842	0,842
|S| IG&DF (I)	0,613	0,729	0,762	0,800	0,812	0,826	0,839	0,838	0,831
|S| IG&Acc2 (I)	0,656	0,743	0,767	0,793	0,815	0,835	0,839	0,839	0,842
|S| DF&Acc2 (I)	0,600	0,732	0,759	0,803	0,816	0,827	0,834	0,835	0,829
MAX	0,665	**0,755**	**0,771**	**0,803**	**0,819**	**0,838**	**0,841**	**0,842**	**0,842**
AVERAGE	0,637	0,735	0,762	0,790	0,810	0,828	0,836	0,835	0,834
Micro-F Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (I)	0,639	0,735	0,753	0,768	0,800	0,809	0,823	0,824	0,827
|R| tf-idf&IG (I)	0,650	0,736	0,750	0,766	0,802	0,828	0,830	0,831	0,833
|R| tf-idf&DF (I)	0,655	0,736	0,758	0,773	0,783	0,802	0,820	0,823	0,830
|R| tf-idf&Acc2 (I)	0,651	0,732	0,749	0,765	0,797	0,829	0,829	0,833	0,832
|R| CHI&IG (I)	0,659	0,745	0,762	0,787	0,812	0,823	0,839	0,836	0,842
|R| CHI&DF (I)	0,656	0,735	0,761	0,778	0,800	0,832	0,831	0,833	0,839
|R| CHI&Acc2 (I)	0,673	0,736	0,769	0,789	0,805	0,834	0,830	0,842	0,841
|R| IG&DF (I)	0,654	0,728	0,756	0,784	0,801	0,825	0,834	0,841	0,837
|R| IG&Acc2 (I)	0,664	0,738	0,768	0,790	0,812	0,830	0,836	0,837	0,842
|R| DF&Acc2 (I)	0,642	0,726	0,759	0,786	0,804	0,818	0,837	0,826	0,830
MAX	**0,673**	**0,745**	**0,769**	0,790	**0,812**	**0,834**	**0,839**	**0,842**	**0,842**
AVERAGE	0,654	0,735	0,759	0,779	0,802	0,823	0,831	0,833	0,835

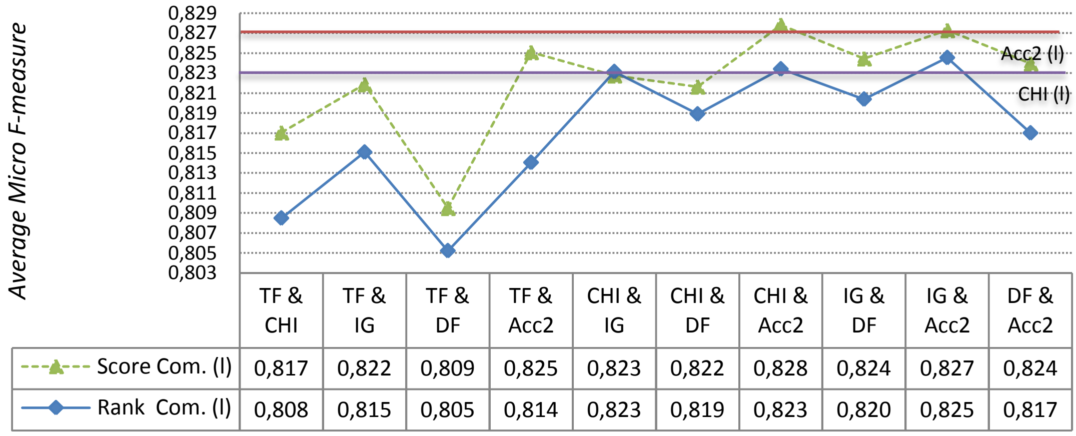
Table 7.30. In global policy, micro-averaged F-measures of the score and rank combinations for LA1 dataset

Macro-F	10	30	50	100	200	500	1000	1500	2000
tf-idf (I)	0,552	0,674	0,706	0,728	0,735	0,755	0,762	0,756	0,764
CHI (I)	**0,607**	0,686	**0,715**	0,741	0,766	0,772	0,778	**0,788**	**0,785**
IG (I)	0,578	**0,688**	0,714	0,743	0,756	**0,781**	**0,779**	0,775	0,777
DF (I)	0,376	0,582	0,665	0,702	0,715	0,755	0,758	0,767	0,768
Acc2 (I)	0,546	0,682	0,712	**0,759**	**0,773**	0,770	0,774	0,776	0,781
MAX	0,607	0,688	0,715	0,759	0,773	0,781	0,779	0,788	0,785
Macro-F Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (I)	0,574	0,689	0,713	0,735	0,755	0,775	0,779	0,777	0,777
|S| tf-idf&IG (I)	0,583	0,684	**0,724**	0,724	0,763	**0,784**	**0,788**	0,785	0,776
|S| tf-idf&DF (I)	0,526	0,649	0,688	0,717	0,730	0,762	0,771	0,772	0,771
|S| tf-idf&Acc2 (I)	0,593	**0,698**	**0,722**	0,741	0,759	0,774	**0,783**	0,781	0,782
|S| CHI&IG (I)	0,595	0,686	**0,717**	0,739	0,764	0,773	**0,787**	0,785	**0,790**
|S| CHI&DF (I)	0,526	0,660	0,701	0,739	0,757	0,773	**0,782**	0,773	0,776
|S| CHI&Acc2 (I)	0,588	**0,688**	**0,716**	0,738	**0,776**	**0,793**	**0,786**	0,782	**0,786**
|S| IG&DF (I)	0,503	0,670	0,711	0,747	0,761	0,765	**0,779**	0,787	0,774
|S| IG&Acc2 (I)	0,571	**0,692**	**0,716**	0,741	0,771	**0,782**	**0,790**	0,781	0,778
|S| DF&Acc2 (I)	0,476	0,669	0,709	0,752	0,769	0,766	0,777	0,774	0,768
MAX	0,595	**0,698**	**0,724**	0,752	**0,776**	**0,793**	**0,790**	0,787	**0,790**
AVERAGE	0,554	0,678	0,712	0,737	0,760	0,775	0,782	0,780	0,778
Macro-F Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (I)	0,570	0,687	0,713	0,721	0,748	0,757	0,769	0,769	0,777
|R| tf-idf&IG (I)	0,583	0,685	0,698	0,714	0,751	0,777	0,777	0,774	0,775
|R| tf-idf&DF (I)	0,559	0,675	0,695	0,715	0,721	0,737	0,765	0,766	0,775
|R| tf-idf&Acc2 (I)	0,581	0,679	0,698	0,710	0,730	0,771	0,775	0,777	0,777
|R| CHI&IG (I)	0,589	**0,692**	0,712	0,742	0,767	0,769	**0,784**	0,782	**0,787**
|R| CHI&DF (I)	0,588	0,687	**0,716**	0,739	0,748	**0,782**	**0,780**	0,781	0,782
|R| CHI&Acc2 (I)	0,602	0,684	**0,725**	0,740	0,758	**0,783**	**0,779**	**0,790**	**0,785**
|R| IG&DF (I)	0,563	0,676	0,703	0,749	0,755	0,774	**0,784**	**0,791**	0,784
|R| IG&Acc2 (I)	0,575	0,686	**0,716**	0,741	0,766	0,780	**0,789**	0,783	0,784
|R| DF&Acc2 (I)	0,545	0,663	0,702	0,727	0,749	0,760	**0,779**	0,767	0,768
MAX	0,602	**0,692**	**0,725**	0,749	0,767	**0,783**	**0,789**	**0,791**	**0,787**
AVERAGE	0,575	0,681	0,708	0,730	0,749	0,769	0,778	0,778	0,779

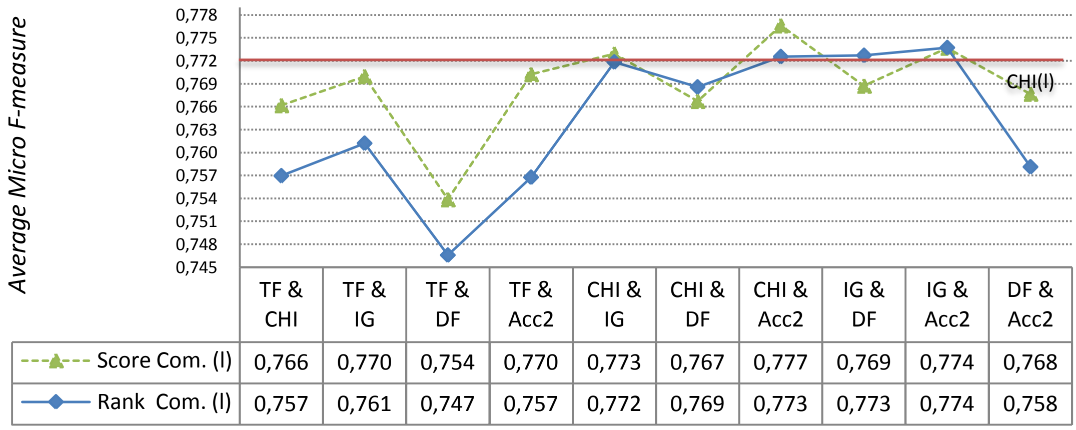
Table 7.31. In global policy, macro-averaged F-measures of the score and rank combinations for LA1 dataset

The classifier achieves the best performance of 84.1% micro- and 77.7% macro-averaged F-measure when the test documents are classified without applying any feature selection method in the LA1 dataset. Only macro-averaged F-measure is improved by the local policy that the closest performance to these results is achieved with 1500 keywords, 84% micro-averaged F-measure by both *CHI* and *Acc2* and 78.8% macro-averaged F-measure by *CHI*. Score combination of *CHI & Acc2* outperforms the highest performance with values of 84.2% micro-averaged F-measure with 1500 keywords and 79.3% macro-averaged F-measure with only 500 keywords.

As seen in Figure 7.23, score and rank combinations of *CHI & IG, CHI & Acc2* and *IG & Acc2* are more successful than other combinations in local policy in the LA1 dataset. In addition, one of the important results of this experiment score combination is significantly better than the rank combination in the case of local policy.



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (l)	0,817	0,822	0,809	0,825	0,823	0,822	0,828	0,824	0,827	0,824
Rank Com. (l)	0,808	0,815	0,805	0,814	0,823	0,819	0,823	0,820	0,825	0,817



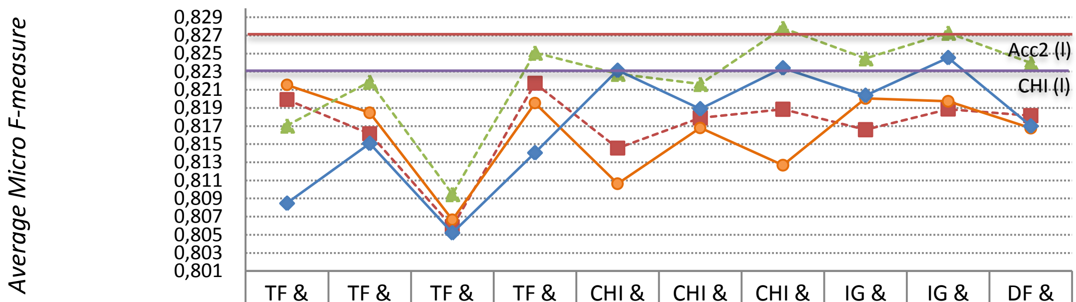
	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (l)	0,766	0,770	0,754	0,770	0,773	0,767	0,777	0,769	0,774	0,768
Rank Com. (l)	0,757	0,761	0,747	0,757	0,772	0,769	0,773	0,773	0,774	0,758

	tf-idf (l)	CHI (l)	IG (l)	DF (l)	Acc2 (l)
Micro	0,804	0,823	0,821	0,805	0,827
Macro	0,750	0,772	0,769	0,744	0,772

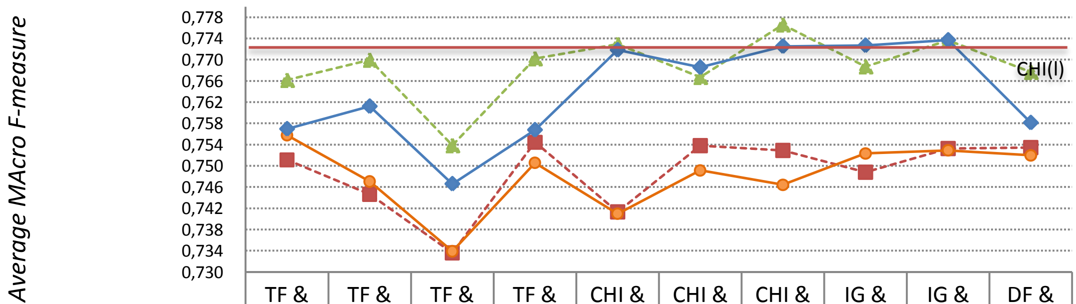
Figure 7.23. In local policy, comparison of score and rank combinations on the LA1 dataset

In LA1 dataset the performance of score combination is better than rank combination in both global and local policy when we compare the average of F-measure results as seen Figure 7.24.

In LA1 dataset the local policy is remarkably better than the global policy as analyzed in Section 7.2.2.2. This result also affects the performance of the combination. The combinations achieve better results in the case of local policy. Finally, score combination in local policy is the best method for LA1 dataset.



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,820	0,816	0,806	0,822	0,815	0,818	0,819	0,817	0,819	0,818
Rank Com. (g)	0,822	0,818	0,807	0,820	0,811	0,817	0,813	0,820	0,820	0,817
Score Com. (l)	0,817	0,822	0,809	0,825	0,823	0,822	0,828	0,824	0,827	0,824
Rank Com. (l)	0,808	0,815	0,805	0,814	0,823	0,819	0,823	0,820	0,825	0,817



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,751	0,745	0,734	0,754	0,741	0,754	0,753	0,749	0,753	0,753
Rank Com. (g)	0,756	0,747	0,734	0,751	0,741	0,749	0,746	0,752	0,753	0,752
Score Com. (l)	0,766	0,770	0,754	0,770	0,773	0,767	0,777	0,769	0,774	0,768
Rank Com. (l)	0,757	0,761	0,747	0,757	0,772	0,769	0,773	0,773	0,774	0,758

Figure 7.24. Comparison of score and rank combinations in global and local policy on the LA1 dataset

7.2.2.4. Analysis of the Proposed Combinations

Before starting the analysis of the proposed combination methods, we summarize what we learned from the previous experiments on the LA1 dataset. Firstly we analyzed the existing metrics and observed that *Acc2* and *IG were* the best metrics in global policy. On the other hand, in local policy *Acc2* and *CHI* were significantly better than other metrics. Then we compared score and rank combinations with existing metric and we found out that rank combinations of *tf-idf & CHI* and *tf-idf & Acc2* improved the performances of the best individual metrics among the 10 possible combinations in global policy. In addition, score combination of *CHI & Acc2* apparently outperformed the best existing metrics *Acc2* and *CHI* in local policy.

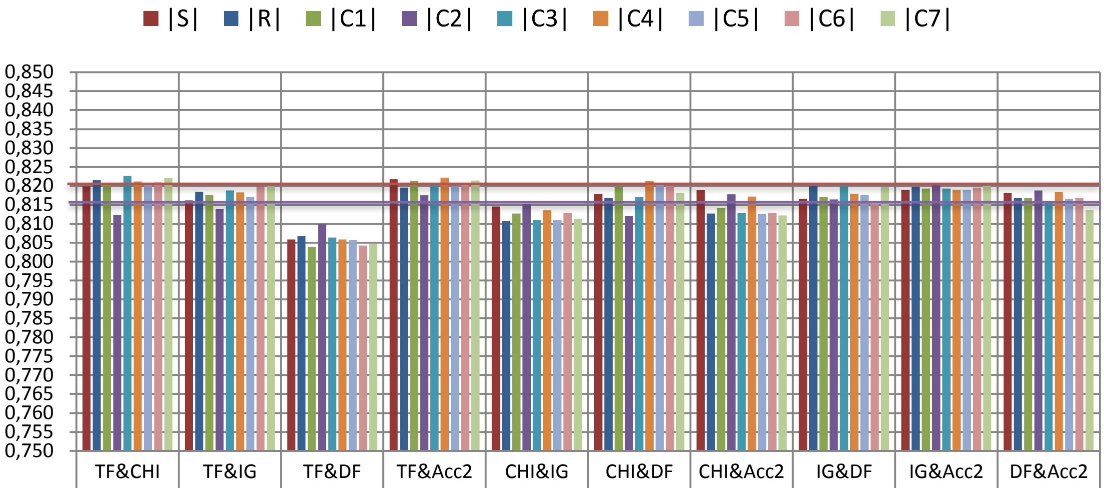
When we evaluated the performance of the existing metrics on the LA1 dataset in Section 7.2.2.2, we also learned that the superiority of the local policy over the global policy became more apparent in this dataset. The performance of the local policy was always better than the global policy. Thus, we especially focus on the performance of the proposed methods in local policy but first of all we examine the results of the experiments in global policy.

Tables 7.34 and 7.35 show the micro- and macro-averaged F-measures, respectively, for all seven proposed methods in global policy for the LA1 dataset.

Before starting the analysis in the global policy, it should be noted that success of the combinations is lower than we expected. The principal reason is that the performance of the score and rank combinations also could not significantly improve the performance of the individual metrics.

In Figures 7.25 and 7.26, we can more explicitly see that the performance of each combination is very close to each other when we analyze the average of the F-measures. Moreover *tf-idf & CHI, tf-idf & Acc2, CHI & DF* and *IG & Acc2* are more successful than the other combinations when we look at the highest F-measure values for each keyword number. We can also see that combination of *tf-idf & DF* and *CHI & IG* are not enough successful and they are the worst combinations in global policy.

Although in global policy the performance of the score and rank combinations on the LA1 dataset was very close to each other, rank combination was still better than score combination. In addition among these proposed methods C3, C5, C6 and C7 also improve the success of the rank combination when we look at the micro- and macro-averaged F-measure values. The most successful combinations are obtained from these proposed methods. C5 and C7 improve the performance of the rank combination with high number of keywords C3 and C6 are very successful when the number of keywords is lower than 1500.

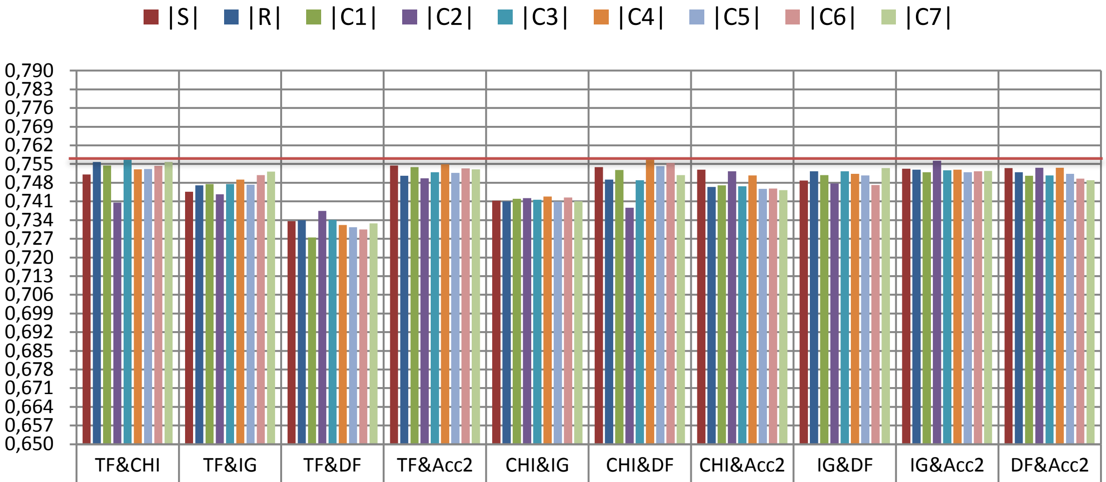


	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,820	0,816	0,806	0,822	0,815	0,818	0,819	0,817	0,819	0,818
|R|	0,822	0,818	0,807	0,820	0,811	0,817	0,813	0,820	0,820	0,817
|C1|	0,821	0,818	0,804	0,821	0,813	0,820	0,814	0,817	0,819	0,817
|C2|	0,812	0,814	0,810	0,817	0,815	0,812	0,818	0,816	0,821	0,819
|C3|	0,823	0,819	0,806	0,821	0,811	0,817	0,813	0,820	0,819	0,816
|C4|	0,821	0,818	0,806	0,822	0,814	0,821	0,817	0,818	0,819	0,818
|C5|	0,820	0,817	0,806	0,820	0,811	0,820	0,813	0,818	0,819	0,817
|C6|	0,821	0,820	0,804	0,821	0,813	0,820	0,813	0,815	0,820	0,817
|C7|	0,822	0,820	0,805	0,821	0,811	0,818	0,812	0,820	0,820	0,814

Figure 7.25. In global policy, averages of the micro-averaged F-measures of all combinations for LA1 dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	**0,549**	0,689	**0,755**	**0,791**	0,817	0,830	0,833	0,833	**0,840**
Combination | R |	0,548	**0,708**	0,752	0,789	**0,821**	**0,833**	**0,834**	**0,837**	**0,840**
Combination |C1|	0,549	0,698	0,743	0,792	0,823	0,832	0,833	0,837	0,836
Combination |C2|	0,549	0,692	0,761	0,791	0,817	0,826	0,832	0,833	0,837
Combination |C3|	0,549	0,708	0,751	0,789	0,825	0,832	0,834	0,836	0,839
Combination |C4|	0,549	0,709	0,745	0,795	0,822	0,828	0,833	0,834	0,837
Combination |C5|	0,549	0,710	0,750	0,789	0,822	0,831	0,832	0,835	0,843
Combination |C6|	0,549	0,709	0,745	0,792	0,819	0,832	0,835	0,837	0,838
Combination |C7|	0,549	0,710	0,750	0,793	0,819	0,830	0,834	0,835	0,841

Table 7.32. In global policy, maximum micro-averaged F-measures of all combinations for LA1 dataset



	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,751	0,745	0,734	0,754	0,741	0,754	0,753	0,749	0,753	0,753
|R|	0,756	0,747	0,734	0,751	0,741	0,749	0,746	0,752	0,753	0,752
|C1|	0,754	0,748	0,727	0,754	0,742	0,753	0,747	0,751	0,752	0,751
|C2|	0,741	0,744	0,737	0,750	0,742	0,739	0,752	0,748	0,756	0,754
|C3|	0,757	0,748	0,734	0,752	0,742	0,749	0,747	0,752	0,753	0,751
|C4|	0,753	0,749	0,732	0,755	0,743	0,757	0,751	0,751	0,753	0,754
|C5|	0,753	0,747	0,731	0,752	0,741	0,754	0,746	0,751	0,752	0,751
|C6|	0,754	0,751	0,730	0,753	0,742	0,755	0,746	0,747	0,752	0,750
|C7|	0,756	0,752	0,733	0,753	0,741	0,751	0,745	0,754	0,752	0,749

Figure 7.26. In global policy, averages of the macro-averaged F-measures of all combinations for LA1 dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	**0,397**	0,570	**0,678**	**0,724**	0,752	**0,771**	0,768	**0,770**	0,772
Combination | R |	0,395	**0,591**	0,645	0,721	**0,754**	0,769	**0,772**	0,767	**0,773**
Combination |C1|	0,397	0,578	0,660	0,719	0,760	0,772	0,769	0,771	0,767
Combination |C2|	0,402	0,573	0,684	0,725	0,752	0,766	0,771	0,766	0,771
Combination |C3|	0,397	0,591	0,644	0,721	0,761	0,778	0,773	0,767	0,772
Combination |C4|	0,397	0,594	0,656	0,729	0,762	0,767	0,768	0,767	0,767
Combination |C5|	0,397	0,593	0,639	0,725	0,756	0,769	0,768	0,764	0,779
Combination |C6|	0,397	0,593	0,656	0,720	0,753	0,770	0,771	0,769	0,775
Combination |C7|	0,397	0,593	0,644	0,722	0,754	0,771	0,773	0,767	0,774

Table 7.33. In global policy, maximum macro-averaged F-measures of all combinations for LA1 dataset

One of the observation from these tests, combination of *tf-idf & CHI* is successful if the number of keywords is higher than 100 and combination of *tf-idf & Acc2* is successful if the number of keywords is less than 1500.

Another observation is the performance of C2 and C4 is very unsuccessful when the keyword numbers are higher than 100. They cannot improve the performance of the individual metrics. Even though they are the worst combination methods for the global policy, C4 has the highest F-measure values when the number of keywords is low.

As we knew that among the individual metrics the best performance of 83.8% micro-averaged and 77.5% macro-averaged F-measure was achieved by *CHI* with 2000 keywords. This performance was improved by the rank combination of *tf-idf & CHI* with value of 84% micro-averaged and 77.3% macro-averaged F-measure in global policy. Among the proposed methods the most successful performance is achieved by C5 of *tf-idf & CHI* with value of 84.3% micro- and 77.9% macro-averaged F-measure value with 2000 keywords. The second best performance is achieved by C7 of *tf-idf & CHI* with value of 84.1% micro- and 77.4% macro-averaged F-measure value with 2000 keywords

Over the LA1 dataset in global policy, we can say that improvement of the macro-average F-measure is more explicit than the improvement of the micro-averaged F-measures. From this observation we can conclude that the proposed methods more apparently improve the performance of the classifier on rare categories in this dataset. For instance C3 of *tf-idf & CHI* achieves 77.8% macro-averaged F-measure value with only 500 keywords.

In the following part, we discuss the performance of the proposed methods in the case of local policy. As we mentioned above the local policy performed significantly better than the global policy in all experiments for the LA1 dataset. For this reason the success of the proposed methods in local policy is more important than the success in the global policy. Firstly we determine which individual metric couples will work best together for the classification. Among the 10 possible combinations of two feature selection metrics, the best combinations are *IG & Acc2, CHI & Acc2, IG & DF* and *CHI & IG* and the worst combinations are *tf-idf & DF* and *tf-idf & CHI* as seen Figures 7.27 and 7.28. As we can see, the same principle is valid for the LA1 dataset. The success of the combination depends on the performance of the individual metrics. Tables 7.38 and 7.39 show the micro- and macro-averaged F-measures, respectively, for all seven proposed methods in local policy for the LA1 dataset.

Table 7.34. Micro F-measure results of the proposed combinations in global policy for LA1 dataset

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,465	0,648	0,722	0,767	0,793	0,816	0,817	0,825	0,833
CHI (g)		0,340	0,635	0,663	0,745	0,789	0,822	0,828	0,824	**0,838**
IG (g)		0,388	0,664	0,724	0,769	0,804	0,828	**0,829**	**0,833**	**0,838**
DF (g)		0,103	0,397	0,642	0,709	0,762	0,799	0,821	0,832	0,827
Acc2 (g)		**0,478**	**0,687**	**0,758**	**0,790**	**0,812**	**0,831**	**0,829**	0,830	0,829
	MAX	0,478	0,687	0,758	0,790	0,812	0,831	0,829	0,833	0,838
Micro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	**0,549**	**0,689**	0,755	**0,791**	**0,817**	0,830	**0,833**	**0,833**	**0,840**
	AVERAGE	0,407	0,641	0,718	0,779	0,808	0,823	0,829	0,829	0,833
Rank Combination	MAX	**0,548**	**0,708**	0,752	0,789	**0,821**	**0,833**	**0,834**	**0,837**	**0,840**
	AVERAGE	0,491	0,663	0,711	0,773	0,808	0,824	0,829	0,830	0,834
Micro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (g)		**0,509**	0,626	0,691	0,783	**0,812**	**0,832**	0,833	**0,833**	0,832
|C1| tf-idf&IG (g)		0,445	0,668	0,727	0,783	**0,816**	0,821	0,823	0,826	0,836
|C1| tf-idf&DF (g)		0,475	0,648	0,675	0,748	0,776	0,810	**0,830**	0,829	0,830
|C1| tf-idf&Acc2 (g)		0,549	0,698	0,743	0,792	0,823	0,827	0,828	0,824	0,835
|C1| CHI&IG (g)		0,388	0,628	0,683	0,761	0,805	0,823	0,824	0,827	0,835
|C1| CHI&DF (g)		**0,528**	0,628	0,734	0,781	**0,816**	0,829	**0,832**	0,826	0,835
|C1| CHI&Acc2 (g)		0,344	0,628	0,697	0,755	0,808	0,829	**0,830**	0,831	0,831
|C1| IG&DF (g)		**0,492**	0,626	0,713	0,787	0,810	0,823	**0,830**	0,824	0,828
|C1| IG&Acc2 (g)		0,476	0,686	0,727	0,787	0,807	0,825	**0,830**	0,837	0,830
|C1| DF&Acc2 (g)		0,456	0,671	0,718	0,782	0,808	0,825	0,828	0,828	0,829
	MAX	**0,549**	**0,698**	0,743	**0,792**	**0,823**	**0,832**	**0,833**	**0,837**	0,836
	AVERAGE	0,466	0,651	0,711	0,776	0,808	0,824	**0,829**	0,828	0,832
Micro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (g)		0,470	0,659	0,710	0,774	0,799	0,821	0,822	0,823	0,835
|C2| tf-idf&IG (g)		0,395	0,637	0,728	0,775	0,804	0,816	0,825	0,827	0,836
|C2| tf-idf&DF (g)		0,418	0,608	0,702	0,776	0,790	0,815	0,819	0,827	0,834
|C2| tf-idf&Acc2 (g)		0,549	0,658	0,745	0,785	0,809	0,820	0,827	0,828	0,836
|C2| CHI&IG (g)		0,388	0,631	0,715	0,764	0,809	0,823	0,828	0,832	0,837
|C2| CHI&DF (g)		0,283	0,608	0,690	0,762	0,805	0,817	**0,829**	0,828	0,831
|C2| CHI&Acc2 (g)		**0,482**	0,634	0,742	0,784	0,810	0,824	0,828	0,829	0,832
|C2| IG&DF (g)		0,291	0,630	0,708	0,781	0,806	0,821	**0,830**	**0,833**	0,829
|C2| IG&Acc2 (g)		0,476	0,692	0,761	0,791	**0,812**	0,826	0,832	**0,833**	0,829
|C2| DF&Acc2 (g)		0,277	0,641	0,751	0,784	0,817	0,824	**0,829**	0,828	0,831
	MAX	**0,549**	**0,692**	**0,761**	**0,791**	**0,817**	0,826	**0,832**	**0,833**	0,837
	AVERAGE	0,403	0,640	0,725	0,778	0,806	0,821	0,827	0,829	0,833
Micro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (g)		**0,541**	0,633	0,706	0,788	**0,815**	0,832	0,827	0,836	0,839
|C3| tf-idf&IG (g)		0,473	0,672	0,721	0,785	**0,815**	0,818	**0,829**	**0,835**	0,832
|C3| tf-idf&DF (g)		0,381	0,666	0,685	0,744	0,788	0,813	0,834	0,827	0,833
|C3| tf-idf&Acc2 (g)		0,549	0,708	0,751	0,786	0,825	0,825	**0,832**	0,823	0,834
|C3| CHI&IG (g)		0,466	0,622	0,683	0,746	0,804	0,826	**0,829**	0,827	0,834
|C3| CHI&DF (g)		**0,534**	0,643	0,696	0,774	0,800	0,832	0,826	0,832	0,837
|C3| CHI&Acc2 (g)		**0,482**	0,636	0,706	0,751	0,807	0,826	**0,832**	0,827	0,834
|C3| IG&DF (g)		0,457	0,674	0,718	0,789	0,808	0,827	**0,832**	0,832	0,830
|C3| IG&Acc2 (g)		0,476	0,686	0,725	0,787	0,805	0,824	**0,829**	**0,835**	0,836
|C3| DF&Acc2 (g)		0,448	0,667	0,711	0,780	0,806	0,822	0,828	0,829	0,829
	MAX	**0,549**	**0,708**	0,751	0,789	**0,825**	**0,832**	**0,834**	**0,836**	**0,839**
	AVERAGE	**0,481**	0,661	0,710	0,773	0,807	0,824	**0,830**	0,830	0,834

Micro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (g)		0,541	0,626	0,697	0,790	0,814	0,827	0,830	0,834	0,832
|C4| tf-idf&IG (g)		0,445	0,669	0,727	0,787	0,817	0,822	0,821	0,829	0,834
|C4| tf-idf&DF (g)		0,487	0,647	0,675	0,749	0,789	0,811	0,826	0,827	0,832
|C4| tf-idf&Acc2 (g)		0,549	0,709	0,745	0,795	0,822	0,828	0,830	0,823	0,834
|C4| CHI&IG (g)		0,466	0,625	0,683	0,761	0,806	0,825	0,826	0,826	0,837
|C4| CHI&DF (g)		0,538	0,634	0,730	0,786	0,818	0,828	0,833	0,831	0,831
|C4| CHI&Acc2 (g)		0,344	0,628	0,697	0,778	0,810	0,826	0,831	0,829	0,828
|C4| IG&DF (g)		0,457	0,648	0,713	0,786	0,817	0,820	0,827	0,827	0,831
|C4| IG&Acc2 (g)		0,476	0,686	0,727	0,789	0,807	0,824	0,829	0,833	0,833
|C4| DF&Acc2 (g)		0,451	0,664	0,730	0,787	0,815	0,826	0,827	0,827	0,828
	MAX	0,549	0,709	0,745	0,795	0,822	0,828	0,833	0,834	0,837
	AVERAGE	0,476	0,654	0,712	0,781	0,811	0,824	0,828	0,829	0,832

Micro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (g)		0,541	0,625	0,697	0,783	0,809	0,828	0,826	0,834	0,843
|C5| tf-idf&IG (g)		0,445	0,664	0,723	0,782	0,817	0,816	0,826	0,829	0,832
|C5| tf-idf&DF (g)		0,382	0,640	0,679	0,752	0,779	0,812	0,832	0,826	0,833
|C5| tf-idf&Acc2 (g)		0,549	0,710	0,750	0,783	0,822	0,830	0,828	0,822	0,835
|C5| CHI&IG (g)		0,466	0,622	0,683	0,747	0,803	0,825	0,829	0,825	0,837
|C5| CHI&DF (g)		0,529	0,646	0,718	0,784	0,811	0,831	0,831	0,831	0,835
|C5| CHI&Acc2 (g)		0,344	0,613	0,698	0,751	0,810	0,829	0,829	0,827	0,830
|C5| IG&DF (g)		0,457	0,649	0,723	0,782	0,808	0,827	0,832	0,826	0,830
|C5| IG&Acc2 (g)		0,454	0,686	0,725	0,789	0,806	0,823	0,828	0,835	0,833
|C5| DF&Acc2 (g)		0,447	0,654	0,716	0,786	0,804	0,825	0,827	0,829	0,829
	MAX	0,549	0,710	0,750	0,789	0,822	0,831	0,832	0,835	0,843
	AVERAGE	0,462	0,651	0,711	0,774	0,807	0,824	0,829	0,828	0,834

Micro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (g)		0,541	0,626	0,691	0,785	0,810	0,832	0,829	0,832	0,838
|C6| tf-idf&IG (g)		0,445	0,668	0,727	0,788	0,814	0,825	0,829	0,831	0,831
|C6| tf-idf&DF (g)		0,487	0,643	0,674	0,745	0,784	0,810	0,830	0,826	0,829
|C6| tf-idf&Acc2 (g)		0,549	0,709	0,745	0,792	0,819	0,827	0,828	0,824	0,834
|C6| CHI&IG (g)		0,466	0,625	0,683	0,761	0,805	0,825	0,826	0,824	0,836
|C6| CHI&DF (g)		0,537	0,625	0,722	0,783	0,814	0,828	0,835	0,829	0,830
|C6| CHI&Acc2 (g)		0,344	0,607	0,690	0,751	0,808	0,829	0,827	0,831	0,830
|C6| IG&DF (g)		0,291	0,645	0,718	0,773	0,809	0,828	0,830	0,825	0,828
|C6| IG&Acc2 (g)		0,476	0,665	0,727	0,787	0,808	0,825	0,828	0,837	0,832
|C6| DF&Acc2 (g)		0,451	0,655	0,721	0,785	0,807	0,827	0,827	0,826	0,828
	MAX	0,549	0,709	0,745	0,792	0,819	0,832	0,835	0,837	0,838
	AVERAGE	0,459	0,647	0,710	0,775	0,808	0,826	0,829	0,829	0,832

Micro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (g)		0,541	0,625	0,691	0,791	0,814	0,825	0,828	0,834	0,841
|C7| tf-idf&IG (g)		0,445	0,664	0,714	0,782	0,814	0,824	0,834	0,835	0,831
|C7| tf-idf&DF (g)		0,382	0,643	0,677	0,743	0,791	0,806	0,828	0,829	0,831
|C7| tf-idf&Acc2 (g)		0,549	0,710	0,750	0,787	0,819	0,828	0,834	0,826	0,833
|C7| CHI&IG (g)		0,466	0,622	0,683	0,746	0,803	0,827	0,828	0,827	0,838
|C7| CHI&DF (g)		0,529	0,646	0,691	0,778	0,806	0,830	0,828	0,833	0,834
|C7| CHI&Acc2 (g)		0,344	0,613	0,695	0,748	0,809	0,826	0,831	0,827	0,832
|C7| IG&DF (g)		0,457	0,649	0,728	0,793	0,808	0,827	0,831	0,829	0,829
|C7| IG&Acc2 (g)		0,476	0,665	0,725	0,787	0,811	0,825	0,829	0,833	0,835
|C7| DF&Acc2 (g)		0,447	0,654	0,711	0,775	0,803	0,820	0,829	0,827	0,829
	MAX	0,549	0,710	0,750	0,793	0,819	0,830	0,834	0,835	0,841
	AVERAGE	0,464	0,649	0,706	0,773	0,808	0,824	0,830	0,830	0,833

Table 7.35. Macro F-measure results of the proposed combinations in global policy for LA1 dataset

Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,284	0,528	0,628	0,692	0,715	0,752	0,748	0,753	0,765
CHI (g)		0,318	0,523	0,549	0,651	0,722	0,762	**0,765**	0,757	**0,775**
IG (g)		0,301	0,510	0,603	0,658	0,745	**0,771**	0,764	0,762	0,772
DF (g)		0,117	0,227	0,515	0,588	0,688	0,724	0,756	0,766	0,760
Acc2 (g)		**0,387**	**0,584**	**0,677**	**0,732**	**0,754**	0,769	0,758	**0,768**	0,759
	MAX	0,387	0,584	0,677	0,732	0,754	0,771	0,765	0,768	0,775
Macro-F	Score Combination	10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	**0,397**	0,570	**0,678**	0,724	0,752	**0,771**	**0,768**	**0,770**	0,772
	AVERAGE	0,250	0,488	**0,614**	**0,701**	**0,745**	**0,759**	0,762	**0,761**	**0,765**
Rank Combination	MAX	**0,395**	**0,591**	0,645	0,721	**0,754**	0,769	**0,772**	0,767	0,773
	AVERAGE	**0,314**	**0,533**	0,597	0,697	0,744	**0,759**	**0,763**	**0,761**	**0,765**
Macro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (g)		0,322	0,467	0,568	0,717	0,744	**0,772**	0,764	0,763	0,767
|C1| tf-idf&IG (g)		0,334	0,530	0,611	0,707	**0,755**	0,750	0,751	0,757	0,765
|C1| tf-idf&DF (g)		0,261	0,502	0,547	0,651	0,689	0,740	0,759	0,763	0,762
|C1| tf-idf&Acc2 (g)		**0,397**	0,578	0,660	0,719	**0,760**	0,758	**0,767**	0,754	0,765
|C1| CHI&IG (g)		0,301	0,470	0,562	0,663	0,740	0,765	0,762	0,758	0,764
|C1| CHI&DF (g)		0,338	0,472	0,618	0,710	0,747	0,765	**0,769**	0,762	0,764
|C1| CHI&Acc2 (g)		0,217	0,469	0,587	0,674	0,750	0,770	**0,765**	0,762	0,762
|C1| IG&DF (g)		0,278	0,464	0,589	0,715	0,748	0,763	0,762	0,757	0,760
|C1| IG&Acc2 (g)		0,298	0,549	0,612	0,712	0,741	0,763	0,764	**0,771**	0,760
|C1| DF&Acc2 (g)		0,262	0,536	0,616	0,712	0,748	0,758	**0,765**	0,760	0,761
	MAX	**0,397**	0,578	0,660	0,719	**0,760**	**0,772**	**0,769**	**0,771**	0,767
	AVERAGE	0,301	0,504	0,597	0,698	0,742	0,760	0,763	0,761	0,763
Macro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (g)		0,293	0,485	0,611	0,691	0,728	0,757	0,749	0,751	0,767
|C2| tf-idf&IG (g)		0,267	0,476	0,641	0,694	0,731	0,751	0,757	0,758	0,771
|C2| tf-idf&DF (g)		0,220	0,422	0,599	0,697	0,715	0,745	0,748	0,756	0,764
|C2| tf-idf&Acc2 (g)		**0,402**	0,523	0,668	0,714	0,745	0,756	0,758	0,757	0,768
|C2| CHI&IG (g)		0,301	0,473	0,589	0,653	0,744	0,766	0,765	0,761	0,765
|C2| CHI&DF (g)		0,136	0,452	0,522	0,651	0,731	0,754	**0,769**	0,766	0,762
|C2| CHI&Acc2 (g)		0,297	0,475	0,652	0,716	0,746	0,762	0,761	0,765	0,763
|C2| IG&DF (g)		0,139	0,448	0,592	0,688	0,748	0,758	**0,765**	0,763	0,764
|C2| IG&Acc2 (g)		0,298	0,573	**0,684**	0,725	0,752	0,761	**0,771**	0,766	0,762
|C2| DF&Acc2 (g)		0,134	0,475	0,675	0,717	0,751	0,761	**0,767**	0,766	0,759
	MAX	**0,402**	0,573	**0,684**	0,725	0,752	0,766	**0,771**	0,766	0,771
	AVERAGE	0,249	0,480	0,623	0,695	0,739	0,757	0,761	0,761	0,765
Macro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (g)		0,383	0,469	0,600	0,721	0,752	**0,778**	0,758	0,764	0,772
|C3| tf-idf&IG (g)		**0,388**	0,552	0,599	0,709	0,749	0,745	0,755	0,764	0,763
|C3| tf-idf&DF (g)		0,191	0,534	0,554	0,649	0,724	0,743	**0,765**	0,760	0,763
|C3| tf-idf&Acc2 (g)		**0,397**	**0,591**	0,644	0,709	**0,761**	0,758	0,764	0,754	0,767
|C3| CHI&IG (g)		0,256	0,463	0,561	0,650	0,740	0,769	**0,771**	0,757	0,763
|C3| CHI&DF (g)		0,356	0,520	0,580	0,701	0,736	0,769	0,759	0,762	0,766
|C3| CHI&Acc2 (g)		0,297	0,515	0,592	0,671	0,749	0,768	**0,773**	0,756	0,763
|C3| IG&DF (g)		0,267	0,540	0,588	0,716	0,742	0,763	**0,768**	0,764	0,761
|C3| IG&Acc2 (g)		0,299	0,549	0,615	0,718	0,741	0,763	0,760	0,767	0,767
|C3| DF&Acc2 (g)		0,233	0,536	0,609	0,710	0,751	0,758	0,764	0,761	0,760
	MAX	**0,397**	**0,591**	0,644	0,721	**0,761**	**0,778**	**0,773**	0,767	0,772
	AVERAGE	0,307	0,527	0,594	0,695	0,744	0,761	0,764	0,761	0,765

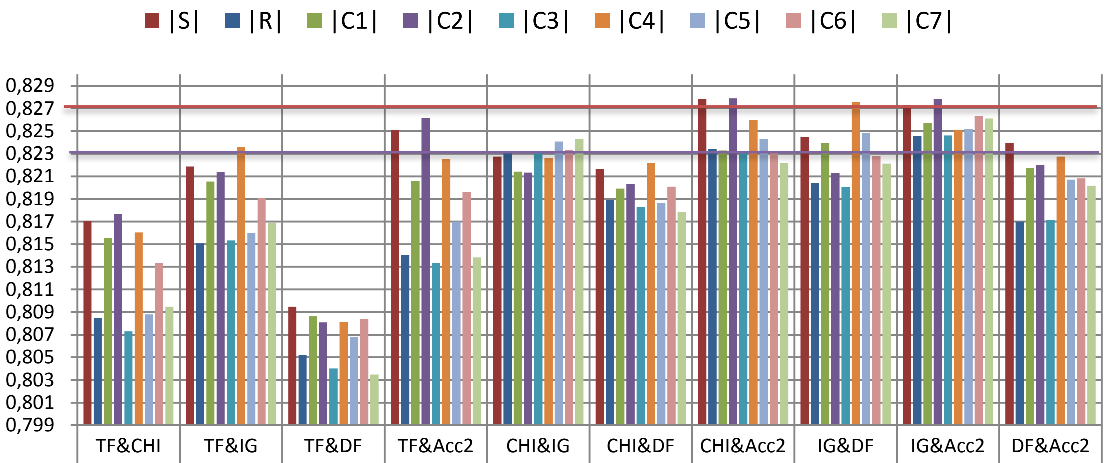
Macro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (g)		0,383	0,467	0,604	0,717	0,750	0,765	0,761	0,767	0,758
|C4| tf-idf&IG (g)		0,334	0,531	0,611	0,713	**0,757**	0,754	0,751	0,757	0,763
|C4| tf-idf&DF (g)		0,273	0,505	0,547	0,656	0,714	0,742	0,760	0,758	0,763
|C4| tf-idf&Acc2 (g)		**0,397**	**0,594**	0,656	0,729	**0,758**	0,758	0,764	0,754	0,767
|C4| CHI&IG (g)		0,256	0,466	0,561	0,662	0,741	0,767	0,764	0,758	0,765
|C4| CHI&DF (g)		0,354	0,478	0,609	0,715	**0,762**	0,766	**0,768**	0,765	0,763
|C4| CHI&Acc2 (g)		0,217	0,469	0,588	0,703	0,750	0,766	**0,767**	0,761	0,759
|C4| IG&DF (g)		0,267	0,497	0,589	0,714	**0,758**	0,756	0,761	0,757	0,762
|C4| IG&Acc2 (g)		0,299	0,549	0,612	0,720	0,746	0,762	**0,766**	0,763	0,761
|C4| DF&Acc2 (g)		0,219	0,530	0,633	0,721	**0,758**	0,762	0,764	0,757	0,759
	MAX	**0,397**	**0,594**	0,656	0,729	**0,762**	0,767	**0,768**	0,767	0,767
	AVERAGE	0,300	0,509	0,601	0,705	0,749	0,760	0,763	0,760	0,762

Macro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (g)		0,383	0,466	0,582	0,715	0,743	0,762	0,755	0,764	**0,779**
|C5| tf-idf&IG (g)		0,334	0,527	0,606	0,709	0,754	0,742	0,754	0,760	0,765
|C5| tf-idf&DF (g)		0,191	0,495	0,551	0,655	0,703	0,743	0,762	0,763	0,763
|C5| tf-idf&Acc2 (g)		**0,397**	**0,593**	0,639	0,706	**0,756**	0,762	0,764	0,754	0,768
|C5| CHI&IG (g)		0,256	0,463	0,561	0,652	0,738	0,768	**0,766**	0,756	0,765
|C5| CHI&DF (g)		0,353	0,521	0,612	0,712	0,752	0,767	**0,768**	0,762	0,766
|C5| CHI&Acc2 (g)		0,217	0,467	0,588	0,671	0,752	0,769	**0,767**	0,754	0,760
|C5| IG&DF (g)		0,267	0,495	0,604	0,709	0,747	0,761	**0,766**	0,757	0,765
|C5| IG&Acc2 (g)		0,252	0,549	0,615	0,720	0,741	0,760	0,764	0,764	0,763
|C5| DF&Acc2 (g)		0,232	0,526	0,611	0,725	0,745	0,760	0,763	0,758	0,757
	MAX	**0,397**	**0,593**	0,639	0,725	**0,756**	0,769	**0,768**	0,764	**0,779**
	AVERAGE	0,288	0,510	0,597	0,697	0,743	0,759	0,763	0,759	0,765

Macro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (g)		0,383	0,467	0,568	0,720	0,741	0,770	0,759	0,762	**0,775**
|C6| tf-idf&IG (g)		0,334	0,530	0,611	0,712	0,751	0,757	0,762	0,762	0,761
|C6| tf-idf&DF (g)		0,273	0,500	0,545	0,647	0,711	0,743	0,760	0,760	0,762
|C6| tf-idf&Acc2 (g)		**0,397**	**0,593**	0,656	0,719	0,753	0,760	**0,768**	0,754	0,767
|C6| CHI&IG (g)		0,256	0,466	0,562	0,663	0,740	0,768	0,764	0,756	0,764
|C6| CHI&DF (g)		0,354	0,464	0,622	0,715	0,752	0,765	**0,771**	0,766	0,763
|C6| CHI&Acc2 (g)		0,217	0,461	0,572	0,670	0,750	0,770	0,763	0,763	0,760
|C6| IG&DF (g)		0,139	0,495	0,596	0,691	0,745	0,764	0,763	0,759	0,761
|C6| IG&Acc2 (g)		0,299	0,533	0,612	0,712	0,744	0,764	0,763	**0,769**	0,762
|C6| DF&Acc2 (g)		0,219	0,529	0,624	0,706	0,747	0,763	0,764	0,759	0,758
	MAX	**0,397**	**0,593**	0,656	0,720	0,753	0,770	**0,771**	**0,769**	**0,775**
	AVERAGE	0,287	0,504	0,597	0,696	0,743	0,762	0,764	0,761	0,763

Macro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (g)		0,383	0,466	0,568	0,722	0,749	0,767	0,759	0,765	0,774
|C7| tf-idf&IG (g)		0,334	0,527	0,594	0,708	**0,754**	0,753	**0,769**	0,767	0,762
|C7| tf-idf&DF (g)		0,191	0,500	0,550	0,646	0,727	0,740	0,760	0,761	0,762
|C7| tf-idf&Acc2 (g)		**0,397**	**0,593**	0,644	0,710	0,753	0,761	**0,773**	0,757	0,764
|C7| CHI&IG (g)		0,256	0,463	0,561	0,649	0,740	**0,771**	0,764	0,756	0,767
|C7| CHI&DF (g)		0,353	0,527	0,574	0,705	0,739	0,767	**0,765**	0,764	0,765
|C7| CHI&Acc2 (g)		0,217	0,467	0,577	0,664	0,751	0,764	**0,772**	0,758	0,763
|C7| IG&DF (g)		0,267	0,495	0,607	0,720	0,746	0,763	**0,766**	0,761	0,765
|C7| IG&Acc2 (g)		0,299	0,533	0,615	0,712	0,748	0,764	0,761	0,766	0,764
|C7| DF&Acc2 (g)		0,232	0,526	0,602	0,708	0,744	0,755	**0,767**	0,759	0,761
	MAX	**0,397**	**0,593**	0,644	0,722	**0,754**	**0,771**	**0,773**	0,767	0,774
	AVERAGE	0,293	0,510	0,589	0,694	0,745	0,760	**0,766**	0,761	0,765

In addition to this, the performances of the combinations *IG & Acc2, CHI & Acc2, IG & DF* and *CHI & IG* are remarkable with some proposed methods especially with a high number of keywords as seen Tables 7.38 and 7.39. Each of them improves the highest F-measure results of the individual metrics and the highest F-measure results of the score and rank combinations. These performances were obtained from the experiments are exactly what we expected from the combining two individual metrics.

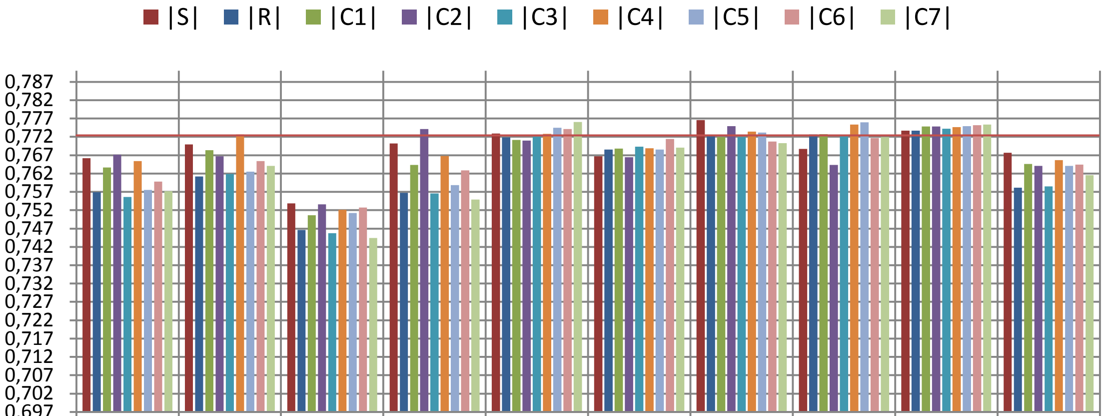


	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,817	0,822	0,809	0,825	0,823	0,822	0,828	0,824	0,827	0,824
|R|	0,808	0,815	0,805	0,814	0,823	0,819	0,823	0,820	0,825	0,817
|C1|	0,816	0,821	0,809	0,821	0,821	0,820	0,823	0,824	0,826	0,822
|C2|	0,818	0,821	0,808	0,826	0,821	0,820	0,828	0,821	0,828	0,822
|C3|	0,807	0,815	0,804	0,813	0,823	0,818	0,823	0,820	0,825	0,817
|C4|	0,816	0,824	0,808	0,823	0,823	0,822	0,826	0,828	0,825	0,823
|C5|	0,809	0,816	0,807	0,817	0,824	0,819	0,824	0,825	0,825	0,821
|C6|	0,813	0,819	0,808	0,820	0,823	0,820	0,823	0,823	0,826	0,821
|C7|	0,809	0,817	0,803	0,814	0,824	0,818	0,822	0,822	0,826	0,820

Figure 7.27. In local policy, averages of the micro-averaged F-measures of all combinations for LA1 dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	0,665	**0,755**	**0,771**	**0,803**	**0,819**	**0,838**	**0,841**	**0,842**	**0,842**
Combination | R |	**0,673**	0,745	0,769	0,790	0,812	0,834	0,839	**0,842**	**0,842**
Combination |C1|	0,662	0,741	0,769	0,802	0,819	0,837	0,839	0,843	0,843
Combination |C2|	0,673	0,746	0,773	0,802	0,822	0,839	0,842	0,843	0,843
Combination |C3|	0,665	0,745	0,769	0,790	0,812	0,833	0,839	0,842	0,844
Combination |C4|	0,661	0,745	0,770	0,802	0,821	0,835	0,838	0,844	0,843
Combination |C5|	0,664	0,744	0,770	0,793	0,814	0,836	0,842	0,843	0,842
Combination |C6|	0,661	0,743	0,767	0,796	0,818	0,835	0,838	0,841	0,843
Combination |C7|	0,664	0,747	0,769	0,798	0,815	0,832	0,838	0,841	0,845

Table 7.36. In local policy, maximum micro-averaged F-measures of all combinations for LA1 dataset



	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,766	0,770	0,754	0,770	0,773	0,767	0,777	0,769	0,774	0,768
|R|	0,757	0,761	0,747	0,757	0,772	0,769	0,773	0,773	0,774	0,758
|C1|	0,764	0,768	0,751	0,764	0,771	0,769	0,772	0,773	0,775	0,765
|C2|	0,767	0,767	0,754	0,774	0,771	0,766	0,775	0,764	0,775	0,764
|C3|	0,756	0,762	0,746	0,757	0,772	0,769	0,772	0,772	0,774	0,759
|C4|	0,765	0,773	0,752	0,767	0,773	0,769	0,773	0,775	0,775	0,766
|C5|	0,758	0,763	0,751	0,759	0,774	0,769	0,773	0,776	0,775	0,764
|C6|	0,760	0,765	0,753	0,763	0,774	0,771	0,771	0,772	0,775	0,764
|C7|	0,757	0,764	0,744	0,755	0,776	0,769	0,770	0,772	0,775	0,762

Figure 7.28. In local policy, averages of the macro-averaged F-measures of all combinations for LA1 dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	0,595	**0,698**	0,724	**0,752**	**0,776**	0,793	**0,790**	0,787	**0,790**
Combination | R |	**0,602**	0,692	**0,725**	0,749	0,767	0,783	0,789	**0,791**	0,787
Combination |C1|	0,592	0,693	0,717	0,748	0,773	0,788	0,791	0,790	0,784
Combination |C2|	0,599	0,698	0,726	0,750	0,779	0,785	0,795	0,783	0,786
Combination |C3|	0,596	0,692	0,726	0,749	0,769	0,782	0,790	0,791	0,788
Combination |C4|	0,592	0,693	0,721	0,751	0,777	0,786	0,791	0,789	0,787
Combination |C5|	0,597	0,691	0,721	0,754	0,768	0,789	0,793	0,791	0,789
Combination |C6|	0,592	0,690	0,722	0,745	0,772	0,786	0,790	0,787	0,794
Combination |C7|	0,597	0,693	0,723	0,748	0,768	0,783	0,790	0,788	0,796

Table 7.37. In local policy, maximum macro-averaged F-measures of all combinations for LA1 dataset

In the case of local policy performance of the rank combination was apparently better than the performance of the score combination as we tested before. According to results of the new experiments, we can say that the performances of the all proposed methods outperform rank combination. Among them C2, C4 and C5 also outperform the success of the score combination when we compare the micro- and macro-averaged F-measure results. Significantly C2 is the best proposed method that improves the existing highest F-measure value of the different number of keywords.

One of the important results of these experiments is that all proposed methods improve the highest F-measure values of the individual metrics as seen in Tables 7.38 and 7.39. From the previous experiments we determined that the highest F-measure values were achieved by CHI in local policy with 1500 keywords, with value of 84% micro- and 78.8% macro-averaged F-measure. Afterward score combination of *CHI & Acc2* outperformed this highest performance with values of 84.2% micro-averaged F-measure with 1500 keywords and 79.3% macro-averaged F-measure with 500 keywords. When we analyze our proposed methods, the highest micro-averaged F-measure of the rank combination is improved by all proposed methods and the highest macro-averaged F-measure of the rank combination improved by C_2, C_6 and C_7 as seen Tables 7.33 and 7.34 Among the proposed methods the best performance is achieved by C_7 of *CHI & IG* with value of 84.3% micro- and 79.6% macro-averaged F-measure with 2000 keywords and another method C_2 of *tf-idf & Acc2* achieves 84.2% micro- and 79.5% macro-averaged F-measure with only 1000 keywords. C_7 of *IG & Acc2* actually has the highest 84.5% micro-averaged F-measure value but its performance in the case of macro-averaged F-measure is low about 78.6%.

Table 7.38. Micro F-measure results of the proposed combinations in local policy for LA1 dataset

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (l)		0,631	0,731	0,761	0,785	0,789	0,807	0,814	0,812	0,815
CHI (l)		0,671	0,736	0,761	0,788	0,813	0,823	0,833	0,840	0,838
IG (l)		0,660	0,739	0,765	0,793	0,807	0,830	0,831	0,831	0,833
DF (l)		0,318	0,688	0,740	0,766	0,782	0,814	0,815	0,827	0,826
Acc2 (l)		0,659	0,742	0,764	0,802	0,817	0,829	0,835	0,840	0,840
	MAX	0,671	0,742	0,765	0,802	0,817	0,830	0,835	0,840	0,840
Micro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	0,665	0,755	0,771	0,803	0,819	0,838	0,841	0,842	0,842
	AVERAGE	0,637	0,735	0,762	0,790	0,810	0,828	0,836	0,835	0,834
Rank Combination	MAX	0,673	0,745	0,769	0,790	0,812	0,834	0,839	0,842	0,842
	AVERAGE	0,654	0,735	0,759	0,779	0,802	0,823	0,831	0,833	0,835
Micro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (l)		0,646	0,735	0,758	0,787	0,799	0,825	0,824	0,825	0,833
|C1| tf-idf&IG (l)		0,650	0,741	0,767	0,769	0,819	0,837	0,830	0,832	0,836
|C1| tf-idf&DF (l)		0,645	0,728	0,766	0,778	0,787	0,808	0,820	0,826	0,833
|C1| tf-idf&Acc2 (l)		0,651	0,732	0,763	0,786	0,802	0,828	0,836	0,836	0,836
|C1| CHI&IG (l)		0,655	0,740	0,763	0,784	0,813	0,823	0,837	0,834	0,838
|C1| CHI&DF (l)		0,651	0,734	0,761	0,781	0,800	0,835	0,829	0,839	0,836
|C1| CHI&Acc2 (l)		0,661	0,737	0,766	0,784	0,807	0,834	0,832	0,843	0,840
|C1| IG&DF (l)		0,638	0,740	0,761	0,791	0,809	0,827	0,839	0,836	0,840
|C1| IG&Acc2 (l)		0,662	0,733	0,769	0,791	0,817	0,826	0,838	0,840	0,843
|C1| DF&Acc2 (l)		0,643	0,730	0,756	0,802	0,815	0,823	0,835	0,827	0,828
	MAX	0,662	0,741	0,769	0,802	0,819	0,837	0,839	0,843	0,843
	AVERAGE	0,650	0,735	0,763	0,785	0,807	0,827	0,832	0,834	0,836
Micro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (l)		0,658	0,734	0,767	0,789	0,805	0,823	0,829	0,829	0,831
|C2| tf-idf&IG (l)		0,655	0,744	0,773	0,795	0,812	0,828	0,834	0,829	0,830
|C2| tf-idf&DF (l)		0,595	0,712	0,751	0,778	0,795	0,812	0,824	0,819	0,820
|C2| tf-idf&Acc2 (l)		0,650	0,746	0,770	0,798	0,821	0,827	0,842	0,834	0,835
|C2| CHI&IG (l)		0,664	0,735	0,761	0,782	0,809	0,823	0,838	0,838	0,837
|C2| CHI&DF (l)		0,604	0,716	0,765	0,796	0,805	0,829	0,831	0,829	0,833
|C2| CHI&Acc2 (l)		0,673	0,741	0,767	0,787	0,822	0,835	0,837	0,843	0,843
|C2| IG&DF (l)		0,622	0,714	0,768	0,799	0,803	0,824	0,835	0,836	0,832
|C2| IG&Acc2 (l)		0,656	0,740	0,765	0,795	0,819	0,839	0,838	0,838	0,838
|C2| DF&Acc2 (l)		0,600	0,720	0,762	0,802	0,815	0,823	0,833	0,831	0,829
	MAX	0,673	0,746	0,773	0,802	0,822	0,839	0,842	0,843	0,843
	AVERAGE	0,638	0,730	0,765	0,792	0,811	0,826	0,834	0,833	0,833
Micro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (l)		0,644	0,731	0,754	0,767	0,795	0,810	0,821	0,823	0,828
|C3| tf-idf&IG (l)		0,649	0,736	0,752	0,770	0,800	0,827	0,830	0,831	0,833
|C3| tf-idf&DF (l)		0,654	0,733	0,760	0,769	0,783	0,801	0,818	0,823	0,830
|C3| tf-idf&Acc2 (l)		0,654	0,730	0,749	0,765	0,794	0,828	0,828	0,833	0,832
|C3| CHI&IG (l)		0,656	0,745	0,762	0,786	0,812	0,823	0,839	0,837	0,842
|C3| CHI&DF (l)		0,656	0,740	0,759	0,777	0,802	0,827	0,834	0,832	0,838
|C3| CHI&Acc2 (l)		0,664	0,736	0,769	0,788	0,805	0,833	0,830	0,842	0,841
|C3| IG&DF (l)		0,650	0,728	0,756	0,784	0,801	0,825	0,833	0,841	0,836
|C3| IG&Acc2 (l)		0,665	0,740	0,769	0,790	0,810	0,831	0,836	0,837	0,844
|C3| DF&Acc2 (l)		0,642	0,725	0,760	0,786	0,805	0,819	0,836	0,825	0,831
	MAX	0,665	0,745	0,769	0,790	0,812	0,833	0,839	0,842	0,844
	AVERAGE	0,653	0,734	0,759	0,778	0,801	0,822	0,831	0,832	0,835

Micro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (I)		0,645	0,735	0,762	0,774	0,802	0,826	0,830	0,830	0,835
|C4| tf-idf&IG (I)		0,650	0,741	**0,767**	0,770	**0,821**	**0,835**	**0,838**	0,837	**0,840**
|C4| tf-idf&DF (I)		0,649	0,730	**0,770**	0,777	0,786	0,806	0,823	0,828	0,829
|C4| tf-idf&Acc2 (I)		0,650	0,729	0,764	0,783	0,812	**0,832**	**0,835**	0,837	0,835
|C4| CHI&IG (I)		0,655	**0,743**	0,764	0,784	0,811	0,825	**0,838**	0,837	**0,841**
|C4| CHI&DF (I)		0,652	**0,742**	0,757	0,784	0,811	0,829	0,832	0,836	**0,840**
|C4| CHI&Acc2 (I)		0,661	0,737	**0,767**	0,788	0,816	**0,835**	0,832	**0,844**	**0,841**
|C4| IG&DF (I)		0,640	**0,745**	0,763	0,794	**0,817**	**0,833**	**0,835**	**0,842**	**0,843**
|C4| IG&Acc2 (I)		0,657	0,736	**0,769**	0,792	0,813	0,829	**0,838**	0,836	**0,843**
|C4| DF&Acc2 (I)		0,642	0,737	0,756	**0,802**	0,813	0,824	0,832	0,835	0,830
	MAX	0,661	**0,745**	**0,770**	0,802	**0,821**	**0,835**	**0,838**	**0,844**	**0,843**
	AVERAGE	0,650	0,737	0,764	0,785	0,810	0,827	0,833	0,836	0,838

Micro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (I)		0,644	0,734	0,757	0,771	0,796	0,814	0,820	0,826	0,826
|C5| tf-idf&IG (I)		0,649	0,739	0,756	0,766	0,809	0,829	0,830	0,828	0,834
|C5| tf-idf&DF (I)		0,662	0,730	0,764	0,770	0,785	0,803	0,825	0,825	0,833
|C5| tf-idf&Acc2 (I)		0,654	0,730	0,752	0,777	0,795	0,827	0,834	0,836	0,832
|C5| CHI&IG (I)		0,655	**0,744**	0,763	0,783	0,814	0,827	**0,842**	0,839	0,839
|C5| CHI&DF (I)		0,648	0,739	0,760	0,778	0,804	**0,830**	0,830	0,830	0,839
|C5| CHI&Acc2 (I)		0,664	0,741	**0,767**	0,787	0,809	**0,836**	0,829	**0,843**	**0,842**
|C5| IG&DF (I)		0,640	0,740	0,755	0,793	0,806	**0,834**	**0,837**	**0,842**	0,837
|C5| IG&Acc2 (I)		0,657	0,740	**0,770**	0,793	0,814	0,827	**0,838**	0,837	**0,842**
|C5| DF&Acc2 (I)		0,642	0,725	0,758	0,792	0,808	0,827	**0,839**	0,829	0,828
	MAX	0,664	**0,744**	**0,770**	0,793	0,814	**0,836**	**0,842**	**0,843**	**0,842**
	AVERAGE	0,651	0,736	0,760	0,781	0,804	0,825	0,833	0,833	0,835

Micro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (I)		0,643	0,733	0,760	0,777	0,799	0,825	0,823	0,826	0,830
|C6| tf-idf&IG (I)		0,650	0,739	0,763	0,771	0,810	**0,835**	0,832	0,834	0,833
|C6| tf-idf&DF (I)		0,649	0,731	0,762	0,774	0,787	0,808	0,819	0,828	0,833
|C6| tf-idf&Acc2 (I)		0,650	0,729	0,758	0,779	0,806	0,826	**0,836**	0,836	0,835
|C6| CHI&IG (I)		0,655	**0,743**	0,763	0,783	0,814	0,824	**0,838**	0,838	**0,843**
|C6| CHI&DF (I)		0,652	0,741	0,756	0,781	0,806	**0,833**	0,830	0,833	0,837
|C6| CHI&Acc2 (I)		0,661	0,740	**0,766**	0,788	0,808	**0,833**	0,833	**0,840**	0,839
|C6| IG&DF (I)		0,640	0,738	0,761	0,783	0,804	**0,830**	**0,838**	**0,841**	**0,841**
|C6| IG&Acc2 (I)		0,656	0,736	**0,767**	0,795	**0,818**	0,828	**0,838**	0,839	**0,840**
|C6| DF&Acc2 (I)		0,642	0,733	0,755	0,796	0,812	0,825	**0,836**	0,829	0,827
	MAX	0,661	**0,743**	**0,767**	0,796	**0,818**	**0,835**	**0,838**	**0,841**	**0,843**
	AVERAGE	0,650	0,736	0,761	0,783	0,806	0,827	0,832	0,834	0,836

Micro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (I)		0,644	0,736	0,757	0,772	0,795	0,818	0,819	0,826	0,828
|C7| tf-idf&IG (I)		0,649	0,738	0,755	0,769	0,805	**0,830**	0,833	0,829	0,835
|C7| tf-idf&DF (I)		0,662	0,733	0,758	0,766	0,780	0,805	0,815	0,825	0,830
|C7| tf-idf&Acc2 (I)		0,650	0,730	0,753	0,774	0,790	0,825	0,834	0,832	0,829
|C7| CHI&IG (I)		0,655	**0,747**	0,763	0,788	0,815	0,826	**0,836**	0,838	**0,843**
|C7| CHI&DF (I)		0,648	0,734	0,763	0,776	0,803	0,827	0,831	0,834	0,836
|C7| CHI&Acc2 (I)		0,664	0,741	**0,766**	0,787	0,804	**0,832**	0,828	**0,841**	**0,841**
|C7| IG&DF (I)		0,636	0,739	0,756	0,783	0,807	**0,831**	0,833	0,839	0,839
|C7| IG&Acc2 (I)		0,657	0,740	**0,769**	0,798	0,812	**0,831**	**0,837**	0,835	**0,845**
|C7| DF&Acc2 (I)		0,642	0,729	0,757	0,795	0,806	0,826	**0,838**	0,829	0,827
	MAX	0,664	**0,747**	**0,769**	0,798	0,815	**0,832**	**0,838**	**0,841**	**0,845**
	AVERAGE	0,651	0,737	0,759	0,781	0,802	0,825	0,830	0,833	0,835

Table 7.39. Macro F-measure results of the proposed combinations in local policy for LA1 dataset

Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (l)		0,552	0,674	0,706	0,728	0,735	0,755	0,762	0,756	0,764
CHI (l)		**0,607**	0,686	**0,715**	0,741	0,766	0,772	0,778	**0,788**	**0,785**
IG (l)		0,578	**0,688**	0,714	0,743	0,756	**0,781**	**0,779**	0,775	0,777
DF (l)		0,376	0,582	0,665	0,702	0,715	0,755	0,758	0,767	0,768
Acc2 (l)		0,546	0,682	0,712	**0,759**	**0,773**	0,770	0,774	0,776	0,781
	MAX	0,607	0,688	0,715	0,759	0,773	0,781	0,779	0,788	0,785
Macro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	**MAX**	0,595	**0,698**	**0,724**	0,752	**0,776**	**0,793**	**0,790**	0,787	**0,790**
	AVERAGE	0,554	0,678	**0,712**	**0,737**	**0,760**	**0,775**	**0,782**	**0,780**	0,778
Rank Combination	**MAX**	0,602	**0,692**	**0,725**	0,749	0,767	**0,783**	**0,789**	**0,791**	**0,787**
	AVERAGE	**0,575**	**0,681**	0,708	0,730	0,749	0,769	0,778	0,778	**0,779**
Macro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (l)		0,574	0,684	0,711	0,741	0,746	0,774	0,766	0,771	0,783
|C1| tf-idf&IG (l)		0,583	**0,693**	0,713	0,709	0,773	**0,788**	0,778	0,780	0,782
|C1| tf-idf&DF (l)		0,563	0,658	0,704	0,718	0,725	0,752	0,765	0,769	0,774
|C1| tf-idf&Acc2 (l)		0,582	0,679	0,712	0,729	0,744	0,770	**0,785**	0,778	0,781
|C1| CHI&IG (l)		0,586	0,686	0,713	0,739	0,764	0,774	**0,781**	0,785	0,784
|C1| CHI&DF (l)		0,582	0,683	0,712	0,735	0,754	**0,785**	0,777	0,786	0,776
|C1| CHI&Acc2 (l)		0,592	0,684	0,715	0,732	0,762	**0,784**	**0,782**	**0,790**	0,781
|C1| IG&DF (l)		0,560	0,678	0,710	0,745	0,763	0,774	**0,791**	0,781	0,783
|C1| IG&Acc2 (l)		0,578	0,682	**0,717**	0,742	0,770	0,776	**0,790**	**0,788**	0,784
|C1| DF&Acc2 (l)		0,549	0,670	0,699	0,748	0,760	0,768	**0,779**	0,768	0,764
	MAX	0,592	**0,693**	**0,717**	0,748	0,773	**0,788**	**0,791**	**0,790**	0,784
	AVERAGE	0,575	0,680	0,711	0,734	0,756	0,774	0,779	0,780	0,779
Macro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (l)		0,594	0,681	**0,717**	0,742	0,761	0,766	0,773	0,780	0,781
|C2| tf-idf&IG (l)		0,582	**0,698**	**0,726**	0,741	0,760	0,774	0,776	0,775	0,774
|C2| tf-idf&DF (l)		0,506	0,641	0,685	0,720	0,733	0,762	0,770	0,766	0,769
|C2| tf-idf&Acc2 (l)		0,579	**0,692**	**0,716**	0,750	0,771	0,769	**0,795**	0,781	0,780
|C2| CHI&IG (l)		0,599	0,681	0,710	0,736	0,762	0,774	**0,785**	0,782	**0,786**
|C2| CHI&DF (l)		0,530	0,653	0,707	0,742	0,757	0,779	0,776	0,766	0,778
|C2| CHI&Acc2 (l)		0,593	**0,689**	**0,715**	0,735	**0,779**	**0,785**	**0,783**	0,783	0,784
|C2| IG&DF (l)		0,516	0,660	0,714	0,743	0,746	0,761	**0,780**	0,782	0,775
|C2| IG&Acc2 (l)		0,571	**0,690**	**0,715**	0,746	**0,777**	**0,782**	**0,788**	0,779	0,777
|C2| DF&Acc2 (l)		0,475	0,652	0,712	0,746	0,766	0,762	0,775	0,773	0,763
	MAX	0,599	**0,698**	**0,726**	0,750	**0,779**	**0,785**	**0,795**	0,783	**0,786**
	AVERAGE	0,555	0,674	0,712	0,740	0,761	0,771	0,780	0,777	0,777
Macro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (l)		0,575	0,684	0,706	0,721	0,744	0,758	0,766	0,768	0,776
|C3| tf-idf&IG (l)		0,581	0,685	0,701	0,717	0,751	0,777	0,777	0,775	0,774
|C3| tf-idf&DF (l)		0,559	0,672	0,697	0,714	0,717	0,737	0,763	0,768	0,775
|C3| tf-idf&Acc2 (l)		0,584	0,676	0,698	0,709	0,729	0,769	0,778	0,777	0,777
|C3| CHI&IG (l)		0,586	**0,692**	0,712	0,740	0,769	0,769	**0,784**	0,783	**0,788**
|C3| CHI&DF (l)		0,588	**0,691**	0,712	0,738	0,755	0,778	**0,782**	0,781	0,782
|C3| CHI&Acc2 (l)		0,596	0,683	**0,726**	0,741	0,757	**0,782**	**0,779**	**0,788**	**0,785**
|C3| IG&DF (l)		0,556	0,676	0,703	0,749	0,755	0,773	**0,783**	**0,791**	0,784
|C3| IG&Acc2 (l)		0,577	0,687	**0,716**	0,742	0,765	**0,781**	**0,790**	0,781	**0,786**
|C3| DF&Acc2 (l)		0,545	0,664	0,704	0,726	0,749	0,761	**0,779**	0,764	0,772
	MAX	0,596	**0,692**	**0,726**	0,749	0,769	**0,782**	**0,790**	**0,791**	**0,788**
	AVERAGE	0,575	0,681	0,707	0,730	0,749	0,768	0,778	0,778	0,780

Macro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (I)		0,575	0,682	0,711	0,731	0,756	0,770	0,773	0,780	0,783
|C4| tf-idf&IG (I)		0,583	0,693	0,714	0,716	0,777	0,784	0,782	0,789	0,786
|C4| tf-idf&DF (I)		0,567	0,661	0,706	0,718	0,726	0,748	0,773	0,772	0,776
|C4| tf-idf&Acc2 (I)		0,581	0,676	0,713	0,726	0,754	0,777	0,785	0,781	0,778
|C4| CHI&IG (I)		0,586	0,690	0,714	0,738	0,763	0,776	0,786	0,787	0,787
|C4| CHI&DF (I)		0,584	0,687	0,705	0,738	0,752	0,778	0,780	0,781	0,785
|C4| CHI&Acc2 (I)		0,592	0,686	0,721	0,736	0,769	0,786	0,781	0,784	0,784
|C4| IG&DF (I)		0,562	0,684	0,712	0,751	0,770	0,778	0,784	0,786	0,783
|C4| IG&Acc2 (I)		0,573	0,684	0,717	0,743	0,766	0,781	0,791	0,783	0,785
|C4| DF&Acc2 (I)		0,545	0,674	0,702	0,749	0,763	0,765	0,772	0,775	0,769
	MAX	0,592	0,693	0,721	0,751	0,777	0,786	0,791	0,789	0,787
	AVERAGE	0,575	0,682	0,712	0,735	0,760	0,774	0,781	0,782	0,781
Macro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (I)		0,575	0,686	0,716	0,722	0,746	0,769	0,761	0,771	0,775
|C5| tf-idf&IG (I)		0,581	0,688	0,710	0,712	0,761	0,779	0,774	0,770	0,780
|C5| tf-idf&DF (I)		0,585	0,663	0,701	0,716	0,725	0,745	0,773	0,770	0,778
|C5| tf-idf&Acc2 (I)		0,584	0,677	0,707	0,717	0,734	0,766	0,782	0,779	0,775
|C5| CHI&IG (I)		0,586	0,691	0,713	0,737	0,768	0,779	0,793	0,787	0,783
|C5| CHI&DF (I)		0,581	0,689	0,713	0,736	0,758	0,783	0,780	0,774	0,781
|C5| CHI&Acc2 (I)		0,597	0,689	0,721	0,737	0,760	0,789	0,779	0,785	0,789
|C5| IG&DF (I)		0,562	0,686	0,702	0,754	0,762	0,782	0,787	0,791	0,780
|C5| IG&Acc2 (I)		0,572	0,687	0,718	0,744	0,767	0,778	0,790	0,785	0,786
|C5| DF&Acc2 (I)		0,545	0,663	0,698	0,738	0,757	0,767	0,782	0,771	0,769
	MAX	0,597	0,691	0,721	0,754	0,768	0,789	0,793	0,791	0,789
	AVERAGE	0,577	0,682	0,710	0,731	0,754	0,774	0,780	0,778	0,779
Macro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (I)		0,573	0,679	0,712	0,732	0,740	0,772	0,764	0,768	0,782
|C6| tf-idf&IG (I)		0,583	0,688	0,713	0,710	0,760	0,786	0,774	0,780	0,781
|C6| tf-idf&DF (I)		0,567	0,661	0,700	0,718	0,729	0,753	0,767	0,773	0,777
|C6| tf-idf&Acc2 (I)		0,581	0,676	0,709	0,719	0,746	0,769	0,786	0,778	0,779
|C6| CHI&IG (I)		0,586	0,690	0,713	0,737	0,765	0,776	0,787	0,787	0,794
|C6| CHI&DF (I)		0,584	0,687	0,707	0,736	0,766	0,785	0,778	0,783	0,781
|C6| CHI&Acc2 (I)		0,592	0,690	0,722	0,736	0,762	0,782	0,782	0,782	0,780
|C6| IG&DF (I)		0,562	0,683	0,714	0,738	0,756	0,778	0,787	0,785	0,786
|C6| IG&Acc2 (I)		0,571	0,684	0,715	0,744	0,772	0,778	0,790	0,784	0,783
|C6| DF&Acc2 (I)		0,545	0,672	0,700	0,745	0,763	0,769	0,779	0,770	0,761
	MAX	0,592	0,690	0,722	0,745	0,772	0,786	0,790	0,787	0,794
	AVERAGE	0,574	0,681	0,711	0,731	0,756	0,775	0,779	0,779	0,780
Macro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (I)		0,575	0,688	0,708	0,723	0,739	0,769	0,764	0,771	0,778
|C7| tf-idf&IG (I)		0,581	0,687	0,708	0,713	0,757	0,783	0,777	0,775	0,781
|C7| tf-idf&DF (I)		0,585	0,670	0,694	0,709	0,715	0,742	0,760	0,768	0,772
|C7| tf-idf&Acc2 (I)		0,581	0,677	0,707	0,715	0,723	0,762	0,785	0,775	0,770
|C7| CHI&IG (I)		0,585	0,693	0,713	0,744	0,768	0,778	0,781	0,788	0,796
|C7| CHI&DF (I)		0,581	0,684	0,715	0,730	0,760	0,778	0,781	0,785	0,781
|C7| CHI&Acc2 (I)		0,597	0,691	0,723	0,737	0,756	0,781	0,776	0,785	0,787
|C7| IG&DF (I)		0,539	0,684	0,700	0,742	0,760	0,780	0,780	0,783	0,786
|C7| IG&Acc2 (I)		0,572	0,687	0,717	0,748	0,767	0,780	0,790	0,782	0,786
|C7| DF&Acc2 (I)		0,545	0,666	0,698	0,743	0,751	0,766	0,780	0,769	0,761
	MAX	0,597	0,693	0,723	0,748	0,768	0,783	0,790	0,788	0,796
	AVERAGE	0,574	0,683	0,708	0,730	0,750	0,772	0,777	0,778	0,780

7.3. Highly Skew Datasets

Highly skew datasets are particularly hard to categorize since the rare classes are dominated by the common classes whereas we know that they are more likely to reflect the real world datasets (models) which are not uniformly distributed. Thus, we evaluate our methods on two document collection which have highly skew class distributions. The first one is the Wap dataset and the second dataset is the Reuters-21578 dataset. Both of them are very popular document collections that have been used by many researchers in the literature.

One of the main difficulty of the classification on highly skew dataset, the standard classification algorithms tend to be overwhelmed by the major categories and ignore the minor ones which lead to misclassification of minority classes [17]. Class distribution of these datasets is not uniform as we mentioned. For the Wap dataset maximum class size is 341 and minimum class size is 5 and for the Reuters dataset this range is higher than Wap, the maximum class size is 3987 and the minimum class size is only 2.

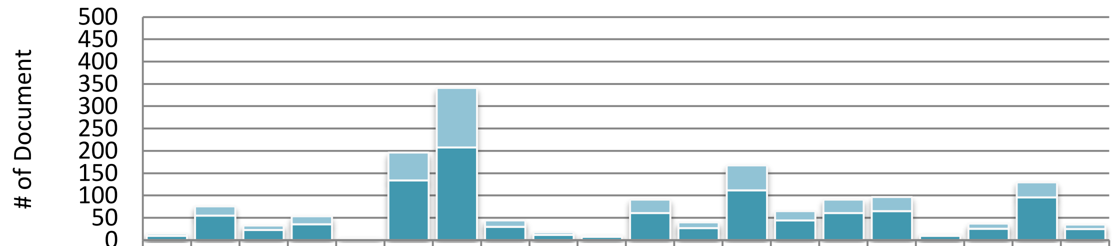
The other difficulty of the classification on highly skew dataset is the overlap between the similar categories. Both Wap and Reuters datasets consist of general topics that are very close to each other and share many common terms. For instance, the Wap dataset contains documents about people and variety which are very general topics and close to each other. Likewise cable and television, multimedia and media or the topics business and industry share many common terms. There is a strong semantic overlap between the topics. This property is similar to the Reuters dataset. For example; the topics related to oil; veg-oil, lin-oil, sun-oil, palm-oil, cotton-oil and castor-oil, and the topics related to metal; gold, platinum, silver, nickel and zine are the good examples of the similar categories. In addition, only about 10% of terms in the Wap dataset and 25% of terms in the Reuters dataset belong to one topic.

Firstly, we begin with the Wap dataset and then we continue to our analysis on the Reuters dataset.

7.3.1. The Wap Dataset

7.3.1.1. Property of the Dataset

The first highly skew corpus chosen for our study is the Wap dataset. This dataset consists of 1560 Web pages from Yahoo! Subject hierarchy collected for the WebACE project, an agent for document categorization and exploration that operates on Web documents. [19] Approximately two thirds of the Web pages (1047) is selected for the training set and the remaining one third (513) is used for testing.



	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20
Test	5	21	10	18	1	62	133	14	6	3	30	13	56	21	30	32	3	11	34	10
Training	10	55	23	36	4	134	208	30	12	8	61	27	112	44	61	65	10	26	96	25
All	15	76	33	54	5	196	341	44	18	11	91	40	168	65	91	97	13	37	130	35

Figure 7.29. Property of the Wap dataset

[C1]	art	15	[C6]	**film**	196	[C11]	**music**	91	[C16]	**sports**	97
[C2]	**business**	76	[C7]	**health**	341	[C12]	online	40	[C17]	stage	13
[C3]	cable	33	[C8]	industry	44	[C13]	**people**	168	[C18]	technology	37
[C4]	**culture**	54	[C9]	media	18	[C14]	**politics**	65	[C19]	**television**	130
[C5]	entertainment	5	[C10]	multimedia	11	[C15]	**review**	91	[C20]	variety	35
Training	1047										
Test	513										
All	1560										

Table 7.40. 20 Categories of the Wap dataset

The Wap dataset is classified into 20 different categories that are very close to each other. Table 7.40 shows these categories along with the number of documents in the dataset and is described in more detail in Figure 7.29. Distribution of the classes is not uniform as seen. Minimum class size is 5 and maximum class size is 341 while average class size is 78. In addition, the 10 top categories with the most number of documents constitute about 85 percent of the dataset and the remaining 10 categories constitute only about 15 percent of all documents.

7.3.1.2. Analysis of the Existing Metrics

In this section we evaluate the performance of the existing metrics over the Wap dataset both in the case of global policy and in the case of local policy. Table 7.41 shows the micro- and macro-averaged F-measure results of the 5 well known feature selection metrics for the Wap dataset with varying number of keywords and the success rates of the existing metrics under feature number criterion is demonstrated in Figure 7.30.

Micro-F	10	30	50	100	200	500	1000	1500	2000	All
tf-idf (g)	0,134	0,496	**0,587**	**0,655**	0,691	0,721	0,740	0,749	0,743	0,752
CHI (g)	0,242	0,523	0,540	0,607	0,631	0,712	0,730	0,741	0,749	0,752
IG (g)	**0,399**	**0,526**	0,577	0,644	**0,746**	**0,753**	0,755	0,756	0,755	0,752
DF (g)	0,000	0,341	0,395	0,543	0,657	0,723	**0,756**	**0,757**	**0,758**	0,752
Acc2 (g)	0,221	0,476	0,529	0,629	0,697	0,730	0,743	0,753	**0,758**	0,752
MAX	0,399	0,526	0,587	0,655	0,746	0,753	0,756	0,757	0,758	
Macro-F	10	30	50	100	200	500	1000	1500	2000	All
tf-idf (g)	0,093	0,208	**0,306**	0,350	0,412	0,442	0,455	0,468	0,455	0,450
CHI (g)	**0,121**	0,185	0,256	0,336	0,375	0,451	**0,486**	0,469	0,478	0,450
IG (g)	0,052	0,185	0,284	0,375	**0,479**	**0,501**	0,473	0,474	0,467	0,450
DF (g)	0,000	0,053	0,095	0,237	0,326	0,430	0,474	0,470	0,462	0,450
Acc2 (g)	0,117	**0,235**	0,278	**0,411**	**0,479**	0,489	0,480	**0,480**	**0,492**	0,450
MAX	0,121	0,235	0,306	0,411	0,479	0,501	0,486	0,480	0,492	
Micro-F	10	30	50	100	200	500	1000	1500	2000	All
tf-idf (l)	0,671	**0,737**	0,741	0,738	0,735	0,722	0,746	0,741	0,749	0,752
CHI (l)	0,440	0,714	0,732	0,732	0,720	0,736	0,742	0,756	**0,758**	0,752
IG (l)	**0,685**	0,735	0,750	0,742	0,747	0,744	0,742	0,758	0,749	0,752
DF (l)	0,000	0,567	0,704	0,751	**0,771**	0,747	**0,760**	0,747	0,747	0,752
Acc2 (l)	0,639	0,728	**0,757**	**0,770**	0,758	**0,755**	0,752	**0,758**	0,752	0,752
MAX	0,685	0,737	0,757	0,770	0,771	0,755	0,760	0,758	0,758	
Macro-F	10	30	50	100	200	500	1000	1500	2000	All
tf-idf (l)	**0,506**	**0,593**	**0,565**	0,532	0,507	0,495	0,509	0,477	0,483	0,450
CHI (l)	0,493	0,511	0,520	0,509	0,462	0,491	0,475	0,488	**0,491**	0,450
IG (l)	0,492	0,531	0,548	0,517	0,508	0,508	0,460	0,490	0,482	0,450
DF (l)	0,000	0,353	0,513	0,550	**0,538**	0,483	**0,524**	0,483	0,481	0,450
Acc2 (l)	0,435	0,554	0,564	**0,551**	0,509	**0,516**	0,488	**0,491**	0,485	0,450
MAX	0,506	0,593	0,565	0,551	0,538	0,516	0,524	0,491	0,491	

Table 7.41. Micro- and macro-averaged F-measures for Wap dataset

The first notable result of the analysis is an apparent increasing gap between micro- and macro-averaged F-measure results. In this dataset, the average of the difference between two types of F-measure values is about 22-27%. This difference is significantly higher than the previous datasets used in our study. As we mentioned the Wap dataset has highly skew class distributions. Thus this marked difference between two values due to the poor performances over rare categories.

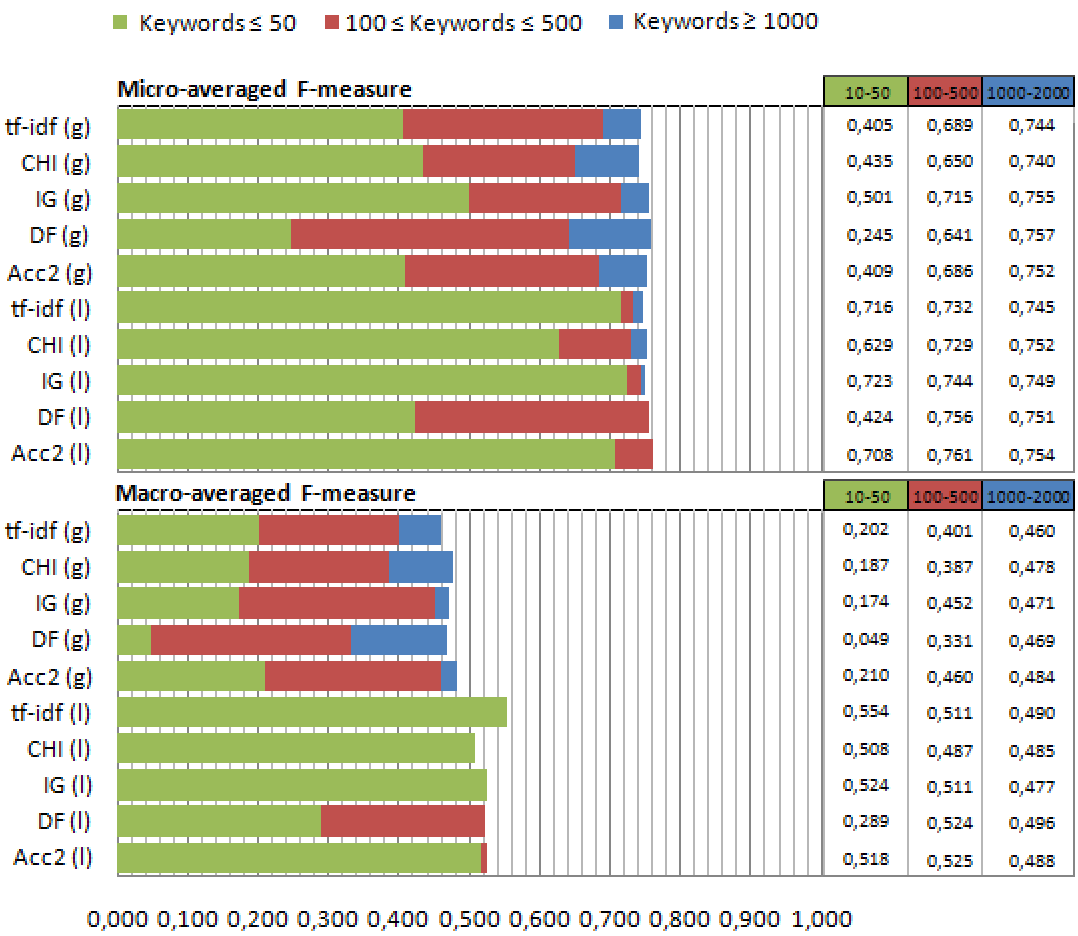


Figure 7.30. Comparison of the averages of micro- and macro-averaged F-measures for Wap dataset

Secondly, the local policy achieves consistently higher micro- and macro-averaged F-measure values than the global policy when we compare the results of the experiments in global policy with the experiments in local policy as seen in Figure 7.30. The superiority of the local policy is over the global policy is emphasized more clearly when we looked at the macro-averaged F-measure results. Especially, this gap is more apparent with a few number of keywords. The average of the difference between the two policies is about 35%

when the keyword number is low. In local policy the highest macro-averaged F-measure values always achieved with a few number of keywords. In Wap dataset, the highest micro-averaged F-measure (77.1%) is achieved by *DF* with 200 keywords in local policy and the highest macro-averaged F-measure (59.3%) is achieved by *tf-idf* with only 30 keywords in local policy. As a result, we can say that rare classes are classified successfully with a few number of keywords in local policy by finding more crucial class keywords, since the approach determines significant terms for each class for the classification. On the other hand the global policy performs worse than the local policy because most of the keywords are selected from the prevailing classes, thus it prevents rare classes to be represented fairly by the general keywords concerning all classes.

When the test documents are categorized without feature selection, the classifier only achieves 75.2% micro- and 45% macro-averaged F-measures. However, both the global and local policies improve this performance with feature selection. For instance, the global policy with 500 and more keywords achieves higher results than the all word approach in terms of micro-averaged F-measure and with 200 and more keywords achieves higher results than the all word approach in terms of macro-averaged F-measure. In addition, this range drops to 50 and 10 keywords in local policy with micro- and macro-averaged F-measure respectively.

In Wap dataset generally *IG*, *DF* and *Acc2* are better than other metrics in global policy. As seen in Figure 7.30, the classifier achieves better micro-averaged F-measure results as the number of keywords increases from 10 to 2000 whereas the macro-averaged F-measure reaches its peak value of 50.1% with 500 keywords. On the other hand *tf-idf*, *DF* and *Acc2* are better than other metrics in local policy and the performance of the classifier is higher with a few number of keywords.

7.3.1.3. Analysis of Score and Rank Combinations

In the previous section we performed the five well-known feature selection metrics on the Wap dataset. In this section we evaluate the performance of the score and rank combinations on the dataset. Tables 7.41 and 7.42 show the micro- and macro-averaged F-measure results of the combinations in global policy for the Wap dataset.

Before starting the evaluation in global policy, it should be remember that the performance of the local policy was significantly better than the performance of the global policy when we analyzed the performance of the existing metrics on the Wap dataset in the previous section. After testing the score and rank combinations, the difference between the global policy and local policy is still high.

One of the different results of the combinations in the Wap dataset is that there is not any consistent successful performance between the score and rank combinations. As seen in Tables 7.41 and 7.42, score combination is more successful when the number of keywords is low but it is not as successful as rank combination when the number of keywords is high. On the other hand, the performance of the score combination is better when we analyze the averages of the F-measures as seen in Figure 7.31. Moreover, the improvement of the performance with combination is more apparent when the number of keyword is low.

In global policy among the 10 possible combinations of the two feature selection metrics, *tf-idf & CHI, tf-idf & Acc2, CHI & IG, CHI & Acc2* and *IG & Acc2* score combinations and *CHI & IG, CHI & Acc2, IG & Acc2* and *DF & Acc2* rank combinations are more successful than other combinations.

In global policy *IG* achieves the best F-measure values with a few number of keywords in the Wap dataset. When we look at the related tables, we can see that the score combination of *tf-idf & CHI* is more successful than individual *IG.* In addition to this, *DF* has higher F-measure values with a high number of keywords and the score combination of *tf-idf & Acc2, CHI & IG, IG & DF, IG & Acc2, DF & Acc2* and the rank combination of *DF & Acc2* improves the performance of individual metric *DF*.

Among the individual metrics *DF* has the highest micro-averaged F-measure 75.8% and it achieves this success rate with 2000 keywords. This value is improved by some combinations and the highest result among them is 76.6% and it is achieved by the rank combination of *DF & Acc2* with 1000 keywords.

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,134	0,496	0,587	0,655	0,691	0,721	0,740	0,749	0,743
CHI (g)		0,242	0,523	0,540	0,607	0,631	0,712	0,730	0,741	0,749
IG (g)		0,399	0,526	0,577	0,644	0,746	0,753	0,755	0,756	0,755
DF (g)		0,000	0,341	0,395	0,543	0,657	0,723	0,756	0,757	0,758
Acc2 (g)		0,221	0,476	0,529	0,629	0,697	0,730	0,743	0,753	0,758
	MAX	0,399	0,526	0,587	0,655	0,746	0,753	0,756	0,757	0,758
Micro-F	Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (g)		0,485	0,585	0,604	0,667	0,720	0,753	0,745	0,743	0,750
|S| tf-idf&IG (g)		0,421	0,546	0,591	0,667	0,722	0,745	0,749	0,754	0,754
|S| tf-idf&DF (g)		0,045	0,414	0,482	0,595	0,680	0,726	0,746	0,755	0,748
|S| tf-idf&Acc2 (g)		0,452	0,523	0,604	0,666	0,698	0,740	0,759	0,758	0,751
|S| CHI&IG (g)		0,452	0,558	0,591	0,631	0,715	0,748	0,756	0,758	0,757
|S| CHI&DF (g)		0,021	0,398	0,543	0,606	0,672	0,737	0,742	0,754	0,753
|S| CHI&Acc2 (g)		0,443	0,481	0,565	0,616	0,702	0,722	0,753	0,755	0,756
|S| IG&DF (g)		0,370	0,383	0,534	0,621	0,716	0,752	0,755	0,764	0,758
|S| IG&Acc2 (g)		0,427	0,558	0,578	0,625	0,708	0,737	0,760	0,757	0,760
|S| DF&Acc2 (g)		0,313	0,414	0,519	0,598	0,684	0,735	0,744	0,759	0,761
	MAX	0,485	0,585	0,604	0,667	0,722	0,753	0,760	0,764	0,761
	AVERAGE	0,343	0,486	0,561	0,629	0,702	0,740	0,751	0,756	0,755
Micro-F	Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (g)		0,470	0,529	0,566	0,613	0,710	0,739	0,746	0,745	0,744
|R| tf-idf&IG (g)		0,431	0,551	0,576	0,654	0,716	0,742	0,756	0,749	0,754
|R| tf-idf&DF (g)		0,323	0,464	0,517	0,627	0,690	0,709	0,754	0,755	0,754
|R| tf-idf&Acc2 (g)		0,418	0,549	0,569	0,644	0,696	0,745	0,758	0,756	0,753
|R| CHI&IG (g)		0,433	0,556	0,576	0,631	0,692	0,752	0,751	0,751	0,749
|R| CHI&DF (g)		0,374	0,503	0,567	0,642	0,676	0,751	0,754	0,752	0,743
|R| CHI&Acc2 (g)		0,440	0,486	0,553	0,608	0,670	0,722	0,747	0,755	0,761
|R| IG&DF (g)		0,404	0,474	0,514	0,673	0,710	0,742	0,755	0,758	0,756
|R| IG&Acc2 (g)		0,407	0,550	0,576	0,648	0,705	0,737	0,755	0,763	0,756
|R| DF&Acc2 (g)		0,344	0,449	0,531	0,645	0,681	0,735	0,766	0,761	0,761
	MAX	0,470	0,556	0,576	0,673	0,716	0,752	0,766	0,763	0,761
	AVERAGE	0,404	0,511	0,555	0,639	0,695	0,737	0,754	0,754	0,753

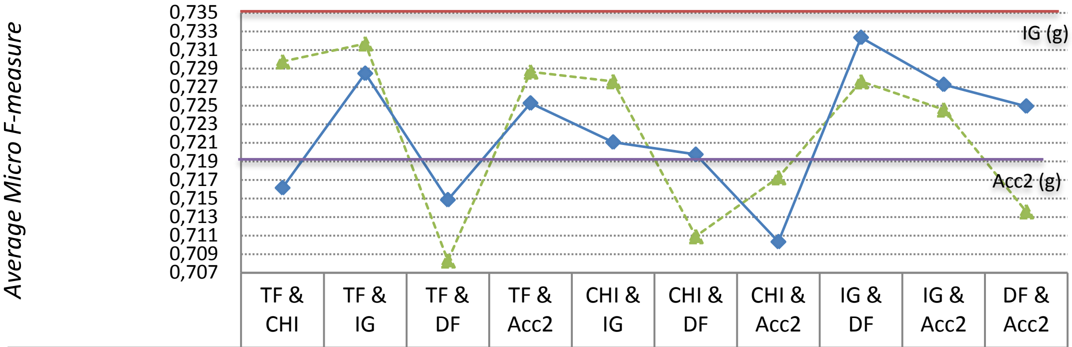
Table 7.42. In global policy, micro-averaged F-measures of the score and rank combinations for Wap dataset

On the other hand, among the individual metrics *IG* has the highest macro-averaged F-measure 50.1% and it achieves this success rate with 500 keywords. This value is improved to 51% by only the rank combination of *CHI & IG* with 500 keywords.

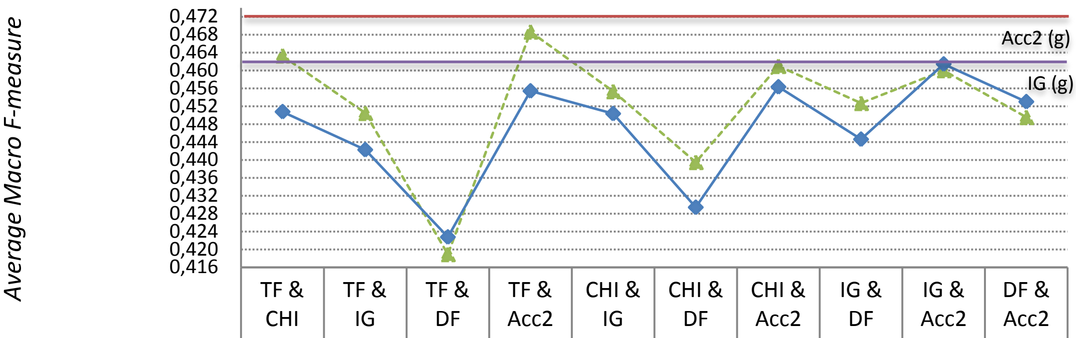
Macro-F	10	30	50	100	200	500	1000	1500	2000
tf-idf (g)	0,093	0,208	**0,306**	0,350	0,412	0,442	0,455	0,468	0,455
CHI (g)	**0,121**	0,185	0,256	0,336	0,375	0,451	**0,486**	0,469	0,478
IG (g)	0,052	0,185	0,284	0,375	**0,479**	**0,501**	0,473	0,474	0,467
DF (g)	0,000	0,053	0,095	0,237	0,326	0,430	0,474	0,470	0,462
Acc2 (g)	0,117	**0,235**	0,278	**0,411**	**0,479**	0,489	0,480	**0,480**	**0,492**
MAX	0,121	0,235	0,306	0,411	0,479	0,501	0,486	0,480	0,492
Macro-F Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (g)	**0,197**	**0,301**	**0,336**	0,408	0,469	0,494	0,466	0,463	0,480
|S| tf-idf&IG (g)	0,077	0,212	0,289	0,365	0,453	0,475	0,469	0,471	0,471
|S| tf-idf&DF (g)	0,014	0,106	0,164	0,307	0,380	0,438	0,459	0,468	0,461
|S| tf-idf&Acc2 (g)	**0,160**	**0,285**	**0,360**	**0,413**	0,465	0,495	**0,486**	**0,480**	0,472
|S| CHI&IG (g)	**0,163**	**0,262**	0,300	0,350	0,450	0,489	0,476	**0,483**	0,485
|S| CHI&DF (g)	0,002	0,139	0,265	0,318	0,398	0,491	0,465	**0,485**	0,480
|S| CHI&Acc2 (g)	**0,177**	**0,242**	0,283	0,341	**0,487**	0,484	0,484	**0,483**	0,487
|S| IG&DF (g)	0,064	0,066	0,179	0,349	0,452	0,498	0,468	0,478	0,472
|S| IG&Acc2 (g)	0,112	**0,284**	0,293	0,356	**0,479**	0,478	**0,487**	0,477	0,483
|S| DF&Acc2 (g)	0,076	0,140	0,262	0,317	0,450	0,490	0,476	0,479	0,486
MAX	**0,197**	**0,301**	**0,360**	**0,413**	**0,487**	0,498	**0,487**	**0,485**	0,487
AVERAGE	0,104	0,204	0,273	0,352	0,448	0,483	0,474	0,477	0,478
Macro-F Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (g)	**0,167**	0,226	0,268	0,323	0,469	0,479	0,478	0,477	0,479
|R| tf-idf&IG (g)	**0,137**	**0,258**	0,277	0,363	0,415	0,459	0,476	0,470	0,471
|R| tf-idf&DF (g)	0,084	0,147	0,190	0,330	0,386	0,413	0,466	0,471	0,471
|R| tf-idf&Acc2 (g)	**0,127**	**0,260**	0,290	0,366	0,440	0,493	**0,486**	0,475	0,472
|R| CHI&IG (g)	**0,149**	**0,268**	0,280	0,350	0,407	**0,510**	0,471	**0,484**	0,480
|R| CHI&DF (g)	0,067	0,156	0,238	0,335	0,398	0,398	**0,495**	0,475	0,474
|R| CHI&Acc2 (g)	**0,177**	**0,246**	0,268	0,345	0,437	0,492	**0,486**	**0,483**	**0,495**
|R| IG&DF (g)	0,064	0,148	0,172	0,363	0,420	0,476	0,467	0,470	0,471
|R| IG&Acc2 (g)	0,099	**0,296**	0,293	0,376	0,467	0,494	0,479	0,478	0,475
|R| DF&Acc2 (g)	0,059	0,148	0,213	0,373	0,417	0,488	0,481	**0,482**	0,477
MAX	**0,177**	**0,296**	0,293	0,376	0,469	**0,510**	**0,495**	**0,484**	**0,495**
AVERAGE	0,113	0,215	0,249	0,352	0,426	0,470	0,478	0,477	0,477

Table 7.43. In global policy, macro-averaged F-measures of the score and rank combinations for Wap dataset

In Figure 7.31, we can see that none of the combinations improve the highest average of the F-measures of the individual metric despite the improvement of the highest F-measures of the existing metrics. The main reason for this result is that the highest micro-averaged F-measure (76.6%) of the individual metric at 200 keywords is not improved by any combinations. Thus *IG* has the highest average of the F-measures in terms of micro-averaged F-measure. The similar situation is valid for macro-averaged F-measure. The highest macro-averaged F-measure (49.2%) of the individual metric at 2000 keywords is not improved by any combinations. Thus *Acc2* has the highest average of the F-measures in terms of micro-averaged F-measure.



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,730	0,732	0,708	0,729	0,728	0,711	0,717	0,728	0,725	0,714
Rank Com. (g)	0,716	0,728	0,715	0,725	0,721	0,720	0,710	0,732	0,727	0,725



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,463	0,450	0,419	0,469	0,455	0,440	0,461	0,453	0,460	0,450
Rank Com. (g)	0,451	0,442	0,423	0,455	0,450	0,429	0,456	0,445	0,461	0,453

	tf-idf (l)	CHI (l)	IG (l)	DF (l)	Acc2 (l)
Micro	0,716	0,695	**0,735**	0,699	0,719
Macro	0,430	0,433	0,462	0,400	**0,472**

Figure 7.31. In global policy, comparison of score and rank combinations on the Wap dataset

When we perform the score and rank combinations on the Wap dataset in local policy, the score combination is significantly better than the rank combination as seen in Figure 7.32. In the previous datasets, there is not such a clear gap between two combinations.

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (l)		0,671	**0,737**	0,741	0,738	0,735	0,722	0,746	0,741	0,749
CHI (l)		0,440	0,714	0,732	0,732	0,720	0,736	0,742	0,756	**0,758**
IG (l)		**0,685**	0,735	0,750	0,742	0,747	0,744	0,742	0,758	0,749
DF (l)		0,000	0,567	0,704	0,751	**0,771**	0,747	**0,760**	0,747	0,747
Acc2 (l)		0,639	0,728	**0,757**	**0,770**	0,758	**0,755**	0,752	**0,758**	0,752
	MAX	0,685	0,737	0,757	0,770	0,771	0,755	0,760	0,758	0,758
Micro-F	Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (l)		**0,692**	**0,740**	0,743	0,758	0,740	0,731	0,739	0,746	0,753
|S| tf-idf&IG (l)		**0,698**	**0,739**	0,738	0,754	0,742	0,732	0,746	0,747	**0,759**
|S| tf-idf&DF (l)		0,623	0,689	0,738	0,764	0,760	0,735	0,737	0,745	0,747
|S| tf-idf&Acc2 (l)		0,671	**0,754**	**0,757**	0,753	0,747	0,735	0,746	0,747	0,756
|S| CHI&IG (l)		**0,711**	0,726	0,735	0,735	0,729	0,742	0,753	**0,763**	0,755
|S| CHI&DF (l)		0,233	0,682	0,743	0,746	0,742	0,738	0,734	0,747	0,754
|S| CHI&Acc2 (l)		0,683	0,736	0,755	0,759	0,744	0,745	0,751	**0,758**	0,753
|S| IG&DF (l)		0,605	0,675	0,737	0,748	0,753	0,743	0,748	0,751	0,755
|S| IG&Acc2 (l)		0,676	0,736	0,752	0,760	0,749	**0,759**	0,754	0,753	0,755
|S| DF&Acc2 (l)		0,606	0,647	0,739	0,739	**0,773**	**0,758**	0,747	0,754	0,748
	MAX	**0,711**	**0,754**	**0,757**	0,764	**0,773**	**0,759**	0,754	**0,763**	**0,759**
	AVERAGE	0,620	0,712	0,744	0,752	0,748	0,742	0,745	0,751	0,753
Micro-F	Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (l)		**0,686**	0,732	0,720	0,747	0,737	0,729	0,733	0,754	**0,762**
|R| tf-idf&IG (l)		0,680	0,728	0,730	0,734	0,740	0,731	0,740	0,753	0,757
|R| tf-idf&DF (l)		0,650	0,734	0,743	0,762	0,745	0,739	0,740	0,742	0,749
|R| tf-idf&Acc2 (l)		0,683	**0,743**	0,751	0,738	0,742	0,741	0,738	0,751	0,752
|R| CHI&IG (l)		**0,700**	0,727	0,717	0,727	0,732	0,739	0,742	**0,763**	**0,758**
|R| CHI&DF (l)		0,669	**0,742**	0,733	0,734	0,741	0,731	0,737	0,749	0,752
|R| CHI&Acc2 (l)		**0,693**	0,725	0,737	0,748	0,737	0,739	0,747	0,751	0,752
|R| IG&DF (l)		0,650	**0,737**	0,736	0,744	0,736	0,733	0,747	0,749	0,749
|R| IG&Acc2 (l)		0,661	**0,740**	0,738	0,749	0,751	0,742	0,754	0,753	0,749
|R| DF&Acc2 (l)		0,627	0,730	0,733	0,758	0,750	**0,758**	0,750	0,744	0,749
	MAX	**0,700**	**0,743**	0,751	0,762	0,751	**0,758**	0,754	**0,763**	**0,762**
	AVERAGE	0,670	0,734	0,734	0,744	0,741	0,738	0,743	0,751	0,753

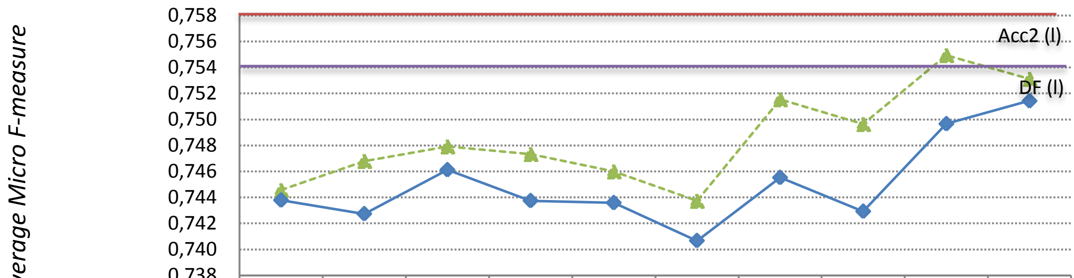
Table 7.44. In local policy, micro-averaged F-measures of the score and rank combinations for Wap dataset

In addition, the local policy achieves consistently higher performance than the global policy. Figure 7.33 demonstrates this difference clearly. Again in the previous datasets, there is not such a high gap between the score combination and the rank combination.

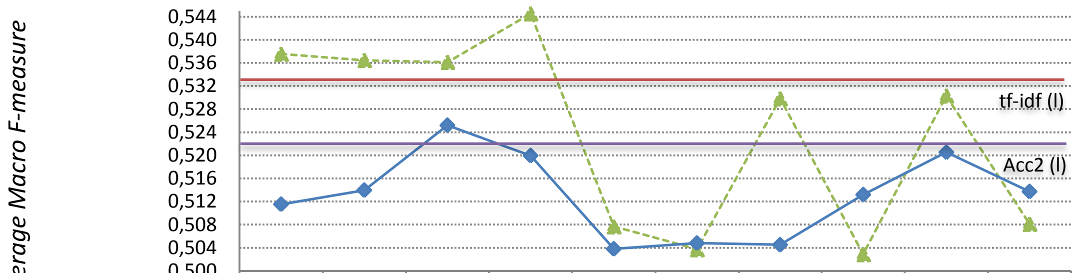
Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (l)		**0,506**	**0,593**	**0,565**	0,532	0,507	0,495	0,509	0,477	0,483
CHI (l)		0,493	0,511	0,520	0,509	0,462	0,491	0,475	0,488	**0,491**
IG (l)		0,492	0,531	0,548	0,517	0,508	0,508	0,460	0,490	0,482
DF (l)		0,000	0,353	0,513	0,550	**0,538**	0,483	**0,524**	0,483	0,481
Acc2 (l)		0,435	0,554	0,564	**0,551**	0,509	**0,516**	0,488	**0,491**	0,485
	MAX	0,506	0,593	0,565	0,551	0,538	0,516	0,524	0,491	0,491
Macro-F	Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (l)		0,503	0,597	**0,568**	**0,562**	0,502	0,494	0,478	0,479	0,485
|S| tf-idf&IG (l)		0,505	0,584	**0,578**	**0,563**	0,498	0,492	0,484	0,478	0,487
|S| tf-idf&DF (l)		0,447	0,531	**0,565**	0,603	0,567	0,503	0,500	0,480	0,482
|S| tf-idf&Acc2 (l)		0,489	0,571	0,614	**0,572**	0,529	0,492	0,483	0,480	0,485
|S| CHI&IG (l)		0,515	0,519	0,514	0,511	0,480	0,507	0,472	0,494	0,486
|S| CHI&DF (l)		0,462	0,487	0,562	0,518	0,499	0,494	0,469	0,480	0,484
|S| CHI&Acc2 (l)		0,505	0,562	**0,569**	0,543	0,492	0,508	0,485	0,492	0,488
|S| IG&DF (l)		0,399	0,494	**0,567**	**0,551**	0,504	0,502	0,483	0,483	0,483
|S| IG&Acc2 (l)		0,481	0,558	0,555	**0,556**	0,510	0,522	0,492	0,488	0,489
|S| DF&Acc2 (l)		0,385	0,481	**0,584**	0,542	**0,538**	**0,520**	0,486	0,484	0,479
	MAX	**0,515**	**0,597**	**0,614**	**0,603**	**0,567**	**0,522**	0,500	**0,494**	0,489
	AVERAGE	0,469	0,538	0,567	0,552	0,512	0,503	0,483	0,484	0,485
Macro-F	Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (l)		0,517	0,519	0,491	0,542	0,508	0,492	0,474	0,489	0,495
|R| tf-idf&IG (l)		0,497	0,543	0,521	0,515	0,509	0,498	0,474	0,488	**0,493**
|R| tf-idf&DF (l)		0,456	0,556	0,545	0,576	0,521	0,497	0,502	0,480	0,484
|R| tf-idf&Acc2 (l)		0,490	0,573	0,540	0,509	0,509	0,498	0,478	0,488	0,485
|R| CHI&IG (l)		**0,510**	0,523	0,492	0,514	0,475	0,509	0,467	0,493	0,489
|R| CHI&DF (l)		**0,508**	0,539	0,505	0,503	0,483	0,492	0,480	0,486	0,487
|R| CHI&Acc2 (l)		**0,507**	0,516	0,509	0,511	0,490	0,493	0,481	0,487	0,488
|R| IG&DF (l)		0,466	0,561	0,527	0,518	0,509	0,498	0,489	0,484	0,485
|R| IG&Acc2 (l)		0,464	0,562	0,530	0,535	0,532	0,500	0,485	0,487	0,485
|R| DF&Acc2 (l)		0,423	0,549	0,554	0,534	0,512	0,512	0,491	0,485	0,482
	MAX	**0,517**	0,573	0,554	**0,576**	0,532	0,512	0,502	**0,493**	**0,495**
	AVERAGE	0,484	0,544	0,521	0,526	0,505	0,499	0,482	0,487	0,487

Table 7.45. In local policy, macro-averaged F-measures of the score and rank combinations for Wap dataset

In local policy among the 10 possible combinations of the two feature selection metrics, *tf-idf & CHI, tf-idf & DF, tf-idf & Acc2, IG & Acc2* and *DF & Acc2* score combinations are more successful than other combinations. It should be noted that the performance of the combinations in terms of the micro-averaged F-measure is not as successful as the performance of the combinations in terms of the macro-averaged F-measure.



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (l)	0,745	0,747	0,748	0,747	0,746	0,744	0,752	0,750	0,755	0,753
Rank Com. (l)	0,744	0,743	0,746	0,744	0,744	0,741	0,746	0,743	0,750	0,751



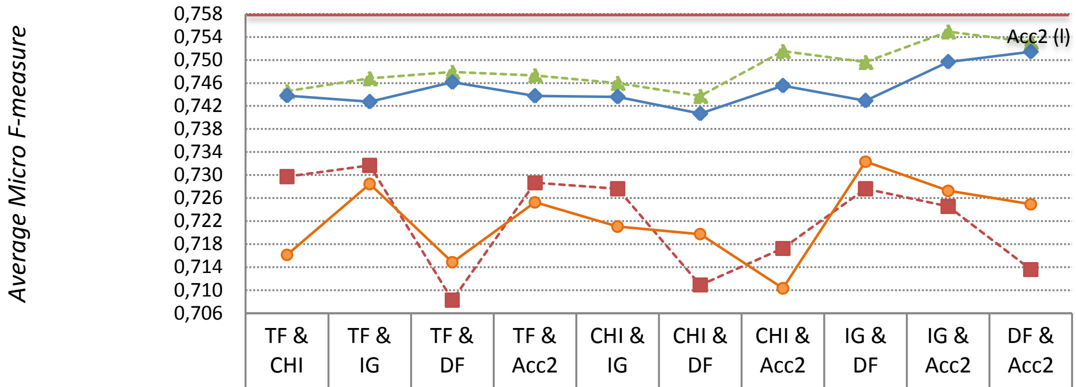
	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (l)	0,538	0,536	0,536	0,545	0,508	0,504	0,530	0,503	0,530	0,508
Rank Com. (l)	0,512	0,514	0,525	0,520	0,504	0,505	0,505	0,513	0,521	0,514

	tf-idf (l)	CHI (l)	IG (l)	DF (l)	Acc2 (l)
Micro	0,739	0,741	0,747	0,754	0,758
Macro	0,533	0,497	0,517	0,406	0,522

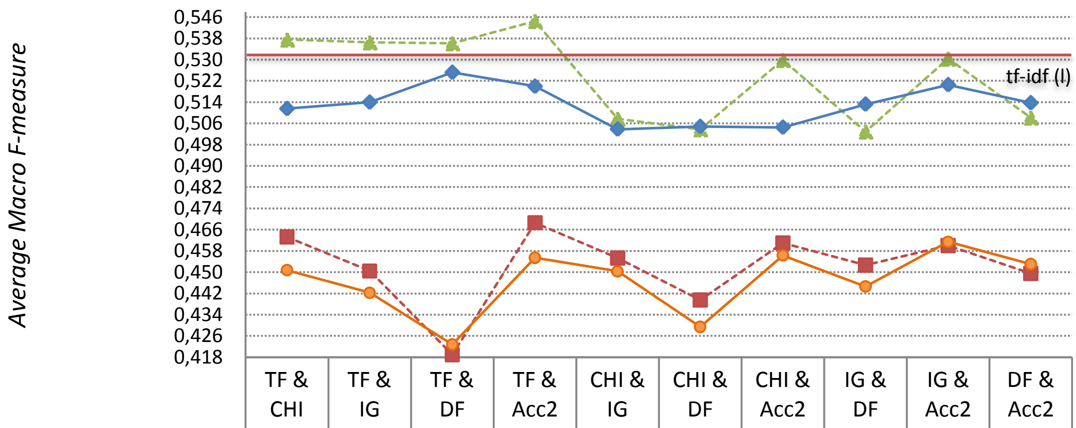
Figure 7.32. In local policy, comparison of score and rank combinations on Wap dataset

In local policy, among the individual feature selection metrics *DF* has the highest micro-averaged F-measure value (77.1%) with 200 keywords. When we look at the Table 7.44, we can see that the only score combination of *DF & Acc2* improves the highest value (77.3%).

Furthermore, the highest macro-averaged F-measure (59.3%) is achieved by *tf-idf* with only 30 keywords in local policy. This performance is improved by three score combinations *tf-idf & CHI, tf-idf & DF* and *tf-idf & Acc2* and the highest macro-averaged F-measure among them (61.4%) is achieved by *tf-idf & Acc2* with 50 keywords.



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,730	0,732	0,708	0,729	0,728	0,711	0,717	0,728	0,725	0,714
Rank Com. (g)	0,716	0,728	0,715	0,725	0,721	0,720	0,710	0,732	0,727	0,725
Score Com. (l)	0,745	0,747	0,748	0,747	0,746	0,744	0,752	0,750	0,755	0,753
Rank Com. (l)	0,744	0,743	0,746	0,744	0,744	0,741	0,746	0,743	0,750	0,751



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,463	0,450	0,419	0,469	0,455	0,440	0,461	0,453	0,460	0,450
Rank Com. (g)	0,451	0,442	0,423	0,455	0,450	0,429	0,456	0,445	0,461	0,453
Score Com. (l)	0,538	0,536	0,536	0,545	0,508	0,504	0,530	0,503	0,530	0,508
Rank Com. (l)	0,512	0,514	0,525	0,520	0,504	0,505	0,505	0,513	0,521	0,514

Figure 7.33. Comparison of score and rank combinations in global and local policy on the Wap dataset

7.3.1.4.Analysis of the Proposed Combinations

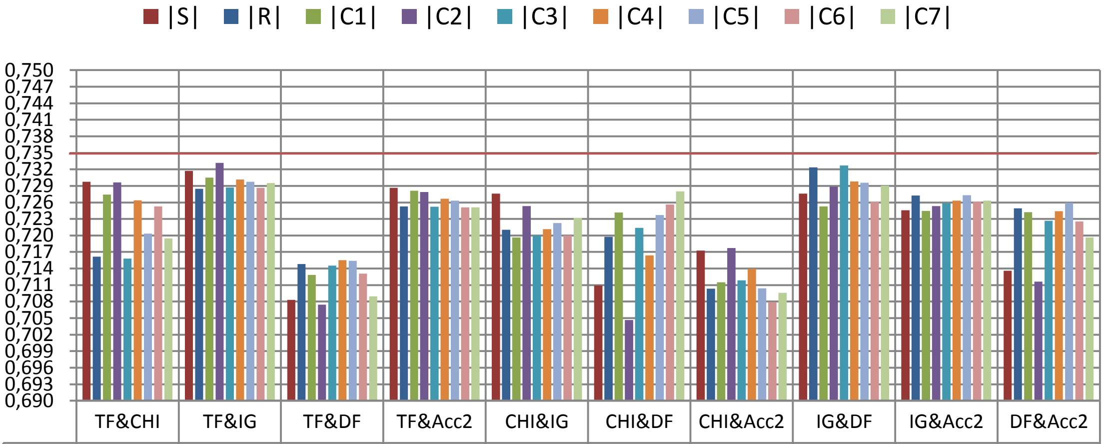
In this section, we analyze and discuss the performance of the proposed methods on the Wap dataset. Firstly, we examine the performance of the proposed methods in global policy. Then, we evaluate the performance of the proposed methods in local policy. Although we firstly share the results of the experiments in global policy, it should not forget that the performance of the local policy is significantly and constantly better than the performance of the global policy in Wap dataset.

In order to compare the proposed methods with the existing methods, we review what we learnt from the previous sections. In global policy *IG* achieved the best F-measure values with a few number of keywords and *DF* had the highest F-measure values with a high number of keywords in the Wap dataset.

Furthermore, we knew, from the previous sections, that the score combination was more successful than the rank combination when the number of keywords is low but it was not as successful as rank combination when the number of keywords is high in global policy in Wap dataset. The score combination of *tf-idf & CHI* and *tf-idf & IG* were more successful than individual *IG* with a few number of keywords. The score combination of *tf-idf & Acc2, CHI & IG, IG & DF, IG & Acc2, DF & Acc2* and the rank combination of *DF & Acc2* and *IG & Acc2* improved the performance of individual metrics *IG* and *DF.* As a result, it showed that combining individual metrics gave more accurate results in Wap dataset.

Tables 7.48 and 7.49 show the micro- and macro-averaged F-measures, respectively, for all seven proposed methods in global policy for the Wap dataset. The first noticeable thing, in these tables, almost all combinations (*tf-idf & CHI, tf-idf & IG, tf-idf & Acc2, CHI & IG, IG & DF, IG & Acc2, DF & Acc2*) in proposed methods are more successful than the individual metrics in terms of micro-averaged F-measure. Although many proposed combinations cannot improve the highest macro-averaged F-measure of the individual metrics, *tf-idf & CHI, tf-idf & Acc2, CHI & DF, CHI & Acc2, CHI & IG* and *IG & Acc2* combinations get successful results with different number of keywords. In addition to, *tf-idf & DF* is the worst combination in dataset.

Among the proposed methods C2 significantly improves the success of the score combination with a few number of keywords and among the proposed methods C5, C6, C7 significantly improves the rank combination with a high number of keywords in terms of micro-averaged F-measure but the rank combination is still has the highest macro-averaged F-measures with a high number of keywords.

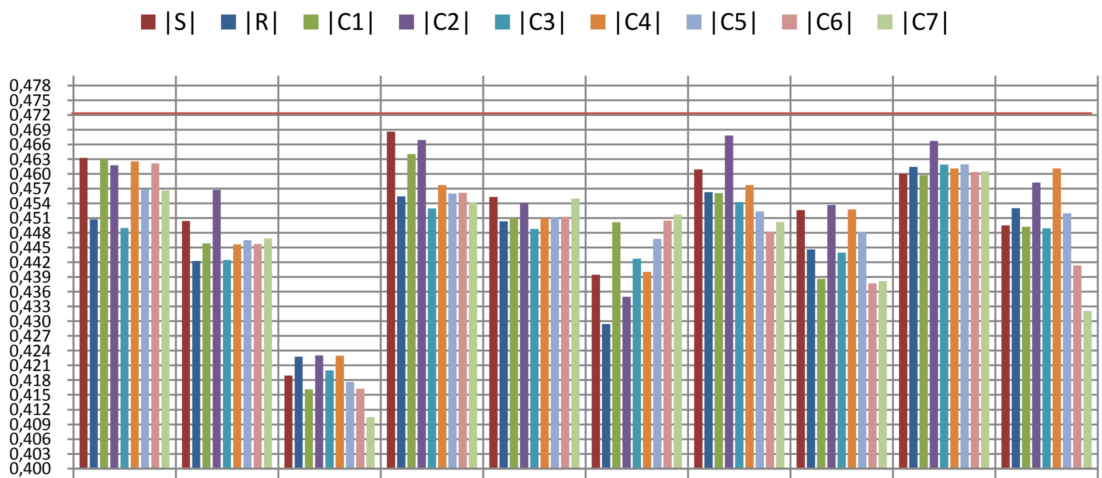


	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,730	0,732	0,708	0,729	0,728	0,711	0,717	0,728	0,725	0,714
|R|	0,716	0,728	0,715	0,725	0,721	0,720	0,710	0,732	0,727	0,725
|C1|	0,727	0,730	0,713	0,728	0,720	0,724	0,712	0,725	0,724	0,724
|C2|	0,730	0,733	0,707	0,728	0,725	0,705	0,718	0,729	0,725	0,712
|C3|	0,716	0,729	0,715	0,725	0,720	0,721	0,712	0,733	0,726	0,723
|C4|	0,726	0,730	0,716	0,727	0,721	0,716	0,714	0,730	0,726	0,724
|C5|	0,720	0,730	0,715	0,726	0,722	0,724	0,710	0,730	0,727	0,726
|C6|	0,725	0,729	0,713	0,725	0,720	0,726	0,708	0,726	0,726	0,723
|C7|	0,719	0,730	0,709	0,725	0,723	0,728	0,710	0,729	0,726	0,720

Figure 7.34. In global policy, averages of the micro-averaged F-measures of all combinations for Wap dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	**0,485**	**0,585**	**0,604**	0,667	**0,722**	**0,753**	0,760	**0,764**	**0,761**
Combination | R |	0,470	0,556	0,576	**0,673**	0,716	0,752	0,766	0,763	**0,761**
Combination |C1|	0,463	0,568	0,592	0,662	0,725	0,752	0,763	0,763	0,763
Combination |C2|	0,456	0,580	0,621	0,681	0,733	0,749	0,754	0,762	0,762
Combination |C3|	0,450	0,552	0,578	0,673	0,717	0,754	0,761	0,761	0,761
Combination |C4|	0,463	0,566	0,606	0,658	0,720	0,748	0,764	0,761	0,763
Combination |C5|	0,450	0,551	0,584	0,662	0,717	0,757	0,768	0,764	0,760
Combination |C6|	0,459	0,570	0,587	0,652	0,723	0,755	0,766	0,764	0,761
Combination |C7|	0,450	0,563	0,587	0,662	0,723	0,764	0,766	0,761	0,760

Table 7.46. In global policy, maximum micro-averaged F-measures of all combinations for Wap dataset



	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,463	0,450	0,419	0,469	0,455	0,440	0,461	0,453	0,460	0,450
|R|	0,451	0,442	0,423	0,455	0,450	0,429	0,456	0,445	0,461	0,453
|C1|	0,463	0,446	0,416	0,464	0,451	0,450	0,456	0,439	0,460	0,449
|C2|	0,462	0,457	0,423	0,467	0,454	0,435	0,468	0,454	0,467	0,458
|C3|	0,449	0,442	0,420	0,453	0,449	0,443	0,454	0,444	0,462	0,449
|C4|	0,463	0,446	0,423	0,458	0,451	0,440	0,458	0,453	0,461	0,461
|C5|	0,457	0,447	0,418	0,456	0,451	0,447	0,452	0,448	0,462	0,452
|C6|	0,462	0,446	0,416	0,456	0,451	0,450	0,448	0,438	0,460	0,441
|C7|	0,457	0,447	0,410	0,454	0,455	0,452	0,450	0,438	0,461	0,432

Figure 7.35. In global policy, averages of the macro-averaged F-measures of all combinations for Wap dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	**0,197**	**0,301**	**0,360**	**0,413**	**0,487**	0,498	0,487	**0,485**	0,487
Combination | R |	0,177	0,296	0,293	0,376	0,469	**0,510**	**0,495**	0,484	**0,495**
Combination |C1|	0,167	0,287	0,318	0,379	**0,493**	0,500	0,490	**0,488**	0,491
Combination |C2|	**0,198**	**0,330**	**0,369**	**0,415**	0,483	0,498	0,490	**0,486**	0,491
Combination |C3|	**0,177**	**0,296**	0,295	0,376	0,465	**0,510**	0,486	0,484	**0,495**
Combination |C4|	0,167	0,273	**0,324**	**0,400**	0,482	0,488	**0,492**	**0,486**	0,486
Combination |C5|	0,162	0,265	0,317	0,375	**0,493**	0,500	**0,493**	0,485	0,492
Combination |C6|	0,167	0,278	0,314	0,365	**0,492**	0,500	**0,492**	**0,488**	0,488
Combination |C7|	**0,177**	0,289	0,317	0,371	**0,493**	**0,505**	0,491	0,483	**0,494**

Table 7.47. In global policy, maximum macro-averaged F-measures of all combinations for Wap dataset

As we knew that among the individual metrics the best micro-averaged 75.8% F-measure was achieved by *DF* with 2000 keywords. This performance was improved by several score and rank combinations and among them the most successful result 76.6% was achieved by the rank combination of *DF & Acc2*. Furthermore this performance is also improved, 76.8%, by C5 of *tf-idf & Acc2* with only 1000 keywords.

Table 7.48. Micro F-measure results of the proposed combinations in global policy for Wap dataset

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,134	0,496	0,587	0,655	0,691	0,721	0,740	0,749	0,743
CHI (g)		0,242	0,523	0,540	0,607	0,631	0,712	0,730	0,741	0,749
IG (g)		0,399	0,526	0,577	0,644	0,746	0,753	0,755	0,756	0,755
DF (g)		0,000	0,341	0,395	0,543	0,657	0,723	0,756	0,757	0,758
Acc2 (g)		0,221	0,476	0,529	0,629	0,697	0,730	0,743	0,753	0,758
	MAX	0,399	0,526	0,587	0,655	0,746	0,753	0,756	0,757	0,758
Micro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	0,485	0,585	0,604	0,667	0,722	0,753	0,760	0,764	0,761
	AVERAGE	0,343	0,486	0,561	0,629	0,702	0,740	0,751	0,756	0,755
Rank Combination	MAX	0,470	0,556	0,576	0,673	0,716	0,752	0,766	0,763	0,761
	AVERAGE	0,404	0,511	0,555	0,639	0,695	0,737	0,754	0,754	0,753
Micro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (g)		0,463	0,568	0,586	0,639	0,725	0,752	0,760	0,743	0,746
|C1| tf-idf&IG (g)		0,416	0,527	0,565	0,662	0,719	0,746	0,751	0,753	0,753
|C1| tf-idf&DF (g)		0,410	0,452	0,503	0,616	0,692	0,709	0,753	0,758	0,750
|C1| tf-idf&Acc2 (g)		0,436	0,543	0,592	0,651	0,700	0,746	0,763	0,759	0,750
|C1| CHI&IG (g)		0,433	0,560	0,574	0,629	0,687	0,749	0,748	0,753	0,753
|C1| CHI&DF (g)		0,429	0,515	0,579	0,619	0,706	0,750	0,760	0,756	0,754
|C1| CHI&Acc2 (g)		0,407	0,521	0,553	0,627	0,679	0,713	0,743	0,748	0,760
|C1| IG&DF (g)		0,370	0,504	0,551	0,640	0,707	0,738	0,754	0,758	0,754
|C1| IG&Acc2 (g)		0,426	0,540	0,587	0,629	0,705	0,733	0,759	0,760	0,760
|C1| DF&Acc2 (g)		0,358	0,444	0,533	0,645	0,691	0,726	0,758	0,763	0,763
	MAX	0,463	0,568	0,592	0,662	0,725	0,752	0,763	0,763	0,763
	AVERAGE	0,415	0,517	0,562	0,636	0,701	0,736	0,755	0,755	0,754
Micro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (g)		0,451	0,580	0,621	0,672	0,710	0,749	0,752	0,745	0,750
|C2| tf-idf&IG (g)		0,379	0,541	0,607	0,681	0,720	0,747	0,751	0,749	0,751
|C2| tf-idf&DF (g)		0,045	0,125	0,488	0,593	0,685	0,723	0,743	0,752	0,748
|C2| tf-idf&Acc2 (g)		0,191	0,525	0,610	0,661	0,699	0,737	0,753	0,761	0,756
|C2| CHI&IG (g)		0,456	0,539	0,592	0,616	0,721	0,745	0,753	0,761	0,756
|C2| CHI&DF (g)		0,021	0,329	0,525	0,591	0,650	0,735	0,750	0,754	0,748
|C2| CHI&Acc2 (g)		0,147	0,489	0,565	0,635	0,696	0,718	0,742	0,756	0,760
|C2| IG&DF (g)		0,322	0,375	0,524	0,621	0,733	0,749	0,746	0,762	0,761
|C2| IG&Acc2 (g)		0,427	0,541	0,571	0,631	0,706	0,742	0,754	0,757	0,762
|C2| DF&Acc2 (g)		0,021	0,147	0,466	0,595	0,681	0,733	0,749	0,756	0,755
	MAX	0,456	0,580	0,621	0,681	0,733	0,749	0,754	0,762	0,762
	AVERAGE	0,246	0,419	0,557	0,630	0,700	0,738	0,749	0,755	0,755
Micro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (g)		0,433	0,529	0,574	0,614	0,712	0,737	0,746	0,742	0,744
|C3| tf-idf&IG (g)		0,431	0,552	0,578	0,656	0,717	0,742	0,756	0,749	0,753
|C3| tf-idf&DF (g)		0,338	0,448	0,511	0,621	0,697	0,706	0,756	0,755	0,754
|C3| tf-idf&Acc2 (g)		0,450	0,549	0,569	0,646	0,698	0,742	0,758	0,756	0,753
|C3| CHI&IG (g)		0,433	0,543	0,576	0,631	0,688	0,752	0,751	0,752	0,746
|C3| CHI&DF (g)		0,374	0,503	0,567	0,642	0,679	0,754	0,756	0,754	0,745
|C3| CHI&Acc2 (g)		0,440	0,486	0,552	0,623	0,663	0,726	0,744	0,755	0,761
|C3| IG&DF (g)		0,404	0,468	0,514	0,673	0,710	0,747	0,754	0,758	0,754
|C3| IG&Acc2 (g)		0,407	0,550	0,577	0,645	0,703	0,733	0,756	0,761	0,757
|C3| DF&Acc2 (g)		0,402	0,442	0,486	0,645	0,672	0,738	0,761	0,761	0,758
	MAX	0,450	0,552	0,578	0,673	0,717	0,754	0,761	0,761	0,761
	AVERAGE	0,411	0,507	0,550	0,640	0,694	0,738	0,754	0,754	0,752

Micro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (g)		0,463	0,532	0,576	0,653	0,718	0,742	0,751	0,745	0,750
|C4| tf-idf&IG (g)		0,431	0,541	0,565	0,658	0,719	0,748	0,751	0,750	0,756
|C4| tf-idf&DF (g)		0,412	0,458	0,509	0,629	0,685	0,715	0,756	0,753	0,756
|C4| tf-idf&Acc2 (g)		0,450	0,535	0,594	0,651	0,695	0,740	0,764	0,759	0,751
|C4| CHI&IG (g)		0,433	0,566	0,587	0,630	0,692	0,746	0,753	0,754	0,751
|C4| CHI&DF (g)		0,410	0,512	0,606	0,618	0,687	0,735	0,751	0,754	0,753
|C4| CHI&Acc2 (g)		0,407	0,519	0,561	0,620	0,690	0,713	0,743	0,760	0,757
|C4| IG&DF (g)		0,383	0,492	0,551	0,645	0,720	0,742	0,757	0,759	0,757
|C4| IG&Acc2 (g)		0,410	0,547	0,587	0,632	0,710	0,735	0,757	0,761	0,763
|C4| DF&Acc2 (g)		0,383	0,436	0,543	0,650	0,690	0,733	0,758	0,760	0,755
	MAX	0,463	0,566	0,606	0,658	0,720	0,748	0,764	0,761	0,763
	AVERAGE	0,418	0,514	0,568	0,639	0,701	0,735	0,754	0,756	0,755

Micro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (g)		0,433	0,523	0,561	0,616	0,717	0,739	0,757	0,746	0,746
|C5| tf-idf&IG (g)		0,431	0,541	0,567	0,657	0,712	0,751	0,753	0,750	0,756
|C5| tf-idf&DF (g)		0,412	0,451	0,501	0,625	0,694	0,709	0,760	0,750	0,754
|C5| tf-idf&Acc2 (g)		0,450	0,543	0,584	0,649	0,692	0,744	0,768	0,754	0,750
|C5| CHI&IG (g)		0,433	0,543	0,577	0,629	0,695	0,751	0,748	0,753	0,757
|C5| CHI&DF (g)		0,374	0,501	0,569	0,628	0,707	0,757	0,756	0,749	0,745
|C5| CHI&Acc2 (g)		0,435	0,519	0,545	0,627	0,663	0,721	0,746	0,750	0,756
|C5| IG&DF (g)		0,372	0,487	0,543	0,662	0,699	0,745	0,756	0,761	0,755
|C5| IG&Acc2 (g)		0,407	0,551	0,571	0,642	0,707	0,735	0,756	0,764	0,760
|C5| DF&Acc2 (g)		0,383	0,426	0,508	0,645	0,692	0,735	0,763	0,763	0,757
	MAX	0,450	0,551	0,584	0,662	0,717	0,757	0,768	0,764	0,760
	AVERAGE	0,413	0,509	0,553	0,638	0,698	0,739	0,756	0,754	0,754

Micro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (g)		0,459	0,518	0,571	0,632	0,723	0,744	0,759	0,747	0,747
|C6| tf-idf&IG (g)		0,431	0,541	0,565	0,652	0,712	0,749	0,753	0,750	0,755
|C6| tf-idf&DF (g)		0,412	0,448	0,503	0,619	0,693	0,709	0,751	0,758	0,748
|C6| tf-idf&Acc2 (g)		0,421	0,545	0,586	0,644	0,686	0,747	0,766	0,759	0,750
|C6| CHI&IG (g)		0,433	0,570	0,571	0,627	0,692	0,750	0,749	0,749	0,754
|C6| CHI&DF (g)		0,374	0,506	0,575	0,617	0,716	0,755	0,758	0,753	0,754
|C6| CHI&Acc2 (g)		0,407	0,521	0,553	0,629	0,656	0,710	0,743	0,754	0,756
|C6| IG&DF (g)		0,368	0,487	0,541	0,635	0,716	0,737	0,755	0,758	0,754
|C6| IG&Acc2 (g)		0,410	0,550	0,587	0,632	0,712	0,735	0,754	0,764	0,760
|C6| DF&Acc2 (g)		0,383	0,432	0,496	0,634	0,686	0,732	0,759	0,764	0,761
	MAX	0,459	0,570	0,587	0,652	0,723	0,755	0,766	0,764	0,761
	AVERAGE	0,410	0,512	0,555	0,632	0,699	0,737	0,755	0,756	0,754

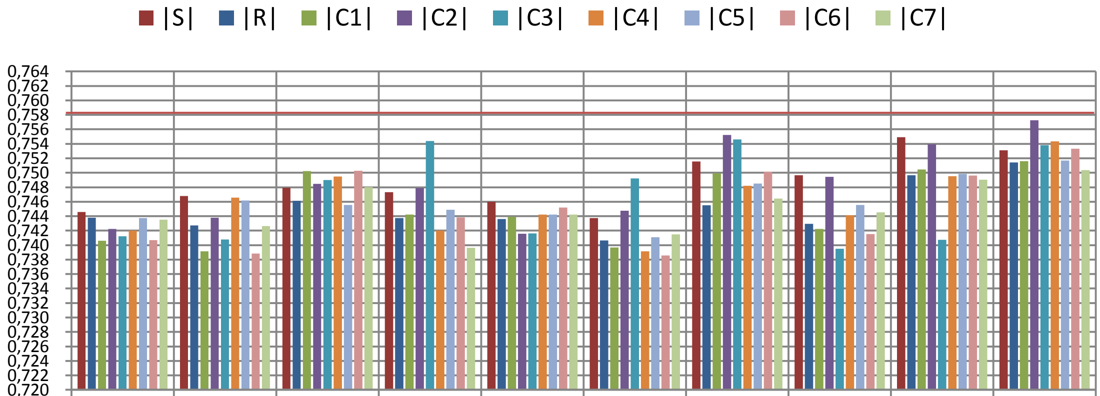
Micro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (g)		0,433	0,523	0,560	0,613	0,723	0,736	0,749	0,751	0,745
|C7| tf-idf&IG (g)		0,431	0,541	0,567	0,657	0,713	0,750	0,754	0,750	0,753
|C7| tf-idf&DF (g)		0,412	0,451	0,504	0,597	0,692	0,701	0,753	0,761	0,750
|C7| tf-idf&Acc2 (g)		0,450	0,543	0,587	0,649	0,691	0,738	0,766	0,756	0,750
|C7| CHI&IG (g)		0,433	0,542	0,577	0,642	0,698	0,746	0,746	0,750	0,756
|C7| CHI&DF (g)		0,374	0,501	0,557	0,646	0,708	0,764	0,752	0,750	0,748
|C7| CHI&Acc2 (g)		0,440	0,514	0,556	0,637	0,661	0,717	0,739	0,747	0,756
|C7| IG&DF (g)		0,372	0,474	0,509	0,662	0,711	0,732	0,756	0,756	0,758
|C7| IG&Acc2 (g)		0,407	0,563	0,571	0,641	0,713	0,735	0,753	0,761	0,756
|C7| DF&Acc2 (g)		0,402	0,432	0,497	0,624	0,684	0,732	0,758	0,761	0,760
	MAX	0,450	0,563	0,587	0,662	0,723	0,764	0,766	0,761	0,760
	AVERAGE	0,415	0,508	0,548	0,637	0,699	0,735	0,753	0,754	0,753

Table 7.49. Macro F-measure results of the proposed combinations in global policy for Wap dataset

Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,093	0,208	**0,306**	0,350	0,412	0,442	0,455	0,468	0,455
CHI (g)		**0,121**	0,185	0,256	0,336	0,375	0,451	**0,486**	0,469	0,478
IG (g)		0,052	0,185	0,284	0,375	**0,479**	**0,501**	0,473	0,474	0,467
DF (g)		0,000	0,053	0,095	0,237	0,326	0,430	0,474	0,470	0,462
Acc2 (g)		0,117	**0,235**	0,278	**0,411**	**0,479**	0,489	0,480	**0,480**	**0,492**
	MAX	0,121	0,235	0,306	0,411	0,479	0,501	0,486	0,480	0,492
Macro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	**0,197**	**0,301**	**0,360**	**0,413**	**0,487**	0,498	**0,487**	**0,485**	0,487
	AVERAGE	0,104	0,204	0,273	0,352	0,448	0,483	0,474	0,477	0,478
Rank Combination	MAX	**0,177**	**0,296**	0,293	0,376	0,469	**0,510**	**0,495**	**0,484**	**0,495**
	AVERAGE	0,113	0,215	0,249	0,352	0,426	0,470	0,478	0,477	0,477
Macro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (g)		0,146	0,287	0,308	0,363	0,493	0,491	0,485	0,471	0,475
|C1| tf-idf&IG (g)		0,076	0,202	0,264	0,365	0,420	0,479	0,470	0,471	0,470
|C1| tf-idf&DF (g)		0,098	0,116	0,190	0,293	0,387	0,414	0,467	0,474	0,462
|C1| tf-idf&Acc2 (g)		0,141	0,263	0,318	0,379	0,468	0,490	0,490	0,488	0,470
|C1| CHI&IG (g)		0,149	0,264	0,280	0,350	0,430	0,500	0,466	0,481	0,480
|C1| CHI&DF (g)		0,113	0,163	0,282	0,330	0,457	0,493	0,479	0,473	0,469
|C1| CHI&Acc2 (g)		0,167	0,252	0,261	0,356	0,458	0,474	0,482	0,475	0,491
|C1| IG&DF (g)		0,064	0,151	0,202	0,347	0,420	0,466	0,467	0,462	0,469
|C1| IG&Acc2 (g)		0,112	0,250	0,293	0,357	0,474	0,481	0,484	0,478	0,484
|C1| DF&Acc2 (g)		0,056	0,124	0,226	0,370	0,421	0,469	0,474	0,484	0,478
	MAX	**0,167**	**0,287**	**0,318**	0,379	**0,493**	0,500	**0,490**	**0,488**	0,491
	AVERAGE	0,112	0,207	0,262	0,351	0,443	0,476	0,476	0,476	0,475
Macro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (g)		0,198	0,330	0,350	0,415	0,449	0,488	0,475	0,463	0,480
|C2| tf-idf&IG (g)		0,070	0,204	0,292	0,410	0,453	0,478	0,468	0,469	0,463
|C2| tf-idf&DF (g)		0,014	0,068	0,193	0,314	0,389	0,443	0,460	0,466	0,466
|C2| tf-idf&Acc2 (g)		0,157	0,280	0,369	0,412	0,467	0,481	0,482	0,482	0,477
|C2| CHI&IG (g)		0,163	0,225	0,301	0,346	0,455	0,491	0,473	0,476	0,484
|C2| CHI&DF (g)		0,002	0,144	0,253	0,294	0,384	0,477	0,490	0,486	0,479
|C2| CHI&Acc2 (g)		0,137	0,243	0,282	0,397	0,480	0,481	0,478	0,479	0,491
|C2| IG&DF (g)		0,030	0,064	0,165	0,345	0,466	0,498	0,465	0,478	0,470
|C2| IG&Acc2 (g)		0,112	0,263	0,287	0,386	0,483	0,492	0,484	0,477	0,479
|C2| DF&Acc2 (g)		0,002	0,137	0,236	0,350	0,463	0,492	0,482	0,482	0,482
	MAX	**0,198**	**0,330**	**0,369**	**0,415**	**0,483**	0,498	**0,490**	**0,486**	0,491
	AVERAGE	0,088	0,196	0,273	0,367	0,449	0,482	0,476	0,476	0,477
Macro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (g)		0,145	0,226	0,275	0,322	0,462	0,477	0,478	0,476	0,479
|C3| tf-idf&IG (g)		0,137	0,259	0,279	0,365	0,414	0,459	0,476	0,470	0,471
|C3| tf-idf&DF (g)		0,088	0,117	0,186	0,306	0,393	0,409	0,470	0,471	0,471
|C3| tf-idf&Acc2 (g)		0,162	0,267	0,290	0,368	0,437	0,480	0,486	0,475	0,472
|C3| CHI&IG (g)		0,149	0,258	0,280	0,350	0,405	0,510	0,469	0,484	0,474
|C3| CHI&DF (g)		0,067	0,156	0,238	0,335	0,401	0,497	0,475	0,476	0,472
|C3| CHI&Acc2 (g)		0,177	0,246	0,267	0,357	0,412	0,495	0,484	0,483	0,495
|C3| IG&DF (g)		0,064	0,142	0,172	0,363	0,413	0,482	0,466	0,470	0,470
|C3| IG&Acc2 (g)		0,099	0,296	0,295	0,376	0,465	0,491	0,486	0,478	0,476
|C3| DF&Acc2 (g)		0,064	0,135	0,163	0,362	0,400	0,494	0,480	0,483	0,474
	MAX	**0,177**	**0,296**	0,295	0,376	0,465	**0,510**	**0,486**	**0,484**	**0,495**
	AVERAGE	0,115	0,210	0,245	0,350	0,420	0,479	0,477	0,477	0,475

Macro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (g)		0,146	0,224	0,277	0,397	0,472	0,479	0,473	0,475	0,480
|C4| tf-idf&IG (g)		0,111	0,228	0,266	0,361	0,420	0,480	0,470	0,469	0,473
|C4| tf-idf&DF (g)		0,097	0,132	0,180	0,332	0,383	0,424	0,470	0,464	0,465
|C4| tf-idf&Acc2 (g)		0,162	0,255	0,324	0,372	0,439	0,488	0,492	0,484	0,471
|C4| CHI&IG (g)		0,149	0,273	0,298	0,354	0,432	0,487	0,474	0,480	0,480
|C4| CHI&DF (g)		0,096	0,173	0,309	0,331	0,396	0,485	0,471	0,479	0,479
|C4| CHI&Acc2 (g)		0,167	0,252	0,267	0,355	0,474	0,477	0,470	0,486	0,485
|C4| IG&DF (g)		0,059	0,144	0,210	0,361	0,455	0,487	0,468	0,475	0,471
|C4| IG&Acc2 (g)		0,100	0,255	0,293	0,362	0,482	0,474	0,481	0,480	0,486
|C4| DF&Acc2 (g)		0,059	0,121	0,237	0,400	0,447	0,482	0,479	0,480	0,479
	MAX	0,167	0,273	0,324	0,400	0,482	0,488	0,492	0,486	0,486
	AVERAGE	0,115	0,206	0,266	0,363	0,440	0,476	0,475	0,477	0,477
Macro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (g)		0,145	0,221	0,268	0,333	0,493	0,484	0,484	0,472	0,476
|C5| tf-idf&IG (g)		0,111	0,228	0,269	0,368	0,413	0,483	0,471	0,470	0,473
|C5| tf-idf&DF (g)		0,097	0,117	0,169	0,305	0,388	0,413	0,469	0,465	0,465
|C5| tf-idf&Acc2 (g)		0,162	0,258	0,317	0,371	0,437	0,489	0,493	0,475	0,471
|C5| CHI&IG (g)		0,149	0,258	0,284	0,350	0,432	0,500	0,467	0,480	0,479
|C5| CHI&DF (g)		0,064	0,158	0,239	0,317	0,466	0,487	0,477	0,469	0,464
|C5| CHI&Acc2 (g)		0,148	0,252	0,260	0,359	0,412	0,493	0,480	0,479	0,492
|C5| IG&DF (g)		0,058	0,144	0,202	0,371	0,415	0,490	0,468	0,475	0,470
|C5| IG&Acc2 (g)		0,099	0,265	0,287	0,375	0,470	0,486	0,476	0,480	0,484
|C5| DF&Acc2 (g)		0,059	0,108	0,194	0,364	0,431	0,477	0,480	0,485	0,475
	MAX	0,162	0,265	0,317	0,375	0,493	0,500	0,493	0,485	0,492
	AVERAGE	0,109	0,201	0,249	0,351	0,436	0,480	0,477	0,475	0,475
Macro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (g)		0,146	0,218	0,278	0,365	0,492	0,484	0,484	0,472	0,476
|C6| tf-idf&IG (g)		0,111	0,228	0,266	0,359	0,418	0,483	0,472	0,470	0,472
|C6| tf-idf&DF (g)		0,097	0,116	0,190	0,294	0,389	0,414	0,466	0,474	0,461
|C6| tf-idf&Acc2 (g)		0,133	0,278	0,314	0,363	0,432	0,492	0,492	0,488	0,470
|C6| CHI&IG (g)		0,149	0,277	0,277	0,349	0,433	0,500	0,468	0,477	0,480
|C6| CHI&DF (g)		0,064	0,159	0,268	0,322	0,464	0,498	0,478	0,472	0,470
|C6| CHI&Acc2 (g)		0,167	0,252	0,261	0,356	0,409	0,478	0,479	0,480	0,488
|C6| IG&DF (g)		0,062	0,144	0,203	0,333	0,423	0,471	0,468	0,462	0,469
|C6| IG&Acc2 (g)		0,100	0,266	0,293	0,362	0,474	0,487	0,475	0,480	0,484
|C6| DF&Acc2 (g)		0,059	0,111	0,185	0,329	0,413	0,471	0,475	0,484	0,476
	MAX	0,167	0,278	0,314	0,365	0,492	0,500	0,492	0,488	0,488
	AVERAGE	0,109	0,205	0,254	0,343	0,435	0,478	0,476	0,476	0,475
Macro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (g)		0,145	0,221	0,264	0,332	0,493	0,477	0,480	0,482	0,476
|C7| tf-idf&IG (g)		0,111	0,228	0,269	0,368	0,414	0,484	0,474	0,470	0,471
|C7| tf-idf&DF (g)		0,097	0,117	0,177	0,264	0,389	0,405	0,467	0,477	0,462
|C7| tf-idf&Acc2 (g)		0,162	0,258	0,317	0,371	0,432	0,477	0,491	0,483	0,470
|C7| CHI&IG (g)		0,149	0,256	0,284	0,358	0,440	0,505	0,467	0,478	0,483
|C7| CHI&DF (g)		0,067	0,156	0,232	0,339	0,460	0,498	0,475	0,471	0,467
|C7| CHI&Acc2 (g)		0,177	0,249	0,260	0,362	0,411	0,482	0,478	0,474	0,494
|C7| IG&DF (g)		0,058	0,142	0,167	0,353	0,414	0,458	0,469	0,461	0,473
|C7| IG&Acc2 (g)		0,099	0,289	0,287	0,371	0,474	0,488	0,477	0,478	0,475
|C7| DF&Acc2 (g)		0,064	0,111	0,195	0,313	0,392	0,462	0,474	0,478	0,474
	MAX	0,177	0,289	0,317	0,371	0,493	0,505	0,491	0,483	0,494
	AVERAGE	0,113	0,203	0,245	0,343	0,432	0,474	0,475	0,475	0,474

When we evaluated the performance of the individual feature selection metrics in local policy in Wap dataset, we saw that *tf-idf, Acc2* and *DF* were better than other metrics. Tables 7.52 and 7.53 show the micro- and macro-averaged F-measure results, respectively, for all seven proposed combination methods in local policy for the Wap dataset. Firstly, we determine which combinations are better than others. Among the 10 possible combinations of the two feature selection metrics, the best combinations are *tf-idf & DF, tf-idf & Acc2, IG & Acc2* and *DF & Acc2*.

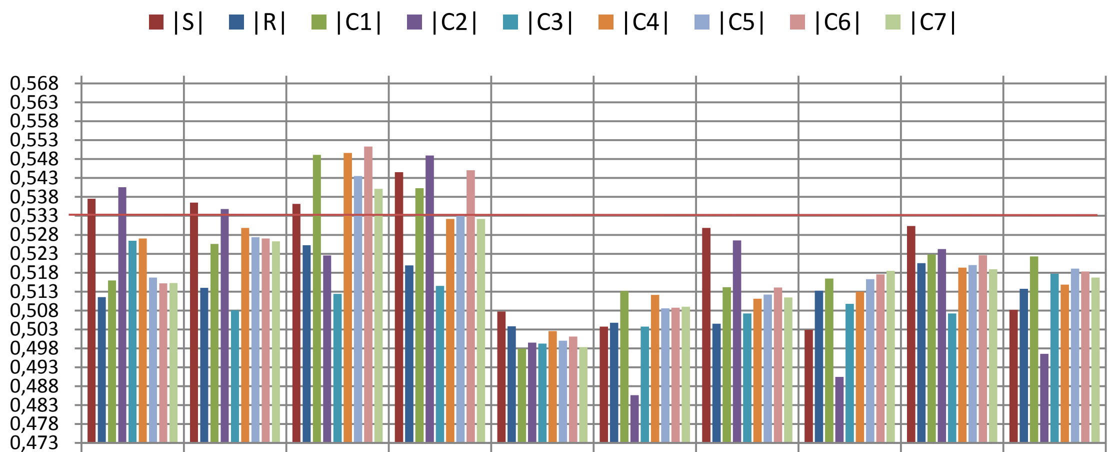


	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,745	0,747	0,748	0,747	0,746	0,744	0,752	0,750	0,755	0,753
|R|	0,744	0,743	0,746	0,744	0,744	0,741	0,746	0,743	0,750	0,751
|C1|	0,741	0,739	0,750	0,744	0,744	0,740	0,750	0,742	0,750	0,752
|C2|	0,742	0,744	0,748	0,748	0,742	0,745	0,755	0,749	0,754	0,757
|C3|	0,741	0,741	0,749	0,754	0,742	0,749	0,755	0,740	0,741	0,754
|C4|	0,742	0,747	0,749	0,742	0,744	0,739	0,748	0,744	0,750	0,754
|C5|	0,744	0,746	0,746	0,745	0,744	0,741	0,748	0,746	0,750	0,752
|C6|	0,741	0,739	0,750	0,744	0,745	0,739	0,750	0,742	0,750	0,753
|C7|	0,744	0,743	0,748	0,740	0,744	0,741	0,746	0,745	0,749	0,750

Figure 7.36. In local policy, averages of the micro-averaged F-measures of all combinations for Wap dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	**0,711**	**0,754**	**0,757**	**0,764**	**0,773**	**0,759**	**0,754**	**0,763**	0,759
Combination | R |	0,700	0,743	0,751	0,762	0,751	0,758	**0,754**	**0,763**	**0,762**
Combination |C1|	0,709	0,751	0,754	0,764	0,759	0,762	0,755	0,766	0,757
Combination |C2|	0,707	0,751	0,757	0,767	0,773	0,755	0,754	0,760	0,761
Combination |C3|	0,715	0,736	0,749	0,752	0,768	0,761	0,752	0,763	0,756
Combination |C4|	0,700	0,749	0,753	0,766	0,758	0,754	0,756	0,761	0,758
Combination |C5|	0,706	0,741	0,756	0,765	0,752	0,754	0,752	0,761	0,759
Combination |C6|	0,701	0,745	0,758	0,766	0,759	0,761	0,754	0,768	0,756
Combination |C7|	0,700	0,738	0,756	0,764	0,755	0,753	0,754	0,763	0,758

Table 7.50. In local policy, maximum micro-averaged F-measures of all combinations for Wap dataset



	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,538	0,536	0,536	0,545	0,508	0,504	0,530	0,503	0,530	0,508
|R|	0,512	0,514	0,525	0,520	0,504	0,505	0,505	0,513	0,521	0,514
|C1|	0,516	0,526	0,549	0,540	0,498	0,513	0,514	0,516	0,523	0,522
|C2|	0,541	0,535	0,523	0,549	0,499	0,486	0,527	0,490	0,524	0,497
|C3|	0,526	0,508	0,512	0,514	0,499	0,504	0,507	0,510	0,507	0,518
|C4|	0,527	0,530	0,550	0,532	0,503	0,512	0,511	0,513	0,519	0,515
|C5|	0,517	0,527	0,544	0,533	0,500	0,509	0,512	0,516	0,520	0,519
|C6|	0,515	0,527	0,551	0,545	0,501	0,509	0,514	0,518	0,523	0,518
|C7|	0,515	0,526	0,540	0,532	0,498	0,509	0,511	0,518	0,519	0,517

Figure 7.37. In local policy, averages of the macro-averaged F-measures of all combinations for Wap dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	0,515	**0,597**	**0,614**	**0,603**	**0,567**	**0,522**	0,500	**0,494**	0,489
Combination | R |	**0,517**	0,573	0,554	0,576	0,532	0,512	**0,502**	0,493	**0,495**
Combination |C1|	0,533	0,581	0,617	0,588	0,542	0,527	0,506	0,494	0,491
Combination |C2|	0,535	0,591	0,611	0,599	0,560	0,517	0,500	0,494	0,490
Combination |C3|	0,525	0,573	0,583	0,528	0,535	0,523	0,507	0,493	0,487
Combination |C4|	0,530	0,590	0,616	0,603	0,530	0,515	0,503	0,492	0,492
Combination |C5|	0,531	0,574	0,588	0,585	0,536	0,514	0,503	0,492	0,493
Combination |C6|	0,527	0,569	0,626	0,585	0,544	0,518	0,504	0,496	0,489
Combination |C7|	0,533	0,568	0,588	0,580	0,536	0,515	0,504	0,493	0,489

Table 7.51. In local policy, maximum macro-averaged F-measures of all combinations for Wap dataset

When we performed score and rank combinations on the dataset in local policy, the first thing that we realized was the score combination improves the performance of the individual metrics and significantly outperforms the rank combination of two features selection metrics. In terms of micro-averaged F-measure, none of the proposed combination is improved the performance of the score combination. Only the C2 reaches the success of the score combination and improves the performance with a few number of keywords. In terms of macro-averaged F-measure, this time C1, C4 and C6 improve the performance of the score combination. It should be noted that the improvement of micro-averaged F-measure is not as successful as the improvement of macro-averaged F-measure. This means that the proposed methods apparently improve the performance of the classifier on rare categories in the Wap dataset.

In local policy, among the individual feature selection metrics *DF* had the highest micro-averaged F-measure value 77.1% with 200 keywords. Only the score combination of *DF & Acc2* improved the highest value 77.3%. This result is only reached by the C2 of the *DF & Acc2.*

Furthermore, the highest macro-averaged F-measure 59.3% was achieved by *tf-idf* with only 30 keywords in local policy. This performance was improved by three score combinations *tf-idf & CHI, tf-idf & DF* and *tf-idf & Acc2* and the highest macro-averaged F-measure 61.4% among them is achieved by *tf-idf & Acc2* with 50 keywords. Furthermore this performance is also improved by C1, C4 and C6. And the highest value 62.6% among them is achieved by C6 of *tf-idf & DF* with only 50 keywords.

Table 7.52. Micro F-measure results of the proposed combinations in local policy for Wap dataset

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (l)		0,671	0,737	0,741	0,738	0,735	0,722	0,746	0,741	0,749
CHI (l)		0,440	0,714	0,732	0,732	0,720	0,736	0,742	0,756	0,758
IG (l)		0,685	0,735	0,750	0,742	0,747	0,744	0,742	0,758	0,749
DF (l)		0,000	0,567	0,704	0,751	0,771	0,747	0,760	0,747	0,747
Acc2 (l)		0,639	0,728	0,757	0,770	0,758	0,755	0,752	0,758	0,752
	MAX	0,685	0,737	0,757	0,770	0,771	0,755	0,760	0,758	0,758
Micro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	0,711	0,754	0,757	0,764	0,773	0,759	0,754	0,763	0,759
	AVERAGE	0,620	0,712	0,744	0,752	0,748	0,742	0,745	0,751	0,753
Rank Combination	MAX	0,700	0,743	0,751	0,762	0,751	0,758	0,754	0,763	0,762
	AVERAGE	0,670	0,734	0,734	0,744	0,741	0,738	0,743	0,751	0,753
Micro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (l)		0,688	0,732	0,736	0,754	0,733	0,729	0,734	0,746	0,749
|C1| tf-idf&IG (l)		0,682	0,736	0,746	0,735	0,741	0,727	0,735	0,744	0,752
|C1| tf-idf&DF (l)		0,661	0,732	0,754	0,764	0,759	0,744	0,743	0,744	0,747
|C1| tf-idf&Acc2 (l)		0,662	0,751	0,746	0,745	0,752	0,743	0,739	0,738	0,748
|C1| CHI&IG (l)		0,707	0,723	0,721	0,740	0,719	0,737	0,747	0,766	0,756
|C1| CHI&DF (l)		0,688	0,734	0,746	0,744	0,744	0,730	0,734	0,740	0,746
|C1| CHI&Acc2 (l)		0,709	0,737	0,740	0,743	0,743	0,748	0,748	0,760	0,757
|C1| IG&DF (l)		0,668	0,748	0,748	0,744	0,740	0,738	0,740	0,744	0,749
|C1| IG&Acc2 (l)		0,675	0,741	0,742	0,755	0,746	0,745	0,755	0,753	0,749
|C1| DF&Acc2 (l)		0,630	0,734	0,744	0,760	0,752	0,762	0,746	0,742	0,748
	MAX	0,709	0,751	0,754	0,764	0,759	0,762	0,755	0,766	0,757
	AVERAGE	0,677	0,737	0,742	0,748	0,743	0,740	0,742	0,748	0,750
Micro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (l)		0,687	0,729	0,752	0,741	0,743	0,727	0,740	0,749	0,753
|C2| tf-idf&IG (l)		0,692	0,738	0,738	0,755	0,726	0,731	0,738	0,752	0,761
|C2| tf-idf&DF (l)		0,559	0,673	0,745	0,762	0,762	0,736	0,737	0,745	0,749
|C2| tf-idf&Acc2 (l)		0,663	0,751	0,754	0,750	0,748	0,747	0,737	0,746	0,759
|C2| CHI&IG (l)		0,707	0,725	0,736	0,720	0,728	0,737	0,748	0,760	0,755
|C2| CHI&DF (l)		0,118	0,695	0,740	0,755	0,736	0,745	0,740	0,742	0,751
|C2| CHI&Acc2 (l)		0,685	0,738	0,757	0,749	0,758	0,753	0,753	0,760	0,758
|C2| IG&DF (l)		0,256	0,638	0,740	0,746	0,753	0,748	0,744	0,753	0,752
|C2| IG&Acc2 (l)		0,671	0,739	0,754	0,767	0,752	0,742	0,754	0,755	0,754
|C2| DF&Acc2 (l)		0,515	0,633	0,727	0,757	0,773	0,755	0,754	0,758	0,747
	MAX	0,707	0,751	0,757	0,767	0,773	0,755	0,754	0,760	0,761
	AVERAGE	0,555	0,706	0,744	0,750	0,748	0,742	0,745	0,752	0,754
Micro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (l)		0,669	0,736	0,724	0,739	0,740	0,728	0,741	0,744	0,756
|C3| tf-idf&IG (l)		0,712	0,726	0,737	0,723	0,738	0,733	0,741	0,756	0,753
|C3| tf-idf&DF (l)		0,660	0,664	0,749	0,741	0,753	0,748	0,745	0,751	0,756
|C3| tf-idf&Acc2 (l)		0,667	0,645	0,727	0,748	0,765	0,755	0,748	0,758	0,752
|C3| CHI&IG (l)		0,711	0,725	0,727	0,725	0,733	0,733	0,742	0,763	0,754
|C3| CHI&DF (l)		0,675	0,654	0,746	0,737	0,758	0,747	0,745	0,754	0,754
|C3| CHI&Acc2 (l)		0,670	0,637	0,726	0,747	0,768	0,760	0,752	0,752	0,748
|C3| IG&DF (l)		0,715	0,728	0,738	0,722	0,733	0,737	0,738	0,756	0,751
|C3| IG&Acc2 (l)		0,708	0,724	0,736	0,727	0,733	0,736	0,747	0,753	0,747
|C3| DF&Acc2 (l)		0,674	0,639	0,732	0,752	0,760	0,761	0,750	0,752	0,747
	MAX	0,715	0,736	0,749	0,752	0,768	0,761	0,752	0,763	0,756
	AVERAGE	0,686	0,688	0,734	0,736	0,748	0,744	0,745	0,754	0,752

Micro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (I)		0,695	0,727	0,735	0,752	0,740	0,723	0,739	0,745	0,753
|C4| tf-idf&IG (I)		0,697	0,735	0,744	0,735	0,749	0,738	0,756	0,746	0,754
|C4| tf-idf&DF (I)		0,660	0,731	0,753	0,766	0,756	0,740	0,741	0,746	0,748
|C4| tf-idf&Acc2 (I)		0,680	0,749	0,740	0,749	0,732	0,733	0,739	0,746	0,753
|C4| CHI&IG (I)		0,700	0,729	0,719	0,732	0,727	0,737	0,751	0,761	0,758
|C4| CHI&DF (I)		0,689	0,745	0,743	0,733	0,723	0,735	0,740	0,752	0,752
|C4| CHI&Acc2 (I)		0,697	0,738	0,739	0,743	0,740	0,737	0,751	0,761	0,758
|C4| IG&DF (I)		0,662	0,740	0,749	0,743	0,729	0,746	0,746	0,745	0,757
|C4| IG&Acc2 (I)		0,674	0,741	0,738	0,750	0,743	0,742	0,754	0,753	0,754
|C4| DF&Acc2 (I)		0,622	0,724	0,739	0,763	0,758	0,754	0,752	0,751	0,749
	MAX	0,700	0,749	0,753	0,766	0,758	0,754	0,756	0,761	0,758
	AVERAGE	0,678	0,736	0,740	0,747	0,740	0,738	0,747	0,751	0,754

Micro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (I)		0,693	0,725	0,729	0,747	0,736	0,729	0,736	0,755	0,759
|C5| tf-idf&IG (I)		0,699	0,730	0,744	0,744	0,743	0,734	0,749	0,752	0,756
|C5| tf-idf&DF (I)		0,671	0,731	0,756	0,759	0,747	0,739	0,740	0,742	0,746
|C5| tf-idf&Acc2 (I)		0,685	0,740	0,748	0,743	0,752	0,736	0,742	0,743	0,753
|C5| CHI&IG (I)		0,706	0,725	0,721	0,723	0,734	0,739	0,751	0,761	0,757
|C5| CHI&DF (I)		0,688	0,739	0,742	0,737	0,738	0,733	0,735	0,749	0,754
|C5| CHI&Acc2 (I)		0,697	0,730	0,744	0,743	0,746	0,740	0,745	0,761	0,755
|C5| IG&DF (I)		0,655	0,741	0,736	0,744	0,744	0,743	0,743	0,747	0,752
|C5| IG&Acc2 (I)		0,674	0,736	0,735	0,751	0,751	0,743	0,752	0,752	0,750
|C5| DF&Acc2 (I)		0,630	0,728	0,739	0,765	0,752	0,754	0,752	0,741	0,747
	MAX	0,706	0,741	0,756	0,765	0,752	0,754	0,752	0,761	0,759
	AVERAGE	0,680	0,732	0,739	0,746	0,744	0,739	0,745	0,750	0,753

Micro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (I)		0,695	0,724	0,741	0,741	0,746	0,725	0,735	0,749	0,749
|C6| tf-idf&IG (I)		0,698	0,737	0,743	0,734	0,734	0,730	0,740	0,743	0,751
|C6| tf-idf&DF (I)		0,660	0,735	0,758	0,764	0,758	0,746	0,741	0,744	0,748
|C6| tf-idf&Acc2 (I)		0,680	0,745	0,749	0,747	0,755	0,734	0,740	0,738	0,748
|C6| CHI&IG (I)		0,700	0,720	0,722	0,738	0,726	0,736	0,748	0,768	0,756
|C6| CHI&DF (I)		0,701	0,733	0,751	0,739	0,743	0,727	0,733	0,744	0,745
|C6| CHI&Acc2 (I)		0,697	0,740	0,742	0,747	0,741	0,746	0,749	0,764	0,754
|C6| IG&DF (I)		0,659	0,739	0,746	0,740	0,737	0,746	0,739	0,745	0,743
|C6| IG&Acc2 (I)		0,671	0,739	0,739	0,751	0,747	0,747	0,754	0,751	0,748
|C6| DF&Acc2 (I)		0,622	0,724	0,744	0,766	0,759	0,761	0,748	0,738	0,748
	MAX	0,701	0,745	0,758	0,766	0,759	0,761	0,754	0,768	0,756
	AVERAGE	0,678	0,734	0,743	0,747	0,745	0,740	0,743	0,748	0,749

Micro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (I)		0,693	0,727	0,722	0,749	0,743	0,730	0,738	0,749	0,753
|C7| tf-idf&IG (I)		0,700	0,730	0,737	0,742	0,740	0,728	0,743	0,747	0,755
|C7| tf-idf&DF (I)		0,654	0,729	0,756	0,761	0,749	0,749	0,743	0,740	0,746
|C7| tf-idf&Acc2 (I)		0,680	0,738	0,748	0,743	0,746	0,732	0,735	0,737	0,744
|C7| CHI&IG (I)		0,696	0,721	0,720	0,731	0,728	0,736	0,749	0,763	0,758
|C7| CHI&DF (I)		0,688	0,738	0,735	0,739	0,737	0,737	0,739	0,747	0,750
|C7| CHI&Acc2 (I)		0,697	0,724	0,739	0,741	0,737	0,742	0,747	0,759	0,752
|C7| IG&DF (I)		0,662	0,738	0,732	0,752	0,738	0,742	0,734	0,748	0,753
|C7| IG&Acc2 (I)		0,671	0,732	0,737	0,753	0,753	0,736	0,754	0,750	0,749
|C7| DF&Acc2 (I)		0,630	0,725	0,737	0,764	0,755	0,753	0,747	0,735	0,748
	MAX	0,700	0,738	0,756	0,764	0,755	0,753	0,754	0,763	0,758
	AVERAGE	0,677	0,730	0,736	0,747	0,743	0,739	0,743	0,748	0,751

Table 7.53. Macro F-measure results of the proposed combinations in local policy for Wap dataset

Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (l)		0,506	0,593	0,565	0,532	0,507	0,495	0,509	0,477	0,483
CHI (l)		0,493	0,511	0,520	0,509	0,462	0,491	0,475	0,488	0,491
IG (l)		0,492	0,531	0,548	0,517	0,508	0,508	0,460	0,490	0,482
DF (l)		0,000	0,353	0,513	0,550	0,538	0,483	0,524	0,483	0,481
Acc2 (l)		0,435	0,554	0,564	0,551	0,509	0,516	0,488	0,491	0,485
	MAX	0,506	0,593	0,565	0,551	0,538	0,516	0,524	0,491	0,491
Macro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	0,515	0,597	0,614	0,603	0,567	0,522	0,500	0,494	0,489
	AVERAGE	0,469	0,538	0,567	0,552	0,512	0,503	0,483	0,484	0,485
Rank Combination	MAX	0,517	0,573	0,554	0,576	0,532	0,512	0,502	0,493	0,495
	AVERAGE	0,484	0,544	0,521	0,526	0,505	0,499	0,482	0,487	0,487
Macro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (l)		0,522	0,509	0,524	0,550	0,507	0,484	0,472	0,478	0,483
|C1| tf-idf&IG (l)		0,512	0,526	0,571	0,531	0,521	0,493	0,473	0,479	0,484
|C1| tf-idf&DF (l)		0,486	0,554	0,617	0,588	0,542	0,509	0,506	0,479	0,482
|C1| tf-idf&Acc2 (l)		0,480	0,574	0,602	0,565	0,523	0,498	0,476	0,475	0,481
|C1| CHI&IG (l)		0,511	0,510	0,488	0,511	0,470	0,496	0,475	0,494	0,488
|C1| CHI&DF (l)		0,507	0,522	0,539	0,521	0,492	0,498	0,473	0,476	0,480
|C1| CHI&Acc2 (l)		0,533	0,531	0,520	0,508	0,492	0,501	0,488	0,491	0,491
|C1| IG&DF (l)		0,462	0,581	0,545	0,520	0,494	0,496	0,478	0,475	0,481
|C1| IG&Acc2 (l)		0,479	0,558	0,547	0,542	0,500	0,512	0,485	0,488	0,486
|C1| DF&Acc2 (l)		0,444	0,551	0,554	0,538	0,519	0,527	0,481	0,477	0,479
	MAX	0,533	0,581	0,617	0,588	0,542	0,527	0,506	0,494	0,491
	AVERAGE	0,494	0,542	0,551	0,537	0,506	0,501	0,481	0,481	0,483
Macro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (l)		0,535	0,576	0,599	0,532	0,513	0,488	0,478	0,481	0,485
|C2| tf-idf&IG (l)		0,506	0,588	0,572	0,562	0,489	0,492	0,480	0,481	0,488
|C2| tf-idf&DF (l)		0,390	0,512	0,569	0,599	0,560	0,506	0,500	0,479	0,482
|C2| tf-idf&Acc2 (l)		0,462	0,591	0,611	0,589	0,530	0,511	0,477	0,478	0,486
|C2| CHI&IG (l)		0,482	0,526	0,511	0,494	0,486	0,497	0,466	0,491	0,487
|C2| CHI&DF (l)		0,309	0,487	0,571	0,543	0,500	0,505	0,474	0,477	0,480
|C2| CHI&Acc2 (l)		0,485	0,568	0,565	0,518	0,506	0,517	0,485	0,494	0,490
|C2| IG&DF (l)		0,364	0,445	0,576	0,539	0,513	0,507	0,479	0,485	0,481
|C2| IG&Acc2 (l)		0,474	0,560	0,556	0,543	0,509	0,504	0,492	0,488	0,489
|C2| DF&Acc2 (l)		0,343	0,448	0,571	0,564	0,536	0,517	0,486	0,485	0,478
	MAX	0,535	0,591	0,611	0,599	0,560	0,517	0,500	0,494	0,490
	AVERAGE	0,435	0,530	0,570	0,548	0,514	0,504	0,482	0,484	0,485
Macro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (l)		0,510	0,573	0,542	0,527	0,510	0,497	0,507	0,480	0,487
|C3| tf-idf&IG (l)		0,512	0,529	0,521	0,499	0,493	0,494	0,461	0,489	0,486
|C3| tf-idf&DF (l)		0,461	0,494	0,578	0,523	0,514	0,505	0,479	0,483	0,484
|C3| tf-idf&Acc2 (l)		0,474	0,470	0,570	0,527	0,528	0,518	0,483	0,484	0,480
|C3| CHI&IG (l)		0,506	0,515	0,494	0,499	0,488	0,494	0,464	0,493	0,486
|C3| CHI&DF (l)		0,470	0,468	0,554	0,512	0,512	0,506	0,479	0,487	0,484
|C3| CHI&Acc2 (l)		0,475	0,443	0,545	0,523	0,535	0,523	0,484	0,482	0,479
|C3| IG&DF (l)		0,525	0,525	0,524	0,500	0,489	0,495	0,460	0,489	0,486
|C3| IG&Acc2 (l)		0,500	0,519	0,524	0,517	0,488	0,495	0,467	0,488	0,483
|C3| DF&Acc2 (l)		0,499	0,447	0,583	0,528	0,529	0,520	0,485	0,481	0,478
	MAX	0,525	0,573	0,583	0,528	0,535	0,523	0,507	0,493	0,487
	AVERAGE	0,493	0,498	0,544	0,516	0,508	0,505	0,477	0,486	0,483

Macro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (I)		0,528	0,530	0,563	0,549	0,508	0,483	0,477	0,478	0,483
|C4| tf-idf&IG (I)		0,530	0,524	0,568	0,533	0,521	0,503	0,489	0,480	0,483
|C4| tf-idf&DF (I)		0,492	0,549	0,616	0,603	0,530	0,509	0,503	0,480	0,482
|C4| tf-idf&Acc2 (I)		0,509	0,590	0,531	0,566	0,505	0,493	0,478	0,479	0,482
|C4| CHI&IG (I)		0,512	0,517	0,488	0,519	0,474	0,506	0,477	0,492	0,489
|C4| CHI&DF (I)		0,521	0,538	0,526	0,508	0,480	0,501	0,475	0,475	0,482
|C4| CHI&Acc2 (I)		0,513	0,532	0,519	0,518	0,488	0,497	0,490	0,491	0,492
|C4| IG&DF (I)		0,462	0,562	0,547	0,516	0,483	0,509	0,485	0,483	0,484
|C4| IG&Acc2 (I)		0,472	0,561	0,540	0,537	0,496	0,510	0,482	0,487	0,489
|C4| DF&Acc2 (I)		0,418	0,542	0,548	0,541	0,525	0,515	0,490	0,484	0,480
	MAX	0,530	0,590	0,616	0,603	0,530	0,515	0,503	0,492	0,492
	AVERAGE	0,496	0,544	0,544	0,539	0,501	0,503	0,485	0,483	0,485

Macro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (I)		0,520	0,506	0,521	0,550	0,516	0,488	0,476	0,489	0,491
|C5| tf-idf&IG (I)		0,531	0,538	0,533	0,543	0,523	0,496	0,489	0,489	0,493
|C5| tf-idf&DF (I)		0,492	0,574	0,588	0,585	0,518	0,504	0,503	0,479	0,482
|C5| tf-idf&Acc2 (I)		0,513	0,565	0,538	0,569	0,521	0,492	0,486	0,477	0,483
|C5| CHI&IG (I)		0,512	0,514	0,496	0,494	0,481	0,503	0,477	0,492	0,488
|C5| CHI&DF (I)		0,510	0,529	0,521	0,509	0,485	0,497	0,475	0,488	0,490
|C5| CHI&Acc2 (I)		0,510	0,524	0,529	0,518	0,499	0,494	0,482	0,491	0,490
|C5| IG&DF (I)		0,458	0,564	0,524	0,536	0,518	0,497	0,485	0,483	0,487
|C5| IG&Acc2 (I)		0,472	0,557	0,523	0,532	0,536	0,500	0,478	0,486	0,486
|C5| DF&Acc2 (I)		0,444	0,546	0,550	0,543	0,517	0,514	0,490	0,478	0,480
	MAX	0,531	0,574	0,588	0,585	0,536	0,514	0,503	0,492	0,493
	AVERAGE	0,496	0,542	0,532	0,538	0,511	0,498	0,484	0,485	0,487

Macro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (I)		0,524	0,504	0,522	0,539	0,519	0,483	0,473	0,481	0,482
|C6| tf-idf&IG (I)		0,527	0,547	0,548	0,529	0,516	0,497	0,475	0,477	0,483
|C6| tf-idf&DF (I)		0,489	0,559	0,626	0,585	0,544	0,506	0,504	0,479	0,482
|C6| tf-idf&Acc2 (I)		0,509	0,569	0,603	0,572	0,527	0,490	0,477	0,476	0,481
|C6| CHI&IG (I)		0,512	0,503	0,498	0,525	0,476	0,494	0,476	0,496	0,488
|C6| CHI&DF (I)		0,491	0,530	0,533	0,513	0,494	0,492	0,472	0,477	0,478
|C6| CHI&Acc2 (I)		0,513	0,534	0,529	0,512	0,489	0,507	0,488	0,494	0,489
|C6| IG&DF (I)		0,459	0,562	0,545	0,524	0,512	0,504	0,477	0,477	0,477
|C6| IG&Acc2 (I)		0,478	0,555	0,528	0,534	0,532	0,509	0,485	0,485	0,484
|C6| DF&Acc2 (I)		0,418	0,544	0,563	0,543	0,524	0,518	0,483	0,475	0,480
	MAX	0,527	0,569	0,626	0,585	0,544	0,518	0,504	0,496	0,489
	AVERAGE	0,492	0,541	0,549	0,538	0,513	0,500	0,481	0,482	0,482

Macro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (I)		0,520	0,510	0,512	0,531	0,524	0,494	0,477	0,480	0,484
|C7| tf-idf&IG (I)		0,533	0,543	0,525	0,544	0,520	0,493	0,481	0,480	0,487
|C7| tf-idf&DF (I)		0,474	0,555	0,588	0,580	0,530	0,515	0,504	0,478	0,481
|C7| tf-idf&Acc2 (I)		0,509	0,568	0,535	0,566	0,519	0,497	0,474	0,474	0,480
|C7| CHI&IG (I)		0,490	0,513	0,491	0,525	0,471	0,501	0,476	0,493	0,489
|C7| CHI&DF (I)		0,523	0,527	0,509	0,510	0,481	0,504	0,478	0,481	0,481
|C7| CHI&Acc2 (I)		0,510	0,520	0,522	0,527	0,486	0,504	0,485	0,491	0,489
|C7| IG&DF (I)		0,476	0,562	0,521	0,534	0,518	0,500	0,474	0,486	0,483
|C7| IG&Acc2 (I)		0,478	0,552	0,527	0,529	0,536	0,491	0,485	0,486	0,485
|C7| DF&Acc2 (I)		0,444	0,544	0,548	0,538	0,514	0,513	0,483	0,473	0,480
	MAX	0,533	0,568	0,588	0,580	0,536	0,515	0,504	0,493	0,489
	AVERAGE	0,496	0,539	0,528	0,538	0,510	0,501	0,482	0,482	0,484

7.3.2. The Reuters-21578 Dataset

7.3.2.1. Property of the Dataset

The final dataset used in this study is Reuters-21578 which is one of the most popular data collections. The Reuters-21578 dataset compiled by David Lewis contains newswire stories in 1987. These documents were manually categorized by the personnel from Reuters Ltd. and Carnegie Group Inc. in 1987. The collection was made available for scientific research in 1990 [43]. Many different versions have been used in past studies and it is considered as the standard benchmark for automatic document organization systems.

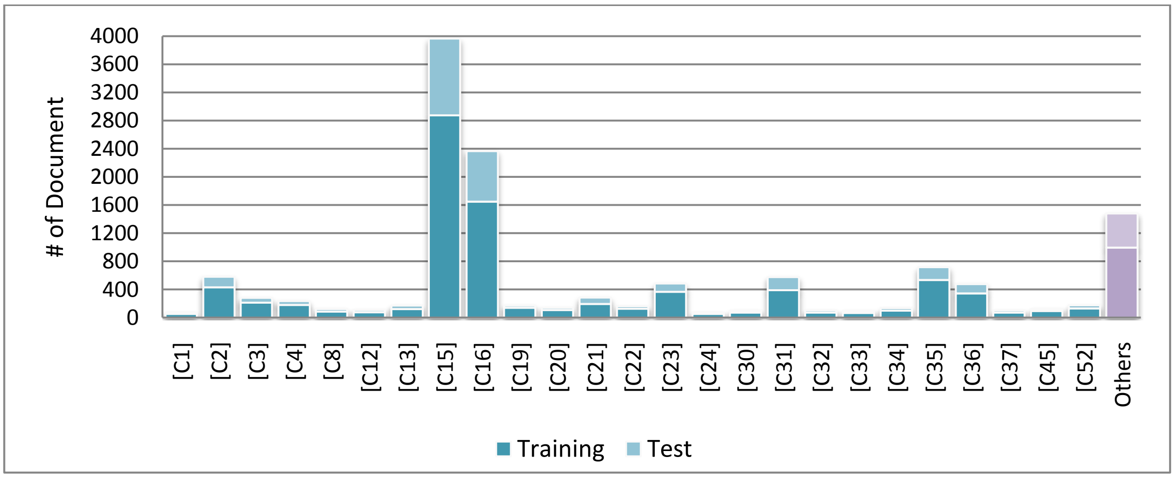


Figure 7.38. Property of the Reuters dataset

In order to divide the Reuters-21578 dataset into training and test sets, we use ModApte splitting method that has been mostly used in the literature [4, 5, 6, 7, 8, 16, 17, 26, 43]. Originally, the Reuters-21578 dataset consists of 21,578 documents that divided into 135 different categories. But with ModApte split the training set consists of 9,603 documents, the test set consists of 3,299 documents and 8,676 documents were unused. The splitting criteria are:

1- The training set consists of any document in the dataset that has at least one category assigned and is dated earlier than April 7th, 1987;
2- The test set consists of any document in the dataset that has at least on category assigned and is dated April 7th, 1987 or later; and
3- The unused set consists of any documents that has no categories assigned to them

After removing the categories that do not exist both in the training set and in the test set, remaining with 90 classes out of 135. Figure 7.38 details the categories of the dataset with ModApte split.

[C1]	cocoa	18	55	73	[C31]	**crude**	189	389	578	[C61]	gas	17	37	54
[C2]	**grain**	149	433	582	[C32]	nat-gas	30	75	105	[C62]	jobs	21	46	67
[C3]	**wheat**	71	212	283	[C33]	cpi	28	69	97	[C63]	lei	3	12	15
[C4]	**corn**	56	181	237	[C34]	gnp	35	101	136	[C64]	yen	14	45	59
[C5]	barley	14	37	51	[C35]	**money-fx**	179	538	717	[C65]	zinc	13	21	34
[C6]	oat	6	8	14	[C36]	**interest**	131	347	478	[C66]	orange	11	16	27
[C7]	sorghum	10	24	34	[C37]	bop	30	75	105	[C67]	pet-chem	12	20	32
[C8]	veg-oil	37	87	124	[C38]	rice	24	35	59	[C68]	fuel	10	13	23
[C9]	lin-oil	1	1	2	[C39]	rubber	12	37	49	[C69]	wpi	10	19	29
[C10]	soy-oil	11	14	25	[C40]	copra-cake	1	2	3	[C70]	potato	3	3	6
[C11]	sun-oil	2	5	7	[C41]	palm-oil	10	30	40	[C71]	lead	14	15	29
[C12]	soybean	33	78	111	[C42]	palmkernel	1	2	3	[C72]	groundnut	4	5	9
[C13]	oilseed	47	124	171	[C43]	tea	4	9	13	[C73]	income	7	9	16
[C14]	sunseed	5	11	16	[C44]	alum	23	35	58	[C74]	palladium	1	2	3
[C15]	**earn**	1087	2877	3964	[C45]	gold	30	94	124	[C75]	nickel	1	8	9
[C16]	**acq**	719	1650	2369	[C46]	platinum	7	5	12	[C76]	lumber	6	10	16
[C17]	copper	18	47	65	[C47]	strategic-metal	11	16	27	[C77]	jet	1	4	5
[C18]	housing	4	16	20	[C48]	tin	12	18	30	[C78]	instal-debt	1	5	6
[C19]	money-supply	34	140	174	[C49]	rapeseed	9	18	27	[C79]	dfl	1	2	3
[C20]	coffee	28	111	139	[C50]	groundnut-oil	1	1	2	[C80]	dmk	4	10	14
[C21]	**ship**	89	197	286	[C51]	rape-oil	3	5	8	[C81]	coconut-oil	3	4	7
[C22]	sugar	36	126	162	[C52]	dlr	44	131	175	[C82]	cpu	1	3	4
[C23]	**trade**	117	369	486	[C53]	l-cattle	2	6	8	[C83]	cotton-oil	2	1	3
[C24]	reserves	18	55	73	[C54]	retail	2	23	25	[C84]	naphtha	4	2	6
[C25]	meal-feed	19	30	49	[C55]	ipi	12	41	53	[C85]	nzdlr	2	2	4
[C26]	soy-meal	13	13	26	[C56]	silver	8	21	29	[C86]	rand	1	2	3
[C27]	rye	1	1	2	[C57]	iron-steel	14	40	54	[C87]	coconut	2	4	6
[C28]	cotton	20	39	59	[C58]	hog	6	16	22	[C88]	castor-oil	1	1	2
[C29]	carcass	18	50	68	[C59]	propane	3	3	6	[C89]	nkr	2	1	3
[C30]	livestock	24	75	99	[C60]	heat	5	14	19	[C90]	sun-meal	1	1	2
Training		9603												
Test		3299												
ALL		12902												

Table 7.54. 20 Categories of the Reuters dataset with ModApte split

These 90 categories are very close to each other thus some of the documents are assigned to more than one category. The maximum number of categories assigned to a document is 14 and the average number of categories per document is 1.24. The 10 top categories which are shown in bold in Table 7.54 constitute about 75 percent of the dataset and the remaining 80 categories constitute only about 25 percent of all documents. In addition, the 2 top categories “earn” and “acq” constitute about 48 percent of the dataset.

7.3.2.2.Analysis of the Existing Metrics

In Table 7.55 we see the micro- and macro-averaged F-measure results of the five well known feature selection metrics with varying number of keywords on the Reuters dataset.

Micro-F	10	30	50	100	200	500	1000	1500	2000	All
tf-idf (g)	0,367	0,565	0,625	0,694	0,760	0,811	0,843	0,858	0,860	0,855
CHI (g)	0,231	0,367	0,531	0,626	0,742	0,798	0,844	0,856	**0,862**	0,855
IG (g)	**0,485**	**0,661**	**0,705**	**0,765**	**0,815**	**0,849**	**0,857**	**0,862**	0,861	0,855
DF (g)	0,412	0,542	0,624	0,679	0,753	0,802	0,839	0,854	0,857	0,855
Acc2 (g)	0,352	0,388	0,513	0,622	0,196	0,814	0,832	0,848	0,860	0,855
MAX	0,485	0,661	0,705	0,765	0,815	0,849	0,857	0,862	0,862	
Macro-F	**10**	**30**	**50**	**100**	**200**	**500**	**1000**	**1500**	**2000**	**All**
tf-idf (g)	0,014	0,031	0,044	0,090	0,163	0,262	0,370	0,417	0,432	0,438
CHI (g)	**0,051**	0,107	**0,163**	**0,242**	**0,377**	0,439	0,476	0,475	0,482	0,438
IG (g)	0,034	0,099	0,140	0,195	0,321	0,392	0,457	0,490	0,476	0,438
DF (g)	0,010	0,034	0,058	0,090	0,147	0,243	0,364	0,411	0,438	0,438
Acc2 (g)	0,039	**0,113**	0,145	0,215	0,193	**0,484**	**0,488**	**0,492**	**0,490**	0,438
MAX	0,051	0,113	0,163	0,242	0,377	0,484	0,488	0,492	0,490	
Micro-F	**10**	**30**	**50**	**100**	**200**	**500**	**1000**	**1500**	**2000**	**All**
tf-idf (l)	0,776	0,812	0,831	0,835	0,838	0,845	0,853	0,850	0,855	0,855
CHI (l)	0,520	**0,823**	**0,840**	0,842	0,839	0,845	0,852	0,855	0,854	0,855
IG (l)	**0,777**	0,820	0,838	0,842	0,845	0,850	0,856	0,858	0,856	0,855
DF (l)	0,725	0,802	0,820	0,841	0,847	0,854	0,859	**0,859**	**0,859**	0,855
Acc2 (l)	0,773	0,811	0,835	**0,846**	**0,855**	**0,860**	**0,862**	**0,859**	**0,859**	0,855
MAX	0,777	0,823	0,840	0,846	0,855	0,860	0,862	0,859	0,859	
Macro-F	**10**	**30**	**50**	**100**	**200**	**500**	**1000**	**1500**	**2000**	**All**
tf-idf (l)	**0,494**	0,512	0,519	0,508	0,514	0,493	0,495	0,491	0,492	0,438
CHI (l)	0,466	0,491	0,493	0,500	0,488	0,493	0,493	0,494	0,491	0,438
IG (l)	**0,494**	**0,530**	0,512	0,517	0,496	0,495	0,493	**0,496**	0,490	0,438
DF (l)	0,463	0,497	0,515	**0,539**	**0,532**	0,511	**0,500**	0,491	**0,493**	0,438
Acc2 (l)	0,492	0,525	**0,524**	0,527	0,515	**0,513**	**0,500**	0,492	0,489	0,438
MAX	0,494	0,530	0,524	0,539	0,532	0,513	0,500	0,496	0,493	

Table 7.55. Micro- and macro-averaged F-measures for Reuters dataset

Like the previous dataset "Wap", the Reuters dataset has highly skew class distribution. When we compare the micro-averaged F-measure results of both datasets, we can see that the Reuters dataset has better F-measures (about 10%) than the Wap dataset.

The main reason for this difference is the number of training documents in the datasets. Although the Reuters dataset has more number of classes (90) than the Wap dataset (20), it has a much larger training set (9603) compared to the Wap dataset (1047). Thus, the classifier more accurately trains the model and gives better results in Reuters dataset rather than Wap in term of the micro-averaged F-measure. On the other hand, when we compare the macro-averaged F-measure results of both datasets, this time we can see that the Reuters dataset has worse F-measure results than the Wap dataset. The main reason for this difference is the number of training documents of the major classes in the Reuters dataset. Only two top classes constitute about 48 percent of the dataset and other 88 classes constitute about 52% in Reuters dataset whereas this percentage drops 35% in Wap dataset and other 18 classes constitute about 65% of the dataset. Thus the model tends to focus on the major classes and ignore the rare ones which lead to more misclassification of minority class in Reuters dataset.

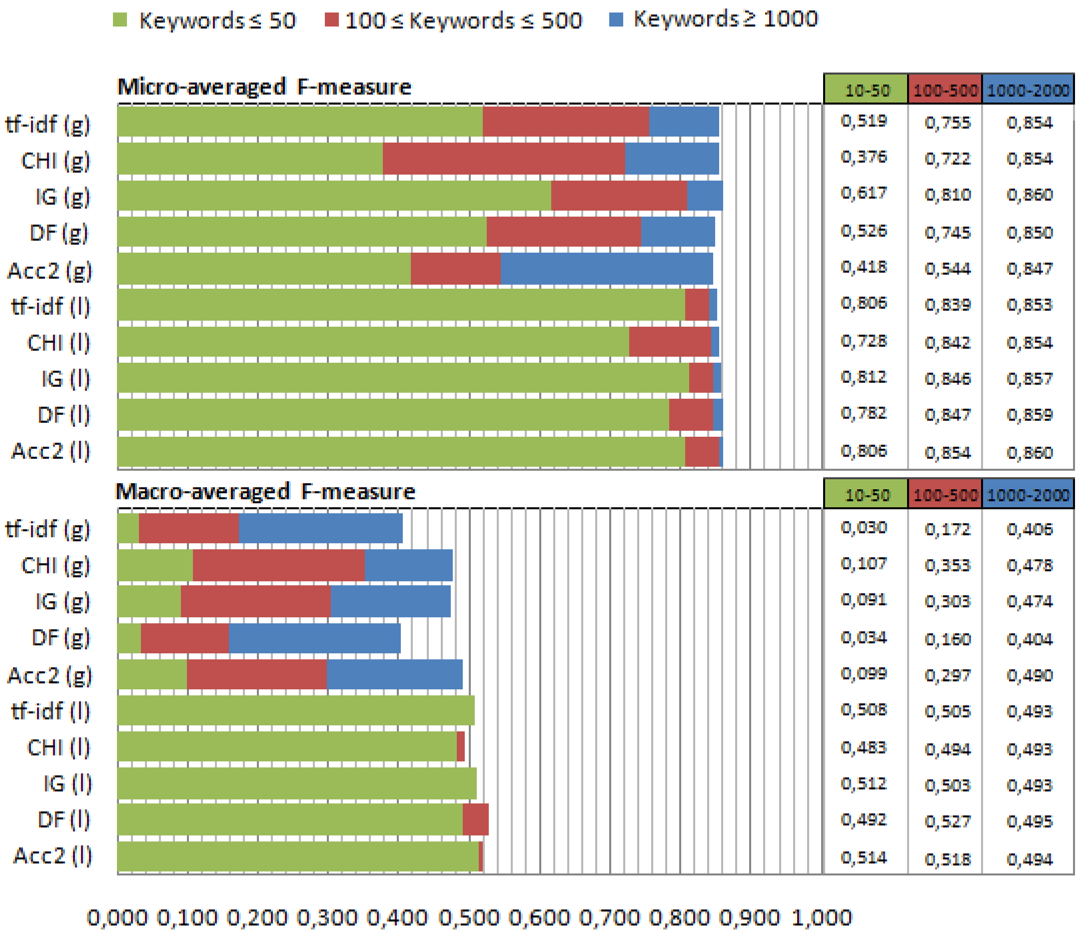


Figure 7.39. Comparison of the averages of micro- and macro-averaged F-measures for Reuters dataset

In the previous experiments, we concluded that micro- and macro-averaged F-measure results of the homogenous datasets were almost equal to each other. On the other hand, in skew (medium) datasets the difference between the micro- and macro-averaged values was clearer, especially when the number of keyword was low. Finally, we tested the existing metrics on the Wap dataset and saw that the difference was significantly higher than the previous datasets used in our study; the average of the difference between two types of F-measures was about 22-27% in this dataset. In Reuters dataset, we observe a similar results related to micro- and macro-averaged F-measures in the other highly skew dataset Wap. The gap between two types of F-measures increases to about 32-44% in the Reuters dataset. As a result, we can conclude that the skewer the class distribution is, the less accurate the model on the rare categories is.

Furthermore, the results in Reuters *dataset* are similar to the results in the Wap *in the case of global and local policies. T*he local policy achieves consistently higher micro- and macro-averaged F-measure values than the global policy when we compare the results of the experiments as seen in Figure 7.39. The superiority of the local policy is over the global policy is emphasized more clearly when we looked at the macro-averaged F-measure results. Especially, this difference is more apparent with a few number of keywords. The average of the difference between the two policies in term of macro-averaged F-measure is about 40% when the keyword number is low. In local policy the highest macro-averaged F-measure values always achieved with a few number of keywords. In Reuters dataset, the highest scores are achieved by the local policy with 1000 keywords in terms of macro-averaged F-measure (86.2%) and with 100 keywords in terms of macro-averaged F-measure (53.9%). As a result, we can say that in highly skew datasets rare classes are classified successfully with a few number of keywords in local policy, since the approach determines significant terms for each class for the classification. Although, the global policy performs worse than the local policy because most of the keywords are selected from the prevailing classes, thus it prevents rare classes to be represented fairly.

When the test documents are categorized without feature selection, the classifier only achieves 85.5% micro- and 43.8% macro-averaged F-measures. On the other hand, both the global and local policies improve the performance with feature selection. For exapmle, the global policy with 1000 and more keywords achieves higher results than the all word approach in terms of micro-averaged F-measure and with 500 and more keywords achieves higher results than the all word approach in terms of macro-averaged F-measure. In addition, this range drops 1000 to 200 and 500 to 10 keywords in local policy in terms of micro- and macro-averaged F-measure respectively.

In Reuters dataset generally *IG* and *Acc2* are better than other metrics in the case of global policy. As seen in Figure 7.39, the classifier achieves better F-measures as the number of keywords increases from 10 to 2000. *IG* especially outperforms other metrics in terms of micro-averaged F-measure but it does not manage to maintain its success in terms of macro-averaged F-measure and it is outperformed by Acc2. On the other hand *DF* and *Acc2* are better than other metrics in the case of local policy. The performance of the classifier is apparently higher with a few number of keywords.

7.3.2.3.Analysis of Score and Rank Combinations

Firstly, the performance of the score and rank combinations are compared with the performance of the existing metrics in the case of global policy. Tables 7.56 and 7.57 show, respectively, the micro- and macro-averaged F-measure results of the score combination and rank combination in global policy for the Reuters dataset. When we look at the first table, the first thing we notice about the results is the poor performance of the score combination in terms of micro-averaged F-measure. Apparently the rank combination is more successful than the score combination. On the other hand, it should be noted that the performance of the score combination is more successful than the rank combination in terms of macro-averaged F-measure when the number of keywords is high as seen in Figure 7.41.

In global policy among the 10 possible combinations of two feature selection metrics, *CHI & IG, CHI & DF, CHI & Acc2* and *IG & Acc2* score combinations and *tf-idf & CHI, tf-idf & Acc2, CHI & IG, CHI & DF* and *IG & Acc2* rank combinations are more successful than other combinations based on the highest micro- and macro-averaged F-measures for each keyword number.

As we evaluated in Section 7.3.2.2, the individual metrics *IG* and *Acc2* were better than others in the case of global policy in Reuters dataset. *IG* especially outperformed other metrics in terms of micro-averaged F-measure but it did not manage to maintain its success in terms of macro-averaged F-measure and it was outperformed by Acc2.

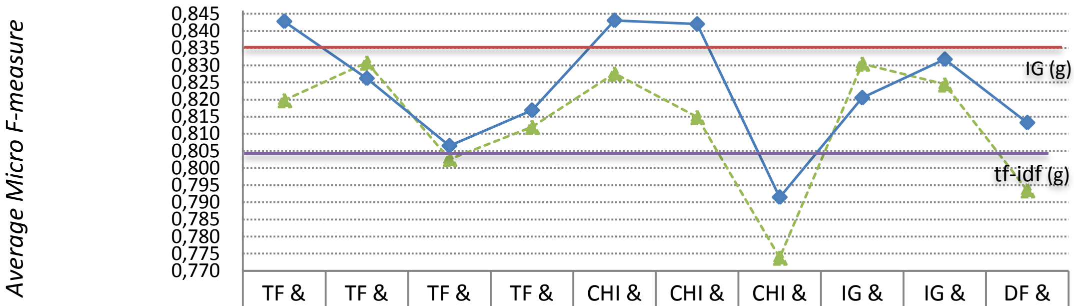
Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,367	0,565	0,625	0,694	0,760	0,811	0,843	0,858	0,860
CHI (g)		0,231	0,367	0,531	0,626	0,742	0,798	0,844	0,856	**0,862**
IG (g)		**0,485**	**0,661**	**0,705**	**0,765**	**0,815**	**0,849**	**0,857**	**0,862**	0,861
DF (g)		0,412	0,542	0,624	0,679	0,753	0,802	0,839	0,854	0,857
Acc2 (g)		0,352	0,388	0,513	0,622	0,196	0,814	0,832	0,848	0,860
	MAX	0,485	0,661	0,705	0,765	0,815	0,849	0,857	0,862	0,862
Micro-F	Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (g)		0,435	0,466	0,581	0,707	0,794	0,845	**0,857**	0,856	0,859
|S| tf-idf&IG (g)		0,423	0,658	0,690	0,764	0,798	0,844	0,855	**0,863**	0,861
|S| tf-idf&DF (g)		0,398	0,587	0,632	0,688	0,759	0,809	0,841	0,859	0,860
|S| tf-idf&Acc2 (g)		0,483	0,563	0,656	0,713	0,772	0,832	0,838	0,853	**0,862**
|S| CHI&IG (g)		0,480	0,527	0,668	0,756	0,788	0,845	0,856	0,859	**0,862**
|S| CHI&DF (g)		0,442	0,462	0,546	0,693	0,785	0,837	0,856	0,856	**0,862**
|S| CHI&Acc2 (g)		0,424	0,508	0,539	0,606	0,666	0,818	0,837	0,855	0,860
|S| IG&DF (g)		0,414	0,653	0,697	0,762	0,800	0,844	0,856	**0,862**	0,859
|S| IG&Acc2 (g)		0,479	0,653	0,691	**0,777**	0,785	0,829	0,842	0,853	0,861
|S| DF&Acc2 (g)		0,319	0,503	0,575	0,654	0,728	0,830	0,838	0,851	0,860
	MAX	0,483	0,658	0,697	**0,777**	0,800	0,845	**0,857**	**0,863**	**0,862**
	AVERAGE	**0,430**	**0,558**	**0,627**	0,712	0,768	0,833	0,848	0,857	**0,861**
Micro-F	Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (g)		**0,503**	**0,686**	**0,724**	**0,786**	**0,832**	**0,857**	**0,862**	**0,862**	0,859
|R| tf-idf&IG (g)		0,493	0,632	0,681	0,749	0,794	0,837	0,855	0,861	0,861
|R| tf-idf&DF (g)		0,465	0,598	0,631	0,707	0,757	0,811	0,847	0,857	0,859
|R| tf-idf&Acc2 (g)		0,193	0,577	0,653	0,715	0,772	0,834	0,853	**0,863**	**0,865**
|R| CHI&IG (g)		**0,499**	**0,696**	**0,746**	**0,784**	**0,836**	**0,859**	**0,860**	0,860	0,860
|R| CHI&DF (g)		**0,496**	**0,689**	**0,730**	**0,788**	**0,831**	**0,855**	**0,860**	0,860	0,859
|R| CHI&Acc2 (g)		0,011	0,039	0,086	0,604	0,745	0,826	0,852	0,859	**0,862**
|R| IG&DF (g)		0,479	0,619	0,669	0,741	0,775	0,834	0,852	0,860	0,861
|R| IG&Acc2 (g)		0,311	0,307	0,649	**0,773**	0,806	0,835	0,856	0,860	0,861
|R| DF&Acc2 (g)		0,067	0,452	0,610	0,705	0,773	0,829	0,852	0,861	0,859
	MAX	**0,503**	**0,696**	**0,746**	**0,788**	**0,836**	**0,859**	**0,862**	**0,863**	**0,865**
	AVERAGE	0,352	0,530	0,618	**0,735**	**0,792**	**0,837**	**0,855**	**0,860**	**0,861**

Table 7.56. In global policy, micro-averaged F-measures of the score and rank combinations for Reuters dataset

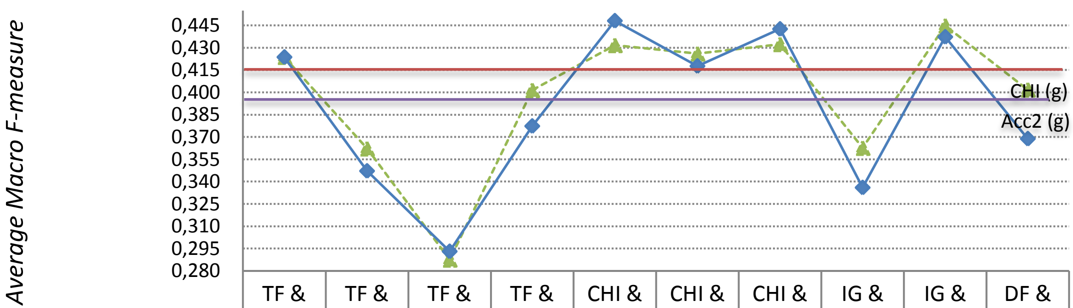
Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,014	0,031	0,044	0,090	0,163	0,262	0,370	0,417	0,432
CHI (g)		**0,051**	0,107	**0,163**	**0,242**	**0,377**	0,439	0,476	0,475	0,482
IG (g)		0,034	0,099	0,140	0,195	0,321	0,392	0,457	0,490	0,476
DF (g)		0,010	0,034	0,058	0,090	0,147	0,243	0,364	0,411	0,438
Acc2 (g)		0,039	**0,113**	0,145	0,215	0,193	**0,484**	**0,488**	**0,492**	**0,490**
	MAX	0,051	0,113	0,163	0,242	0,377	0,484	0,488	0,492	0,490
Macro-F	Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (g)		0,047	0,098	0,133	0,240	0,372	0,463	**0,490**	0,489	0,488
|S| tf-idf&IG (g)		0,012	0,077	0,103	0,174	0,239	0,379	0,430	0,472	0,479
|S| tf-idf&DF (g)		0,010	0,040	0,061	0,093	0,159	0,247	0,373	0,429	0,425
|S| tf-idf&Acc2 (g)		0,020	0,070	0,100	0,177	0,303	0,472	0,476	0,491	**0,490**
|S| CHI&IG (g)		**0,069**	**0,125**	**0,180**	**0,301**	**0,382**	0,467	0,469	0,481	**0,490**
|S| CHI&DF (g)		0,048	0,100	0,137	0,237	**0,380**	0,454	**0,491**	**0,500**	**0,494**
|S| CHI&Acc2 (g)		**0,094**	**0,172**	**0,227**	**0,302**	0,351	0,476	0,482	**0,492**	**0,491**
|S| IG&DF (g)		0,010	0,073	0,112	0,174	0,234	0,380	0,444	0,471	0,474
|S| IG&Acc2 (g)		**0,057**	**0,139**	**0,201**	**0,324**	**0,414**	0,468	0,481	0,490	**0,492**
|S| DF&Acc2 (g)		0,009	0,045	0,071	0,154	0,319	0,479	0,477	0,490	**0,491**
	MAX	**0,094**	**0,172**	**0,227**	**0,324**	**0,414**	0,479	**0,491**	**0,500**	**0,494**
	AVERAGE	0,038	0,094	0,132	0,217	0,315	0,428	0,461	0,481	0,482
Macro-F	Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (g)		0,035	0,110	0,138	**0,253**	0,353	0,464	0,475	**0,498**	**0,498**
|R| tf-idf&IG (g)		0,017	0,061	0,092	0,152	0,219	0,349	0,428	0,466	0,469
|R| tf-idf&DF (g)		0,013	0,041	0,061	0,109	0,156	0,247	0,380	0,421	0,446
|R| tf-idf&Acc2 (g)		0,012	0,079	0,117	0,170	0,256	0,401	0,461	0,482	**0,494**
|R| CHI&IG (g)		**0,056**	**0,179**	**0,254**	**0,335**	**0,439**	0,473	0,479	0,478	0,485
|R| CHI&DF (g)		0,040	0,106	0,153	**0,246**	0,352	0,467	0,473	0,484	0,484
|R| CHI&Acc2 (g)		0,031	0,066	0,149	**0,298**	**0,413**	0,481	0,484	0,488	**0,491**
|R| IG&DF (g)		0,021	0,050	0,089	0,143	0,196	0,336	0,403	0,467	0,471
|R| IG&Acc2 (g)		**0,173**	**0,165**	**0,227**	**0,292**	**0,404**	0,471	0,484	0,484	0,488
|R| DF&Acc2 (g)		0,008	0,044	0,072	0,131	0,256	0,390	0,467	0,485	0,483
	MAX	**0,173**	**0,179**	**0,254**	**0,335**	**0,439**	0,481	0,484	**0,498**	**0,498**
	AVERAGE	0,041	0,090	0,135	0,213	0,304	0,408	0,453	0,475	0,481

Table 7.57. In global policy, macro-averaged F-measures of the score and rank combinations for Reuters dataset

When we look at the Tables 7.56 and 7.57, we can see that *CHI & IG* and *CHI & DF* rank combinations are better than *IG* and *Acc2* in terms of both micro-averaged and macro-averaged F-measures when the number of keywords is lower than 1000. In addition, the rank combination *tf-idf & Acc2* is better than *IG* and *Acc2* in terms of both micro-averaged and macro-averaged F-measures when the number of keywords is higher than 1000 and the rank combination *tf-idf & CHI* apparently outperforms *IG* in terms of micro-averaged F-measure and outperforms *Acc2* in terms of macro-averaged F-measure when the number of keywords is high.



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,820	0,831	0,803	0,812	0,828	0,815	0,774	0,830	0,824	0,793
Rank Com. (g)	0,843	0,826	0,806	0,817	0,843	0,842	0,791	0,820	0,832	0,813



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,424	0,362	0,288	0,401	0,432	0,426	0,432	0,363	0,445	0,402
Rank Com. (g)	0,424	0,347	0,293	0,377	0,448	0,418	0,443	0,336	0,437	0,369

	tf-idf (l)	CHI (l)	IG (l)	DF (l)	Acc2 (l)
Micro	0,804	0,788	0,835	0,797	0,695
Macro	0,289	0,415	0,389	0,282	0,394

Figure 7.40. In global policy, comparison of score and rank combinations on the Reuters dataset

Among the individual metrics *IG* has the highest micro-averaged F-measure 86.2% with 1500 keywords. This value is improved by *tf-idf & CHI* and *tf-idf & Acc2* rank combinations and the highest result among them is 86.5% and it is achieved by the rank combination of *tf-idf & Acc2* with 2000 keywords.

On the other hand, among the individual metrics *Acc2* has the highest macro-averaged F-measure 49.2% with 1500 keywords. This value is also improved by *tf-idf & CHI* and *tf-idf & Acc2* rank combinations and the highest result among them is 49.8% and it is achieved by the rank combination of *tf-idf & CHI* with 1500 keywords. But it should be noted that the performance of *CHI & DF* score combination is significantly better than all combinations when the number of keywords is higher and it has the highest macro-averaged F-measure 50.0% with 1500 keywords.

One of the important results of these experiments is that the performance gap between the existing metrics and the combinations enlarges as the number of keywords decreases on Reuters dataset. Thus, we can say that combining feature selection metrics especially improves the performance of the classifier with a few number of keywords.

Another important result is that the improvement of the macro-averaged F-measure is higher than the improvement of the micro-averaged F-measure on Reuters dataset. Thus, we can say that combining feature selection metrics especially improves the performance of the classifier on rare categories in the dataset. As a result, we can conclude that combining feature selection metrics according to their ranks in global policy improves the performance of the classifier.

Tables 7.56 and 7.57 show, respectively, the micro- and macro-averaged F-measure results of the score combination and rank combination in local policy for the Reuters dataset. Before evaluating the results of the combinations, the first noticeable thing in Tables 7.58 and 7.59 the score combination significantly outperforms the rank combination almost all combinations of two feature selection metrics. In Figure 7.41, we can see the average of the combinations from 100 keywords to 2000. The figure also supports this conclusion. Secondly, combining the existing metrics in local policy is not a good approach that we expected when the number of keywords is high.

In local policy among the 10 possible combinations of two feature selection metrics, *tf-idf & DF, CHI & Acc2, IG & DF, IG & Acc2* and *DF & Acc2* score combinations are more successful than other combinations.

Micro-F	10	30	50	100	200	500	1000	1500	2000
tf-idf (l)	0,776	0,812	0,831	0,835	0,838	0,845	0,853	0,850	0,855
CHI (l)	0,520	**0,823**	**0,840**	0,842	0,839	0,845	0,852	0,855	0,854
IG (l)	**0,777**	0,820	0,838	0,842	0,845	0,850	0,856	0,858	0,856
DF (l)	0,725	0,802	0,820	0,841	0,847	0,854	0,859	**0,859**	**0,859**
Acc2 (l)	0,773	0,811	0,835	**0,846**	**0,855**	**0,860**	**0,862**	**0,859**	**0,859**
MAX	0,777	0,823	0,840	0,846	0,855	0,860	0,862	0,859	0,859
Micro-F Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (l)	**0,780**	**0,825**	0,833	0,838	0,842	0,849	0,855	0,855	0,854
|S| tf-idf&IG (l)	**0,782**	**0,829**	0,838	0,843	0,844	0,852	0,857	0,856	0,856
|S| tf-idf&DF (l)	0,751	0,813	0,827	0,840	0,845	0,847	0,855	0,854	0,856
|S| tf-idf&Acc2 (l)	0,770	0,820	0,837	**0,849**	0,854	0,856	0,858	0,858	0,858
|S| CHI&IG (l)	0,523	0,820	0,836	0,840	0,843	0,848	0,854	0,857	0,857
|S| CHI&DF (l)	0,514	**0,823**	0,833	**0,849**	0,852	0,853	0,856	0,858	0,856
|S| CHI&Acc2 (l)	0,525	0,820	0,838	**0,849**	0,852	0,858	0,859	**0,859**	0,857
|S| IG&DF (l)	0,761	0,815	0,839	**0,848**	0,852	0,855	0,856	**0,859**	**0,859**
|S| IG&Acc2 (l)	0,773	0,816	0,834	**0,848**	0,852	0,859	0,861	0,858	0,858
|S| DF&Acc2 (l)	0,744	0,815	0,832	**0,846**	**0,859**	0,858	0,857	**0,859**	**0,860**
MAX	**0,782**	**0,829**	0,839	**0,849**	**0,859**	0,859	0,861	**0,859**	**0,860**
AVERAGE	0,692	**0,820**	**0,835**	**0,845**	**0,850**	**0,853**	**0,857**	**0,857**	**0,857**
Micro-F Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (l)	**0,784**	0,819	0,831	0,827	0,834	0,845	0,850	0,850	0,851
|R| tf-idf&IG (l)	**0,784**	0,822	0,829	0,830	0,837	0,849	0,854	0,855	0,855
|R| tf-idf&DF (l)	0,764	0,815	0,830	0,842	0,844	0,845	0,857	0,855	0,855
|R| tf-idf&Acc2 (l)	0,765	0,817	0,830	0,838	0,848	0,847	0,855	0,855	0,856
|R| CHI&IG (l)	0,525	0,821	0,838	0,842	0,843	0,848	0,854	0,855	0,857
|R| CHI&DF (l)	0,771	0,821	0,826	0,841	0,845	0,850	0,857	0,855	0,853
|R| CHI&Acc2 (l)	0,776	0,820	0,839	**0,846**	0,851	0,852	0,859	0,857	0,858
|R| IG&DF (l)	0,771	0,819	0,829	0,842	0,849	0,851	0,859	0,856	0,858
|R| IG&Acc2 (l)	0,773	0,816	0,835	**0,847**	0,851	0,857	0,857	**0,859**	0,858
|R| DF&Acc2 (l)	0,760	0,812	0,830	0,838	0,850	0,854	0,858	**0,859**	0,856
MAX	**0,784**	0,822	0,839	**0,847**	0,851	0,857	0,859	**0,859**	0,858
AVERAGE	**0,747**	0,818	0,832	0,839	0,845	0,850	0,856	0,855	0,856

Table 7.58. In local policy, micro-averaged F-measures of the score and rank combinations for Reuters dataset

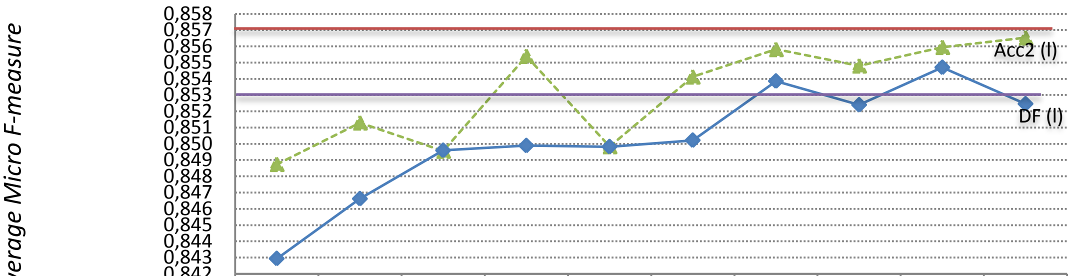
Macro-F	10	30	50	100	200	500	1000	1500	2000
tf-idf (l)	**0,494**	0,512	0,519	0,508	0,514	0,493	0,495	0,491	0,492
CHI (l)	0,466	0,491	0,493	0,500	0,488	0,493	0,493	0,494	0,491
IG (l)	**0,494**	**0,530**	0,512	0,517	0,496	0,495	0,493	**0,496**	0,490
DF (l)	0,463	0,497	0,515	**0,539**	**0,532**	0,511	**0,500**	0,491	**0,493**
Acc2 (l)	0,492	0,525	**0,524**	0,527	0,515	**0,513**	**0,500**	0,492	0,489
MAX	0,494	0,530	0,524	0,539	0,532	0,513	0,500	0,496	0,493
Macro-F Score Combination	10	30	50	100	200	500	1000	1500	2000
|S| tf-idf&CHI (l)	0,493	0,515	0,520	0,506	0,498	0,492	0,492	0,490	0,488
|S| tf-idf&IG (l)	**0,499**	0,524	0,514	0,515	0,502	0,498	0,496	0,490	0,488
|S| tf-idf&DF (l)	**0,497**	**0,536**	0,523	0,536	**0,533**	0,494	0,496	0,490	0,492
|S| tf-idf&Acc2 (l)	**0,494**	0,529	**0,524**	0,529	0,519	0,505	0,495	0,490	0,489
|S| CHI&IG (l)	0,476	0,506	0,498	0,503	0,488	0,495	0,492	**0,496**	0,491
|S| CHI&DF (l)	0,485	0,526	0,505	0,526	0,522	0,506	0,495	0,490	0,488
|S| CHI&Acc2 (l)	0,491	**0,536**	0,513	0,521	0,515	0,510	0,499	0,492	0,489
|S| IG&DF (l)	0,492	**0,541**	0,519	0,531	0,514	0,509	0,497	0,490	0,489
|S| IG&Acc2 (l)	**0,500**	**0,531**	0,520	0,528	0,513	0,512	0,499	0,491	0,490
|S| DF&Acc2 (l)	0,471	**0,532**	0,515	0,531	0,524	0,511	0,495	0,490	0,489
MAX	**0,500**	**0,541**	**0,524**	0,536	**0,533**	0,512	0,499	**0,496**	0,492
AVERAGE	**0,490**	**0,528**	**0,515**	**0,523**	**0,513**	**0,503**	**0,496**	0,491	0,489
Macro-F Rank Combination	10	30	50	100	200	500	1000	1500	2000
|R| tf-idf&CHI (l)	0,482	0,500	0,519	0,500	0,491	0,497	0,493	0,494	0,491
|R| tf-idf&IG (l)	**0,511**	0,510	0,519	0,504	0,494	0,498	0,495	0,495	0,492
|R| tf-idf&DF (l)	0,493	0,525	**0,531**	0,530	0,523	0,493	0,495	0,490	0,492
|R| tf-idf&Acc2 (l)	0,484	**0,533**	0,523	0,521	0,516	0,501	0,494	0,492	0,489
|R| CHI&IG (l)	0,484	0,501	0,501	0,503	0,485	0,497	0,494	0,495	0,491
|R| CHI&DF (l)	0,488	0,507	0,502	0,511	0,501	0,498	0,496	0,493	0,492
|R| CHI&Acc2 (l)	**0,494**	0,528	0,508	0,519	0,502	0,502	0,497	0,492	0,491
|R| IG&DF (l)	0,489	**0,530**	0,508	0,517	0,500	0,502	0,496	0,494	0,490
|R| IG&Acc2 (l)	**0,495**	0,529	0,516	0,523	0,515	0,506	0,495	0,495	0,492
|R| DF&Acc2 (l)	0,477	0,529	0,520	0,527	0,515	0,509	0,495	0,491	0,489
MAX	**0,511**	**0,533**	**0,531**	0,530	0,523	0,509	0,497	0,495	0,492
AVERAGE	**0,490**	0,519	**0,515**	0,516	0,504	0,500	0,495	**0,493**	**0,491**

Table 7.59. In local policy, macro-averaged F-measures of the score and rank combinations for Reuters dataset

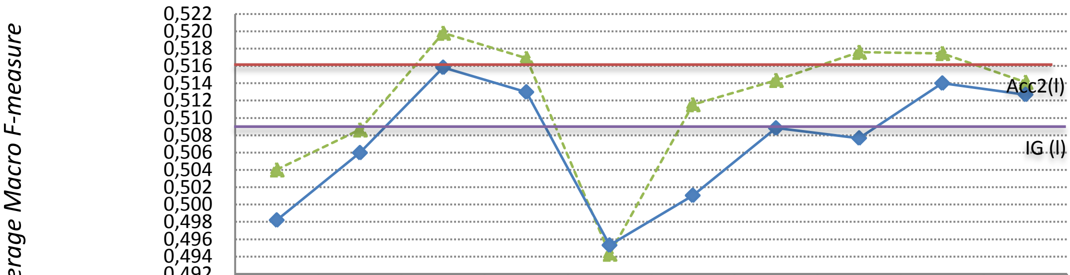
In the previous section, DF and Acc2 were better than other metrics in the case of local policy. In terms of micro-averaged F-measure *Acc2* significantly outperformed other metrics and had the highest result 86.2% with 1000 keywords. When we look at the Table

7.58, we can see that none of the combinations improves the highest value. Despite the worsening of the highest value, tf-idf & IG and *DF & Acc2* improve the performance of the classifier with a few number of keywords.

Furthermore, DF had the highest result macro-averaged F-measure 53.9% with 100 keywords. When we look at the Table 7.58, we can see that the only score combination of *IG & DF* improves the value 54.1% with only 30 keywords.



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (I)	0,849	0,851	0,850	0,855	0,850	0,854	0,856	0,855	0,856	0,857
Rank Com. (I)	0,843	0,847	0,850	0,850	0,850	0,850	0,854	0,852	0,855	0,852



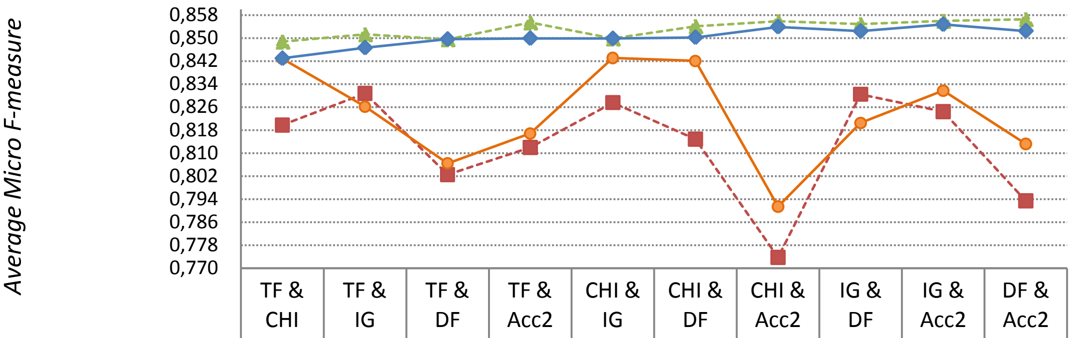
	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (I)	0,504	0,509	0,520	0,517	0,494	0,512	0,514	0,518	0,517	0,514
Rank Com. (I)	0,498	0,506	0,516	0,513	0,495	0,501	0,509	0,508	0,514	0,513

	tf-idf (I)	CHI (I)	IG (I)	DF (I)	Acc2 (I)
Micro	0,846	0,848	0,851	0,853	0,857
Macro	0,507	0,489	0,507	0,509	0,516

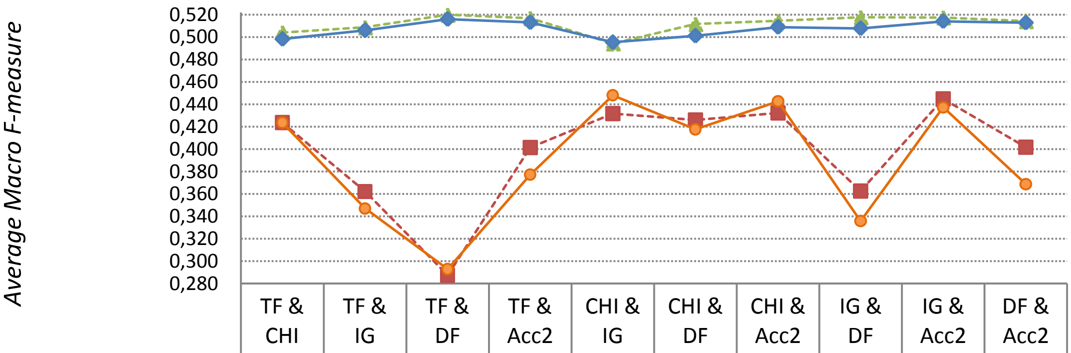
Figure 7.41. In local policy, comparison of score and rank combinations on the Reuters dataset

In Figure 7.42 we can clearly see that the local policy achieves consistently higher performance than the global policy but it should be noted that the global policy is more successful than the local policy when the number of keywords is high.

As a result, the improvement in local policy is not as high as we expected in Reuters dataset and we can conclude that combining feature selection metrics according to their ranks in global policy improves the performance of the classifier whereas combining feature selection metrics according to their scores in local policy improves the performance of the classifier with only a few number of keywords.



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,820	0,831	0,803	0,812	0,828	0,815	0,774	0,830	0,824	0,793
Rank Com. (g)	0,843	0,826	0,806	0,817	0,843	0,842	0,791	0,820	0,832	0,813
Score Com. (l)	0,849	0,851	0,850	0,855	0,850	0,854	0,856	0,855	0,856	0,857
Rank Com. (l)	0,843	0,847	0,850	0,850	0,850	0,850	0,854	0,852	0,855	0,852



	TF & CHI	TF & IG	TF & DF	TF & Acc2	CHI & IG	CHI & DF	CHI & Acc2	IG & DF	IG & Acc2	DF & Acc2
Score Com. (g)	0,424	0,362	0,288	0,401	0,432	0,426	0,432	0,363	0,445	0,402
Rank Com. (g)	0,424	0,347	0,293	0,377	0,448	0,418	0,443	0,336	0,437	0,369
Score Com. (l)	0,504	0,509	0,520	0,517	0,494	0,512	0,514	0,518	0,517	0,514
Rank Com. (l)	0,498	0,506	0,516	0,513	0,495	0,501	0,509	0,508	0,514	0,513

Figure 7.42. Comparison of score and rank combinations in global and local policy on the Reuters dataset

7.3.2.4. Analysis of the Proposed Combinations

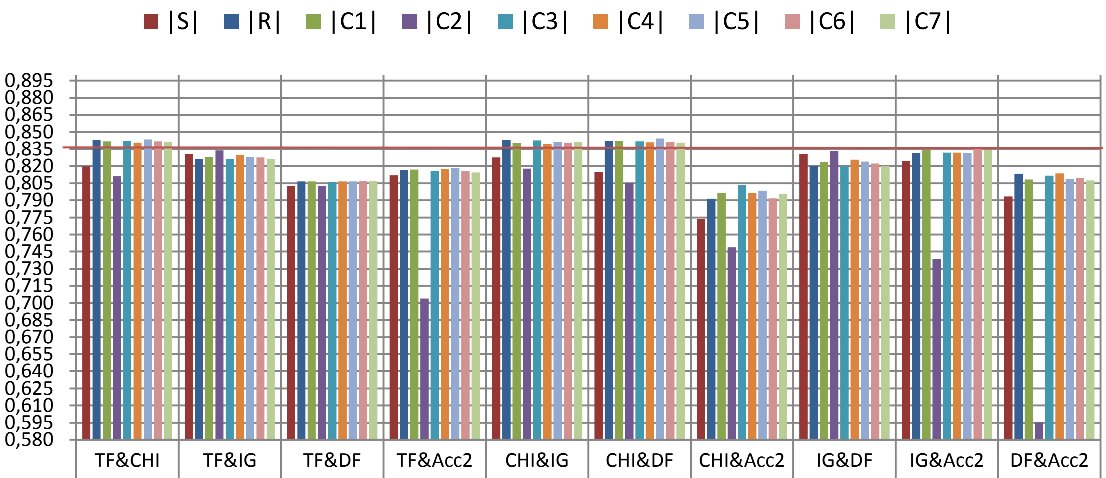
We knew, from the previous sections, that the individual metrics *IG* and *Acc2* were better than others in the case of global policy in Reuters dataset. The former outperformed other metrics in terms of micro-averaged F-measure and the latter outperformed other metrics in terms of macro-averaged F-measure. When we evaluated the performance of the score and rank combinations, we saw that the rank combination was more successful than the score combination, especially in terms of micro-averaged F-measure, in global policy. After testing the score and rank combinations, we now test the proposed methods on the Reuters dataset.

Figures 7.43 and 7.44 demonstrate the averages of the micro- and macro-averaged F-measures of all combinations in global policy for the Reuters dataset. The first thing that we realize in Figure 7.43 is the averages of all combinations are very close to each other and are quite high according to general. On the other hand, the difference of the averages becomes clearer when we look at the Figure 7.44.

Tables 7.62 and 7.63 show the micro- and macro-averaged F-measures, respectively, for all seven proposed methods in global policy for the Reuters dataset. In terms of micro-averaged F-measure, among the 10 possible combinations of two metrics, the combinations of *tf-idf & CHI, tf-idf & Acc2, CHI & IG, CHI & DF* and *IG & Acc2* are significantly better than others and in terms of macro-averaged F-measure the combinations of *tf-idf & CHI, CHI & IG, CHI & DF, CHI & Acc2, IG & Acc2* are significantly better than others. *Tf-idf & DF* and *IG & DF* are the worst combinations in this dataset.

In terms of micro-averaged F-measure all proposed methods except C2 improve the performance of the individual methods and among the proposed methods C3 and C7 are slightly better than others but as we mentioned before the rank combination was the best method when we compared the score and rank combinations. It is still the best method after testing all proposed methods. Only the C3 is more successful than the rank combination with a few number of keywords.

In terms of macro-averaged F-measure the performance of the rank combination was better than the score combination with a few number of keywords whereas the performance of the score combination was better than the rank combination with a high number of keywords. After testing the proposed methods, it is seen that the success of the rank combination is outperformed by the C3 and the success of the score combination is outperformed by the C2.

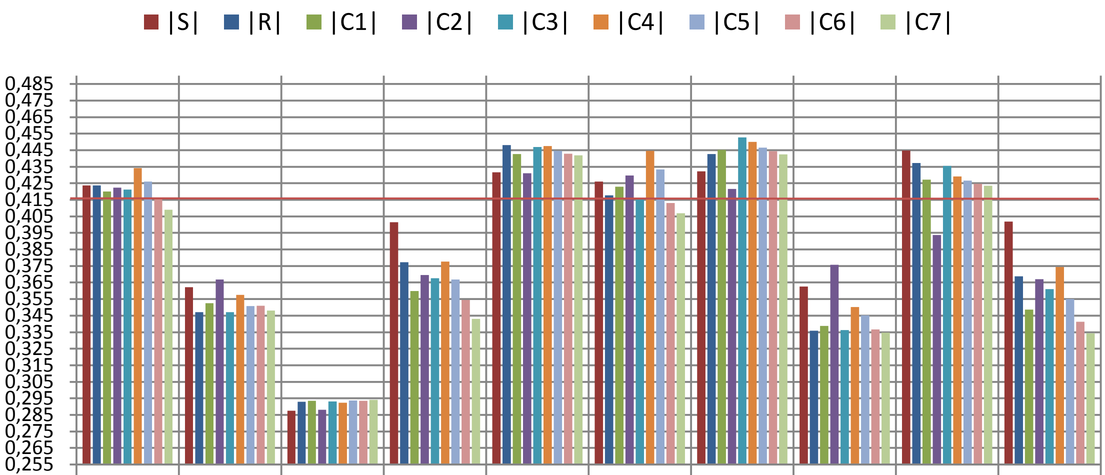


	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,820	0,831	0,803	0,812	0,828	0,815	0,774	0,830	0,824	0,793
|R|	0,843	0,826	0,806	0,817	0,843	0,842	0,791	0,820	0,832	0,813
|C1|	0,842	0,828	0,807	0,817	0,840	0,842	0,797	0,823	0,834	0,808
|C2|	0,811	0,834	0,802	0,704	0,818	0,805	0,749	0,833	0,739	0,595
|C3|	0,842	0,826	0,806	0,816	0,843	0,842	0,803	0,820	0,832	0,812
|C4|	0,841	0,830	0,807	0,817	0,839	0,841	0,797	0,826	0,832	0,814
|C5|	0,843	0,828	0,807	0,818	0,841	0,844	0,798	0,824	0,832	0,808
|C6|	0,842	0,828	0,807	0,816	0,841	0,841	0,792	0,822	0,835	0,810
|C7|	0,841	0,826	0,807	0,814	0,841	0,841	0,796	0,820	0,834	0,807

Figure 7.43. In global policy, averages of the micro-averaged F-measures of all combinations for Reuters dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	0,483	0,658	0,697	0,777	0,800	0,845	0,857	**0,863**	0,862
Combination | R |	**0,503**	**0,696**	**0,746**	**0,788**	**0,836**	**0,859**	**0,862**	**0,863**	**0,865**
Combination |C1|	0,491	0,690	0,739	0,788	0,829	0,856	0,860	0,863	0,863
Combination |C2|	0,476	0,666	0,700	0,772	0,811	0,847	0,858	0,862	0,861
Combination |C3|	0,504	0,702	0,747	0,786	0,834	0,858	0,861	0,862	0,865
Combination |C4|	0,496	0,693	0,741	0,791	0,826	0,858	0,859	0,863	0,863
Combination |C5|	0,503	0,690	0,747	0,796	0,833	0,857	0,861	0,863	0,862
Combination |C6|	0,489	0,696	0,745	0,784	0,829	0,859	0,861	0,864	0,863
Combination |C7|	0,493	0,677	0,745	0,781	0,833	0,858	0,861	0,863	0,864

Table 7.60. Maximum micro-averaged F-measures for all combinations in global policy for Reuters dataset



	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,424	0,362	0,288	0,401	0,432	0,426	0,432	0,363	0,445	0,402
|R|	0,424	0,347	0,293	0,377	0,448	0,418	0,443	0,336	0,437	0,369
|C1|	0,420	0,352	0,293	0,360	0,443	0,423	0,445	0,339	0,427	0,349
|C2|	0,422	0,367	0,288	0,370	0,431	0,430	0,422	0,376	0,394	0,367
|C3|	0,421	0,347	0,293	0,368	0,447	0,416	0,453	0,336	0,436	0,361
|C4|	0,434	0,357	0,292	0,378	0,447	0,445	0,450	0,350	0,429	0,374
|C5|	0,426	0,351	0,294	0,367	0,445	0,433	0,447	0,345	0,427	0,355
|C6|	0,416	0,351	0,294	0,354	0,443	0,413	0,444	0,337	0,425	0,341
|C7|	0,409	0,348	0,294	0,343	0,442	0,407	0,442	0,335	0,424	0,334

Figure 7.44. In global policy, averages of the macro-averaged F-measures of all combinations for Reuters dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	0,094	0,172	0,227	0,324	0,414	0,479	**0,491**	**0,500**	0,494
Combination | R |	**0,173**	**0,179**	**0,254**	**0,335**	**0,439**	**0,481**	0,484	0,498	**0,498**
Combination |C1|	0,094	0,157	0,211	0,328	0,414	0,482	0,487	0,495	0,491
Combination |C2|	0,094	0,170	0,231	0,292	0,387	0,486	0,491	0,500	0,497
Combination |C3|	0,064	0,175	0,282	0,350	0,438	0,485	0,485	0,498	0,498
Combination |C4|	0,083	0,165	0,218	0,343	0,410	0,485	0,493	0,493	0,491
Combination |C5|	0,083	0,185	0,238	0,339	0,431	0,478	0,489	0,493	0,492
Combination |C6|	0,083	0,158	0,221	0,326	0,419	0,478	0,490	0,495	0,490
Combination |C7|	0,083	0,179	0,231	0,320	0,415	0,477	0,486	0,493	0,502

Table 7.61. Maximum macro-averaged F-measures for all combinations in global policy for Reuters dataset

As we knew that among the individual metrics *IG* has the highest micro-averaged F-measure 86.2% with 1500 keywords. That performance was improved by *tf-idf & CHI* and *tf-idf & Acc2* rank combinations and the highest result among them was 86.5% and it was achieved by the rank combination of *tf-idf & Acc2* with 2000 keywords. Only C3 of *tf-idf & Acc2* reaches this performance with 2000.

Table 7.62. Micro F-measure results of the proposed combinations in global policy for Reuters dataset

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,367	0,565	0,625	0,694	0,760	0,811	0,843	0,858	0,860
CHI (g)		0,231	0,367	0,531	0,626	0,742	0,798	0,844	0,856	0,862
IG (g)		0,485	0,661	0,705	0,765	0,815	0,849	0,857	0,862	0,861
DF (g)		0,412	0,542	0,624	0,679	0,753	0,802	0,839	0,854	0,857
Acc2 (g)		0,352	0,388	0,513	0,622	0,196	0,814	0,832	0,848	0,860
	MAX	0,485	0,661	0,705	0,765	0,815	0,849	0,857	0,862	0,862
Micro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	0,483	0,658	0,697	0,777	0,800	0,845	0,857	0,863	0,862
	AVERAGE	0,430	0,558	0,627	0,712	0,768	0,833	0,848	0,857	0,861
Rank Combination	MAX	0,503	0,696	0,746	0,788	0,836	0,859	0,862	0,863	0,865
	AVERAGE	0,352	0,530	0,618	0,735	0,792	0,837	0,855	0,860	0,861
Micro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (g)		0,457	0,664	0,722	0,788	0,829	0,856	0,858	0,860	0,860
|C1| tf-idf&IG (g)		0,477	0,659	0,686	0,761	0,792	0,838	0,854	0,861	0,861
|C1| tf-idf&DF (g)		0,399	0,596	0,635	0,706	0,760	0,811	0,847	0,858	0,859
|C1| tf-idf&Acc2 (g)		0,454	0,584	0,645	0,715	0,774	0,834	0,853	0,863	0,863
|C1| CHI&IG (g)		0,491	0,690	0,739	0,782	0,829	0,856	0,858	0,859	0,859
|C1| CHI&DF (g)		0,454	0,678	0,738	0,787	0,828	0,856	0,860	0,862	0,860
|C1| CHI&Acc2 (g)		0,424	0,457	0,554	0,623	0,752	0,834	0,853	0,857	0,860
|C1| IG&DF (g)		0,416	0,647	0,692	0,741	0,788	0,836	0,854	0,860	0,862
|C1| IG&Acc2 (g)		0,472	0,649	0,691	0,772	0,811	0,844	0,855	0,860	0,863
|C1| DF&Acc2 (g)		0,461	0,557	0,638	0,700	0,750	0,825	0,851	0,861	0,861
	MAX	0,491	0,690	0,739	0,788	0,829	0,856	0,860	0,863	0,863
	AVERAGE	0,451	0,618	0,674	0,737	0,791	0,839	0,854	0,860	0,861
Micro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (g)		0,443	0,466	0,501	0,679	0,778	0,838	0,854	0,859	0,858
|C2| tf-idf&IG (g)		0,413	0,666	0,700	0,772	0,807	0,844	0,857	0,861	0,861
|C2| tf-idf&DF (g)		0,398	0,6052	0,637	0,688	0,759	0,810	0,840	0,858	0,859
|C2| tf-idf&Acc2 (g)		0,376	0,376	0,377	0,388	0,463	0,824	0,835	0,853	0,861
|C2| CHI&IG (g)		0,428	0,490	0,548	0,718	0,780	0,834	0,856	0,859	0,860
|C2| CHI&DF (g)		0,231	0,462	0,491	0,645	0,785	0,830	0,853	0,859	0,860
|C2| CHI&Acc2 (g)		0,424	0,446	0,494	0,548	0,577	0,816	0,837	0,853	0,860
|C2| IG&DF (g)		0,431	0,656	0,692	0,761	0,811	0,847	0,858	0,862	0,860
|C2| IG&Acc2 (g)		0,476	0,476	0,481	0,509	0,558	0,817	0,835	0,853	0,860
|C2| DF&Acc2 (g)		0,352	0,388	0,003	0,028	0,188	0,814	0,832	0,850	0,860
	MAX	0,476	0,666	0,700	0,772	0,811	0,847	0,858	0,862	0,861
	AVERAGE	0,397	0,503	0,493	0,573	0,651	0,827	0,846	0,857	0,860
Micro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (g)		0,504	0,686	0,722	0,783	0,832	0,857	0,861	0,862	0,859
|C3| tf-idf&IG (g)		0,493	0,632	0,682	0,749	0,794	0,836	0,855	0,861	0,861
|C3| tf-idf&DF (g)		0,464	0,598	0,631	0,707	0,757	0,811	0,847	0,857	0,859
|C3| tf-idf&Acc2 (g)		0,041	0,460	0,652	0,713	0,773	0,830	0,852	0,862	0,865
|C3| CHI&IG (g)		0,499	0,702	0,747	0,783	0,834	0,858	0,860	0,860	0,860
|C3| CHI&DF (g)		0,496	0,696	0,730	0,786	0,831	0,855	0,860	0,860	0,859
|C3| CHI&Acc2 (g)		0,049	0,206	0,569	0,660	0,750	0,837	0,853	0,859	0,862
|C3| IG&DF (g)		0,479	0,619	0,669	0,741	0,775	0,834	0,852	0,860	0,861
|C3| IG&Acc2 (g)		0,173	0,305	0,509	0,768	0,807	0,839	0,855	0,860	0,862
|C3| DF&Acc2 (g)		0,220	0,372	0,539	0,704	0,772	0,821	0,852	0,861	0,859
	MAX	0,504	0,702	0,747	0,786	0,834	0,858	0,861	0,862	0,865
	AVERAGE	0,342	0,528	0,645	0,739	0,792	0,838	0,855	0,860	0,861

Micro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (g)		0,458	0,669	0,720	0,785	0,826	0,855	0,859	0,858	0,860
|C4| tf-idf&IG (g)		0,476	0,660	0,686	0,761	0,794	0,844	0,855	0,863	0,861
|C4| tf-idf&DF (g)		0,402	0,569	0,632	0,706	0,759	0,811	0,846	0,858	0,860
|C4| tf-idf&Acc2 (g)		0,468	0,582	0,647	0,714	0,780	0,839	0,850	0,858	0,863
|C4| CHI&IG (g)		0,494	0,693	0,740	0,779	0,824	0,858	0,858	0,859	0,860
|C4| CHI&DF (g)		0,496	0,678	0,741	0,791	0,823	0,853	0,857	0,860	0,861
|C4| CHI&Acc2 (g)		0,075	0,532	0,554	0,643	0,743	0,828	0,851	0,856	0,859
|C4| IG&DF (g)		0,479	0,651	0,693	0,742	0,793	0,843	0,854	0,860	0,862
|C4| IG&Acc2 (g)		0,484	0,644	0,693	0,770	0,808	0,842	0,850	0,860	0,861
|C4| DF&Acc2 (g)		0,459	0,554	0,632	0,700	0,781	0,835	0,848	0,854	0,863
	MAX	0,496	0,693	0,741	0,791	0,826	0,858	0,859	0,863	0,863
	AVERAGE	0,429	0,623	0,674	0,739	0,793	0,841	0,853	0,859	0,861

Micro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (g)		0,503	0,668	0,722	0,791	0,832	0,857	0,861	0,859	0,860
|C5| tf-idf&IG (g)		0,476	0,661	0,671	0,755	0,796	0,837	0,856	0,862	0,862
|C5| tf-idf&DF (g)		0,464	0,596	0,632	0,707	0,758	0,810	0,848	0,857	0,859
|C5| tf-idf&Acc2 (g)		0,467	0,585	0,645	0,715	0,783	0,834	0,853	0,863	0,862
|C5| CHI&IG (g)		0,491	0,678	0,747	0,782	0,832	0,856	0,858	0,860	0,859
|C5| CHI&DF (g)		0,496	0,690	0,732	0,796	0,833	0,857	0,859	0,860	0,860
|C5| CHI&Acc2 (g)		0,075	0,519	0,567	0,644	0,742	0,836	0,852	0,859	0,858
|C5| IG&DF (g)		0,479	0,636	0,675	0,744	0,785	0,838	0,855	0,861	0,862
|C5| IG&Acc2 (g)		0,202	0,654	0,690	0,766	0,805	0,842	0,856	0,859	0,861
|C5| DF&Acc2 (g)		0,190	0,543	0,630	0,698	0,756	0,822	0,852	0,860	0,862
	MAX	0,503	0,690	0,747	0,796	0,833	0,857	0,861	0,863	0,862
	AVERAGE	0,384	0,623	0,671	0,740	0,792	0,839	0,855	0,860	0,861

Micro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (g)		0,465	0,660	0,725	0,784	0,829	0,855	0,861	0,861	0,860
|C6| tf-idf&IG (g)		0,475	0,661	0,670	0,757	0,794	0,839	0,854	0,862	0,861
|C6| tf-idf&DF (g)		0,397	0,569	0,634	0,707	0,759	0,811	0,847	0,858	0,859
|C6| tf-idf&Acc2 (g)		0,455	0,583	0,647	0,718	0,766	0,832	0,852	0,864	0,863
|C6| CHI&IG (g)		0,484	0,696	0,745	0,781	0,828	0,859	0,859	0,858	0,859
|C6| CHI&DF (g)		0,455	0,674	0,742	0,781	0,828	0,856	0,860	0,861	0,860
|C6| CHI&Acc2 (g)		0,075	0,501	0,563	0,622	0,723	0,835	0,852	0,859	0,858
|C6| IG&DF (g)		0,479	0,636	0,685	0,739	0,783	0,836	0,853	0,860	0,862
|C6| IG&Acc2 (g)		0,489	0,652	0,693	0,770	0,814	0,845	0,858	0,860	0,863
|C6| DF&Acc2 (g)		0,458	0,554	0,631	0,704	0,761	0,822	0,850	0,861	0,860
	MAX	0,489	0,696	0,745	0,784	0,829	0,859	0,861	0,864	0,863
	AVERAGE	0,423	0,619	0,673	0,736	0,789	0,839	0,855	0,860	0,861

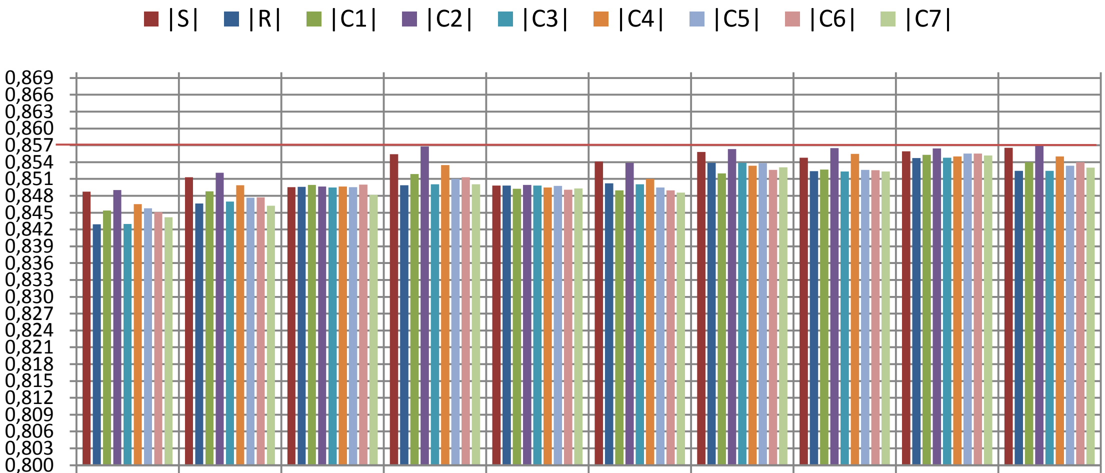
Micro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (g)		0,471	0,670	0,721	0,778	0,830	0,856	0,860	0,861	0,862
|C7| tf-idf&IG (g)		0,476	0,633	0,671	0,748	0,793	0,837	0,856	0,861	0,862
|C7| tf-idf&DF (g)		0,465	0,598	0,631	0,707	0,759	0,812	0,847	0,858	0,859
|C7| tf-idf&Acc2 (g)		0,440	0,605	0,645	0,713	0,766	0,828	0,854	0,861	0,864
|C7| CHI&IG (g)		0,491	0,674	0,745	0,780	0,833	0,858	0,857	0,859	0,859
|C7| CHI&DF (g)		0,493	0,677	0,730	0,781	0,829	0,853	0,861	0,860	0,861
|C7| CHI&Acc2 (g)		0,075	0,517	0,578	0,622	0,740	0,836	0,852	0,863	0,860
|C7| IG&DF (g)		0,479	0,614	0,675	0,744	0,772	0,833	0,852	0,859	0,862
|C7| IG&Acc2 (g)		0,201	0,617	0,691	0,768	0,812	0,847	0,857	0,860	0,862
|C7| DF&Acc2 (g)		0,190	0,543	0,630	0,697	0,757	0,821	0,846	0,860	0,863
	MAX	0,493	0,677	0,745	0,781	0,833	0,858	0,861	0,863	0,864
	AVERAGE	0,378	0,615	0,672	0,734	0,789	0,838	0,854	0,860	0,861

Table 7.63. Macro F-measure results of the proposed combinations in global policy for Reuters dataset

Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (g)		0,014	0,031	0,044	0,090	0,163	0,262	0,370	0,417	0,432
CHI (g)		0,051	0,107	0,163	0,242	0,377	0,439	0,476	0,475	0,482
IG (g)		0,034	0,099	0,140	0,195	0,321	0,392	0,457	0,490	0,476
DF (g)		0,010	0,034	0,058	0,090	0,147	0,243	0,364	0,411	0,438
Acc2 (g)		0,039	0,113	0,145	0,215	0,193	0,484	0,488	0,492	0,490
	MAX	0,051	0,113	0,163	0,242	0,377	0,484	0,488	0,492	0,490
Macro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	0,094	0,172	0,227	0,324	0,414	0,479	0,491	0,500	0,494
	AVERAGE	0,038	0,094	0,132	0,217	0,315	0,428	0,461	0,481	0,482
Rank Combination	MAX	0,173	0,179	0,254	0,335	0,439	0,481	0,484	0,498	0,498
	AVERAGE	0,041	0,090	0,135	0,213	0,304	0,408	0,453	0,475	0,481
Macro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (g)		0,024	0,107	0,156	0,261	0,350	0,463	0,487	0,480	0,479
|C1| tf-idf&IG (g)		0,012	0,076	0,098	0,173	0,220	0,356	0,429	0,466	0,470
|C1| tf-idf&DF (g)		0,009	0,041	0,061	0,108	0,160	0,246	0,377	0,426	0,444
|C1| tf-idf&Acc2 (g)		0,006	0,040	0,079	0,137	0,223	0,357	0,461	0,495	0,487
|C1| CHI&IG (g)		0,052	0,157	0,211	0,328	0,414	0,469	0,477	0,484	0,484
|C1| CHI&DF (g)		0,024	0,107	0,140	0,253	0,362	0,470	0,486	0,490	0,477
|C1| CHI&Acc2 (g)		0,094	0,125	0,201	0,324	0,407	0,482	0,478	0,489	0,491
|C1| IG&DF (g)		0,011	0,072	0,097	0,139	0,208	0,346	0,408	0,461	0,470
|C1| IG&Acc2 (g)		0,056	0,104	0,158	0,258	0,378	0,469	0,481	0,488	0,489
|C1| DF&Acc2 (g)		0,015	0,044	0,068	0,121	0,195	0,349	0,461	0,481	0,484
	MAX	0,094	0,157	0,211	0,328	0,414	0,482	0,487	0,495	0,491
	AVERAGE	0,030	0,087	0,127	0,210	0,292	0,401	0,455	0,476	0,478
Macro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (g)		0,048	0,098	0,131	0,247	0,370	0,447	0,491	0,491	0,488
|C2| tf-idf&IG (g)		0,008	0,094	0,104	0,181	0,258	0,378	0,446	0,463	0,475
|C2| tf-idf&DF (g)		0,010	0,045	0,059	0,095	0,156	0,253	0,366	0,427	0,432
|C2| tf-idf&Acc2 (g)		0,010	0,010	0,011	0,077	0,212	0,467	0,478	0,492	0,491
|C2| CHI&IG (g)		0,038	0,110	0,154	0,292	0,387	0,460	0,482	0,482	0,483
|C2| CHI&DF (g)		0,051	0,100	0,132	0,251	0,384	0,461	0,485	0,500	0,497
|C2| CHI&Acc2 (g)		0,094	0,170	0,231	0,269	0,299	0,486	0,491	0,494	0,490
|C2| IG&DF (g)		0,017	0,094	0,112	0,174	0,270	0,383	0,463	0,492	0,474
|C2| IG&Acc2 (g)		0,080	0,080	0,100	0,158	0,251	0,483	0,488	0,492	0,490
|C2| DF&Acc2 (g)		0,039	0,113	0,002	0,064	0,188	0,483	0,486	0,491	0,490
	MAX	0,094	0,170	0,231	0,292	0,387	0,486	0,491	0,500	0,497
	AVERAGE	0,040	0,091	0,104	0,181	0,278	0,430	0,467	0,482	0,481
Macro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (g)		0,037	0,110	0,131	0,243	0,348	0,464	0,475	0,498	0,498
|C3| tf-idf&IG (g)		0,017	0,061	0,092	0,152	0,219	0,349	0,428	0,466	0,469
|C3| tf-idf&DF (g)		0,014	0,041	0,061	0,109	0,156	0,247	0,380	0,421	0,446
|C3| tf-idf&Acc2 (g)		0,012	0,056	0,102	0,171	0,248	0,375	0,442	0,476	0,494
|C3| CHI&IG (g)		0,056	0,175	0,245	0,328	0,438	0,473	0,479	0,478	0,485
|C3| CHI&DF (g)		0,040	0,106	0,151	0,236	0,352	0,467	0,473	0,484	0,483
|C3| CHI&Acc2 (g)		0,055	0,166	0,282	0,350	0,417	0,485	0,484	0,487	0,494
|C3| IG&DF (g)		0,021	0,050	0,089	0,143	0,196	0,337	0,403	0,467	0,471
|C3| IG&Acc2 (g)		0,064	0,123	0,180	0,297	0,394	0,461	0,485	0,488	0,490
|C3| DF&Acc2 (g)		0,014	0,042	0,078	0,131	0,258	0,357	0,451	0,487	0,484
	MAX	0,064	0,175	0,282	0,350	0,438	0,485	0,485	0,498	0,498
	AVERAGE	0,033	0,093	0,141	0,216	0,303	0,402	0,450	0,475	0,481

Macro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (g)		0,022	0,125	0,182	0,294	0,383	0,477	0,488	0,486	0,477
|C4| tf-idf&IG (g)		0,013	0,077	0,098	0,173	0,228	0,382	0,426	0,465	0,470
|C4| tf-idf&DF (g)		0,009	0,034	0,061	0,108	0,159	0,246	0,373	0,421	0,447
|C4| tf-idf&Acc2 (g)		0,012	0,044	0,082	0,140	0,245	0,427	0,476	0,489	0,490
|C4| CHI&IG (g)		0,054	0,165	0,218	0,343	0,410	0,475	0,477	0,493	0,487
|C4| CHI&DF (g)		0,041	0,133	0,200	0,330	0,403	0,460	0,493	0,491	0,491
|C4| CHI&Acc2 (g)		0,083	0,149	0,201	0,334	0,408	0,485	0,493	0,490	0,490
|C4| IG&DF (g)		0,021	0,072	0,097	0,140	0,223	0,379	0,419	0,469	0,472
|C4| IG&Acc2 (g)		0,047	0,094	0,156	0,260	0,387	0,468	0,483	0,489	0,489
|C4| DF&Acc2 (g)		0,017	0,036	0,068	0,122	0,250	0,420	0,478	0,486	0,490
	MAX	0,083	0,165	0,218	0,343	0,410	0,485	0,493	0,493	0,491
	AVERAGE	0,032	0,093	0,136	0,224	0,310	0,422	0,461	0,478	0,480
Macro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (g)		0,036	0,094	0,146	0,263	0,370	0,468	0,485	0,483	0,488
|C5| tf-idf&IG (g)		0,013	0,075	0,090	0,162	0,219	0,356	0,430	0,466	0,471
|C5| tf-idf&DF (g)		0,014	0,041	0,061	0,109	0,160	0,246	0,380	0,422	0,447
|C5| tf-idf&Acc2 (g)		0,021	0,046	0,082	0,152	0,233	0,369	0,465	0,493	0,489
|C5| CHI&IG (g)		0,047	0,173	0,238	0,329	0,431	0,470	0,478	0,475	0,484
|C5| CHI&DF (g)		0,040	0,096	0,156	0,274	0,387	0,476	0,489	0,482	0,492
|C5| CHI&Acc2 (g)		0,083	0,185	0,233	0,339	0,406	0,478	0,474	0,492	0,490
|C5| IG&DF (g)		0,021	0,061	0,089	0,144	0,209	0,359	0,419	0,468	0,472
|C5| IG&Acc2 (g)		0,052	0,100	0,162	0,263	0,379	0,460	0,482	0,488	0,487
|C5| DF&Acc2 (g)		0,011	0,042	0,068	0,121	0,203	0,354	0,480	0,484	0,488
	MAX	0,083	0,185	0,238	0,339	0,431	0,478	0,489	0,493	0,492
	AVERAGE	0,034	0,091	0,132	0,216	0,300	0,404	0,458	0,475	0,481
Macro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (g)		0,019	0,096	0,150	0,226	0,353	0,468	0,481	0,489	0,478
|C6| tf-idf&IG (g)		0,013	0,075	0,087	0,163	0,221	0,357	0,429	0,466	0,470
|C6| tf-idf&DF (g)		0,009	0,034	0,061	0,108	0,160	0,246	0,377	0,426	0,444
|C6| tf-idf&Acc2 (g)		0,020	0,040	0,080	0,140	0,207	0,342	0,451	0,495	0,490
|C6| CHI&IG (g)		0,045	0,158	0,221	0,326	0,419	0,472	0,477	0,481	0,482
|C6| CHI&DF (g)		0,024	0,080	0,149	0,213	0,341	0,468	0,490	0,490	0,477
|C6| CHI&Acc2 (g)		0,083	0,155	0,210	0,318	0,403	0,478	0,483	0,494	0,490
|C6| IG&DF (g)		0,021	0,061	0,094	0,138	0,198	0,344	0,407	0,461	0,471
|C6| IG&Acc2 (g)		0,055	0,105	0,164	0,260	0,383	0,453	0,482	0,484	0,487
|C6| DF&Acc2 (g)		0,017	0,036	0,066	0,119	0,205	0,330	0,433	0,481	0,480
	MAX	0,083	0,158	0,221	0,326	0,419	0,478	0,490	0,495	0,490
	AVERAGE	0,031	0,084	0,128	0,201	0,289	0,396	0,451	0,477	0,477
Macro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (g)		0,022	0,100	0,140	0,213	0,329	0,447	0,475	0,487	0,502
|C7| tf-idf&IG (g)		0,013	0,062	0,090	0,152	0,216	0,356	0,428	0,466	0,470
|C7| tf-idf&DF (g)		0,013	0,041	0,061	0,109	0,160	0,248	0,377	0,426	0,445
|C7| tf-idf&Acc2 (g)		0,013	0,056	0,082	0,138	0,206	0,322	0,429	0,473	0,492
|C7| CHI&IG (g)		0,048	0,156	0,229	0,320	0,415	0,470	0,477	0,485	0,485
|C7| CHI&DF (g)		0,040	0,081	0,137	0,214	0,338	0,446	0,479	0,483	0,481
|C7| CHI&Acc2 (g)		0,083	0,179	0,231	0,316	0,404	0,477	0,486	0,493	0,479
|C7| IG&DF (g)		0,021	0,049	0,089	0,143	0,188	0,339	0,407	0,460	0,470
|C7| IG&Acc2 (g)		0,060	0,122	0,167	0,262	0,381	0,450	0,479	0,484	0,485
|C7| DF&Acc2 (g)		0,011	0,042	0,068	0,122	0,203	0,322	0,406	0,470	0,482
	MAX	0,083	0,179	0,231	0,320	0,415	0,477	0,486	0,493	0,502
	AVERAGE	0,032	0,089	0,129	0,199	0,284	0,388	0,444	0,473	0,479

On the other hand, among the individual metrics *Acc2* had the highest macro-averaged F-measure 49.2% with 1500 keywords. That performance was improved by several score and rank combinations and among them the most successful result 50.0% was achieved by the score combination of *CHI & DF* with 1500 keywords. Furthermore this performance is also improved, 50.2%, by C7 of *tf-idf & CHI* with only 2000 keywords.

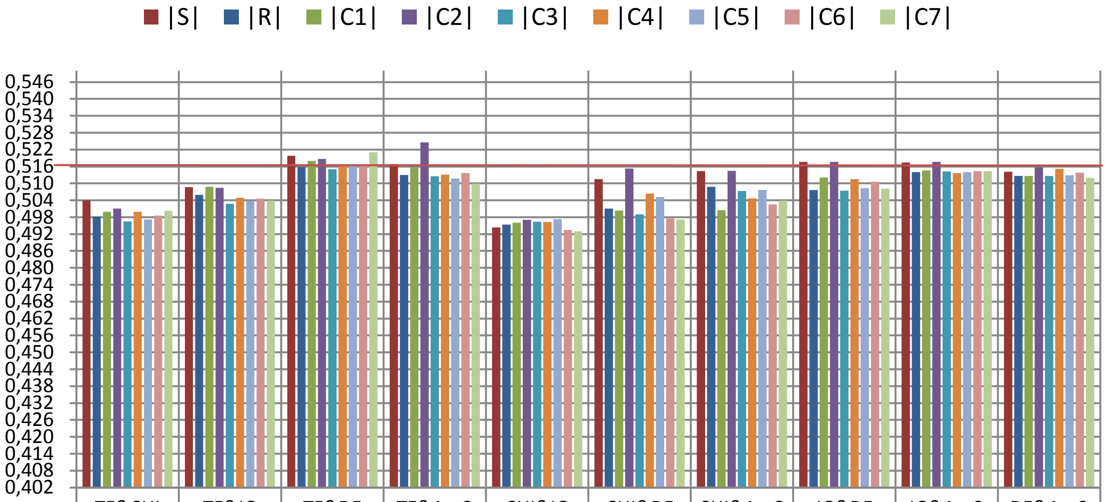


	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,849	0,851	0,850	0,855	0,850	0,854	0,856	0,855	0,856	0,857
|R|	0,843	0,847	0,850	0,850	0,850	0,850	0,854	0,852	0,855	0,852
|C1|	0,845	0,849	0,850	0,852	0,849	0,849	0,852	0,853	0,855	0,854
|C2|	0,849	0,852	0,850	0,857	0,850	0,854	0,856	0,857	0,856	0,857
|C3|	0,843	0,847	0,849	0,850	0,850	0,850	0,854	0,852	0,855	0,852
|C4|	0,847	0,850	0,850	0,853	0,849	0,851	0,853	0,855	0,855	0,855
|C5|	0,846	0,848	0,850	0,851	0,850	0,849	0,854	0,853	0,856	0,853
|C6|	0,845	0,848	0,850	0,851	0,849	0,849	0,853	0,853	0,856	0,854
|C7|	0,844	0,846	0,848	0,850	0,849	0,849	0,853	0,852	0,855	0,853

Figure 7.45. In local policy, averages of the micro-averaged F-measures of all combinations for Reuters dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	0,782	**0,829**	**0,839**	**0,849**	**0,859**	**0,859**	**0,861**	**0,859**	**0,860**
Combination | R |	**0,784**	0,822	**0,839**	0,847	0,851	0,857	0,859	**0,859**	0,858
Combination |C1|	0,785	0,825	0,841	0,849	0,854	0,856	0,859	0,858	0,858
Combination |C2|	0,786	0,824	0,840	0,854	0,859	0,860	0,860	0,861	0,861
Combination |C3|	0,783	0,822	0,841	0,847	0,851	0,857	0,860	0,859	0,858
Combination |C4|	0,785	0,823	0,841	0,850	0,855	0,858	0,860	0,859	0,858
Combination |C5|	0,785	0,827	0,839	0,847	0,853	0,856	0,860	0,860	0,859
Combination |C6|	0,785	0,825	0,840	0,848	0,854	0,856	0,860	0,859	0,859
Combination |C7|	0,785	0,824	0,839	0,847	0,851	0,857	0,859	0,859	0,859

Table 7.64. In local policy, maximum micro-averaged F-measures of all combinations for Reuters dataset



	TF&CHI	TF&IG	TF&DF	TF&Acc2	CHI&IG	CHI&DF	CHI&Acc2	IG&DF	IG&Acc2	DF&Acc2
|S|	0,504	0,509	0,520	0,517	0,494	0,512	0,514	0,518	0,517	0,514
|R|	0,498	0,506	0,516	0,513	0,495	0,501	0,509	0,508	0,514	0,513
|C1|	0,500	0,509	0,518	0,516	0,496	0,500	0,501	0,512	0,515	0,513
|C2|	0,501	0,508	0,519	0,525	0,497	0,515	0,514	0,518	0,518	0,516
|C3|	0,496	0,503	0,515	0,513	0,496	0,499	0,507	0,507	0,514	0,513
|C4|	0,500	0,505	0,517	0,513	0,496	0,506	0,505	0,512	0,514	0,515
|C5|	0,497	0,504	0,517	0,512	0,497	0,505	0,508	0,508	0,514	0,513
|C6|	0,499	0,505	0,516	0,514	0,493	0,498	0,503	0,511	0,514	0,514
|C7|	0,500	0,504	0,521	0,510	0,493	0,497	0,504	0,508	0,514	0,512

Figure 7.46. In local policy, averages of the macro-averaged F-measures of all combinations for Reuters dataset

	10	30	50	100	200	500	1000	1500	2000
Combination | S |	0,500	**0,541**	0,524	**0,536**	**0,533**	**0,512**	**0,499**	**0,496**	**0,492**
Combination | R |	**0,511**	0,533	**0,531**	0,530	0,523	0,509	0,497	0,495	**0,492**
Combination |C1|	0,503	0,533	0,531	0,534	0,528	0,511	0,496	0,495	0,492
Combination |C2|	0,507	0,539	0,534	0,537	0,538	0,513	0,500	0,496	0,492
Combination |C3|	0,503	0,533	0,532	0,529	0,523	0,508	0,497	0,496	0,492
Combination |C4|	0,505	0,531	0,530	0,534	0,527	0,512	0,497	0,496	0,492
Combination |C5|	0,503	0,531	0,532	0,536	0,525	0,510	0,498	0,496	0,492
Combination |C6|	0,504	0,531	0,533	0,534	0,527	0,509	0,497	0,494	0,492
Combination |C7|	0,503	0,532	0,534	0,534	0,523	0,523	0,497	0,497	0,492

Table 7.65. In local policy, maximum macro-averaged F-measures of all combinations for Reuters dataset

When we evaluated the performance of the individual feature selection metrics in local policy in Reuters dataset, it was seen that *DF* and *Acc2* were better than other metrics in the case of local policy. After testing the combination methods, we concluded that the success of the combination of the existing metrics was not as high as we expected when the number of keywords was high but when the number of keyword was low, the score combination outperformed the rank combination and was better than the individual metrics especially in terms of macro-averaged.

Tables 7.66 and 7.67 show the micro- and macro-averaged F-measure results, respectively, for all seven proposed combination methods in local policy for the Reuters dataset. If we want to determine which combinations are better than others, among the 10 possible combinations of two feature selection metrics, the best combinations are *tf-idf & DF, tf-idf & Acc2, CHI & Acc2, IG & DF* and *IG & Acc2.*

As we determined the score combination significantly outperformed the rank combinations. Now we can say that the performance of the score combination is only improved by the C2 with different number of keywords as seen in related tables. Although C2 is more successful than the score combination with many number of keywords, it cannot improve the highest micro- and macro-averaged F-measure values of the score combination.

Table 7.66. Micro F-measure results of the proposed combinations in local policy for Reuters dataset

Micro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (l)		0,776	0,812	0,831	0,835	0,838	0,845	0,853	0,850	0,855
CHI (l)		0,520	**0,823**	**0,840**	0,842	0,839	0,845	0,852	0,855	0,854
IG (l)		**0,777**	0,820	0,838	0,842	0,845	0,850	0,856	0,858	0,856
DF (l)		0,725	0,802	0,820	0,841	0,847	0,854	0,859	**0,859**	**0,859**
Acc2 (l)		0,773	0,811	0,835	**0,846**	**0,855**	**0,860**	**0,862**	**0,859**	**0,859**
	MAX	0,777	0,823	0,840	0,846	0,855	0,860	0,862	0,859	0,859
Micro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	**0,782**	**0,829**	0,839	**0,849**	**0,859**	0,859	0,861	**0,859**	**0,860**
	AVERAGE	0,692	0,820	0,835	0,845	0,850	0,853	0,857	0,857	0,857
Rank Combination	MAX	**0,784**	0,822	0,839	**0,847**	0,851	0,857	0,859	**0,859**	0,858
	AVERAGE	0,747	0,818	0,832	0,839	0,845	0,850	0,856	0,855	0,856
Micro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (l)		**0,785**	0,823	0,838	0,834	0,836	0,846	0,852	0,851	0,853
|C1| tf-idf&IG (l)		**0,783**	0,822	0,831	0,840	0,841	0,849	0,855	0,854	0,854
|C1| tf-idf&DF (l)		0,766	0,813	0,828	0,842	0,847	0,845	0,855	0,855	0,856
|C1| tf-idf&Acc2 (l)		0,768	0,820	0,831	0,841	0,849	0,849	0,856	0,857	0,858
|C1| CHI&IG (l)		0,525	**0,825**	0,836	0,840	0,843	0,846	0,854	0,855	0,857
|C1| CHI&DF (l)		0,521	**0,824**	0,832	0,839	0,840	0,850	0,855	0,855	0,854
|C1| CHI&Acc2 (l)		0,526	**0,824**	**0,841**	0,842	0,847	0,853	0,858	0,856	0,856
|C1| IG&DF (l)		0,769	0,819	0,836	0,844	0,849	0,854	0,857	0,857	0,855
|C1| IG&Acc2 (l)		0,773	0,816	0,837	**0,849**	0,853	0,856	0,859	0,858	0,857
|C1| DF&Acc2 (l)		0,749	0,812	0,830	0,841	0,854	0,856	0,857	0,858	0,858
	MAX	**0,785**	**0,825**	**0,841**	**0,849**	0,854	0,856	0,859	0,858	0,858
	AVERAGE	0,697	0,820	0,834	0,841	0,846	0,850	0,856	0,856	0,856
Micro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (l)		0,776	**0,824**	0,838	0,838	0,841	0,852	0,854	0,854	0,854
|C2| tf-idf&IG (l)		**0,786**	0,814	**0,840**	**0,847**	0,845	0,853	0,856	0,857	0,855
|C2| tf-idf&DF (l)		0,740	0,813	0,826	0,838	0,847	0,850	0,854	0,855	0,854
|C2| tf-idf&Acc2 (l)		0,775	0,815	**0,839**	**0,854**	**0,859**	0,856	0,857	0,858	0,858
|C2| CHI&IG (l)		0,523	**0,823**	0,835	0,841	0,843	0,848	0,854	0,857	0,857
|C2| CHI&DF (l)		0,508	0,818	0,833	**0,847**	0,852	0,853	0,856	0,857	0,857
|C2| CHI&Acc2 (l)		0,523	0,822	**0,839**	**0,851**	0,853	0,857	0,860	**0,859**	0,858
|C2| IG&DF (l)		0,759	0,811	0,836	**0,850**	**0,855**	0,857	0,857	**0,859**	**0,861**
|C2| IG&Acc2 (l)		0,772	0,815	0,834	**0,848**	**0,855**	**0,860**	0,860	0,857	**0,859**
|C2| DF&Acc2 (l)		0,759	0,812	0,833	**0,850**	0,854	**0,860**	0,858	**0,861**	**0,860**
	MAX	**0,786**	**0,824**	**0,840**	**0,854**	**0,859**	**0,860**	0,860	**0,861**	**0,861**
	AVERAGE	0,692	0,817	0,835	0,846	0,850	0,854	0,857	0,857	0,857
Micro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (l)		**0,783**	0,819	0,831	0,827	0,835	0,845	0,850	0,850	0,851
|C3| tf-idf&IG (l)		**0,782**	0,821	0,829	0,832	0,837	0,849	0,854	0,855	0,855
|C3| tf-idf&DF (l)		0,752	0,815	0,830	0,841	0,844	0,845	0,856	0,855	0,855
|C3| tf-idf&Acc2 (l)		0,766	0,817	0,829	0,839	0,849	0,847	0,855	0,855	0,856
|C3| CHI&IG (l)		0,524	0,822	0,836	0,842	0,843	0,848	0,854	0,855	0,857
|C3| CHI&DF (l)		0,771	0,821	0,827	0,839	0,845	0,851	0,858	0,855	0,853
|C3| CHI&Acc2 (l)		0,774	0,819	**0,841**	**0,847**	0,850	0,852	0,860	0,857	0,858
|C3| IG&DF (l)		0,771	0,817	0,829	0,842	0,849	0,851	0,859	0,855	0,857
|C3| IG&Acc2 (l)		0,772	0,817	0,834	**0,847**	0,851	0,857	0,857	**0,859**	0,858
|C3| DF&Acc2 (l)		0,747	0,811	0,829	0,838	0,850	0,854	0,858	**0,859**	0,856
	MAX	**0,783**	0,822	**0,841**	**0,847**	0,851	0,857	0,860	0,859	0,858
	AVERAGE	0,744	0,818	0,832	0,839	0,845	0,850	0,856	0,855	0,856

Micro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (I)		0,785	0,821	0,828	0,834	0,840	0,846	0,854	0,852	0,854
|C4| tf-idf&IG (I)		0,783	0,821	0,830	0,840	0,844	0,851	0,856	0,855	0,854
|C4| tf-idf&DF (I)		0,765	0,813	0,829	0,842	0,845	0,846	0,855	0,855	0,855
|C4| tf-idf&Acc2 (I)		0,768	0,817	0,828	0,846	0,852	0,851	0,857	0,856	0,858
|C4| CHI&IG (I)		0,525	0,822	0,837	0,840	0,845	0,848	0,853	0,855	0,856
|C4| CHI&DF (I)		0,515	0,823	0,828	0,843	0,846	0,850	0,856	0,857	0,855
|C4| CHI&Acc2 (I)		0,526	0,822	0,841	0,844	0,849	0,854	0,860	0,856	0,856
|C4| IG&DF (I)		0,769	0,818	0,837	0,850	0,852	0,856	0,858	0,858	0,858
|C4| IG&Acc2 (I)		0,774	0,816	0,833	0,848	0,853	0,855	0,860	0,858	0,857
|C4| DF&Acc2 (I)		0,748	0,812	0,830	0,843	0,855	0,858	0,858	0,859	0,858
	MAX	0,785	0,823	0,841	0,850	0,855	0,858	0,860	0,859	0,858
	AVERAGE	0,696	0,819	0,832	0,843	0,848	0,851	0,857	0,856	0,856

Micro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (I)		0,785	0,822	0,832	0,838	0,838	0,845	0,852	0,851	0,852
|C5| tf-idf&IG (I)		0,782	0,822	0,829	0,832	0,840	0,848	0,856	0,855	0,855
|C5| tf-idf&DF (I)		0,763	0,814	0,830	0,842	0,845	0,845	0,855	0,855	0,855
|C5| tf-idf&Acc2 (I)		0,766	0,817	0,828	0,839	0,849	0,849	0,856	0,856	0,858
|C5| CHI&IG (I)		0,525	0,822	0,837	0,841	0,844	0,848	0,854	0,855	0,856
|C5| CHI&DF (I)		0,516	0,827	0,830	0,841	0,841	0,850	0,856	0,854	0,855
|C5| CHI&Acc2 (I)		0,523	0,822	0,839	0,846	0,852	0,853	0,858	0,856	0,857
|C5| IG&DF (I)		0,772	0,816	0,833	0,843	0,850	0,854	0,857	0,855	0,856
|C5| IG&Acc2 (I)		0,773	0,816	0,834	0,847	0,853	0,856	0,860	0,858	0,859
|C5| DF&Acc2 (I)		0,747	0,812	0,830	0,839	0,851	0,855	0,858	0,860	0,857
	MAX	0,785	0,827	0,839	0,847	0,853	0,856	0,860	0,860	0,859
	AVERAGE	0,695	0,819	0,832	0,841	0,846	0,850	0,856	0,855	0,856

Micro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (I)		0,785	0,821	0,828	0,831	0,838	0,844	0,852	0,852	0,853
|C6| tf-idf&IG (I)		0,783	0,822	0,830	0,837	0,840	0,848	0,854	0,854	0,853
|C6| tf-idf&DF (I)		0,766	0,812	0,830	0,844	0,845	0,845	0,856	0,856	0,855
|C6| tf-idf&Acc2 (I)		0,769	0,820	0,829	0,840	0,849	0,848	0,856	0,857	0,858
|C6| CHI&IG (I)		0,526	0,820	0,836	0,840	0,844	0,846	0,853	0,855	0,856
|C6| CHI&DF (I)		0,517	0,825	0,831	0,839	0,839	0,851	0,856	0,854	0,854
|C6| CHI&Acc2 (I)		0,526	0,823	0,840	0,842	0,847	0,853	0,860	0,856	0,857
|C6| IG&DF (I)		0,772	0,819	0,834	0,843	0,849	0,854	0,858	0,857	0,854
|C6| IG&Acc2 (I)		0,773	0,816	0,837	0,848	0,853	0,856	0,860	0,858	0,859
|C6| DF&Acc2 (I)		0,748	0,813	0,831	0,842	0,854	0,855	0,857	0,859	0,857
	MAX	0,785	0,825	0,840	0,848	0,854	0,856	0,860	0,859	0,859
	AVERAGE	0,696	0,819	0,832	0,841	0,846	0,850	0,856	0,856	0,856

Micro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (I)		0,785	0,821	0,828	0,828	0,836	0,845	0,852	0,854	0,851
|C7| tf-idf&IG (I)		0,782	0,822	0,828	0,830	0,838	0,847	0,855	0,853	0,854
|C7| tf-idf&DF (I)		0,763	0,815	0,830	0,844	0,844	0,844	0,846	0,856	0,855
|C7| tf-idf&Acc2 (I)		0,766	0,816	0,828	0,839	0,846	0,848	0,854	0,856	0,857
|C7| CHI&IG (I)		0,525	0,820	0,837	0,842	0,842	0,847	0,855	0,855	0,856
|C7| CHI&DF (I)		0,774	0,824	0,827	0,836	0,841	0,848	0,855	0,857	0,854
|C7| CHI&Acc2 (I)		0,783	0,822	0,839	0,842	0,850	0,853	0,858	0,857	0,859
|C7| IG&DF (I)		0,773	0,817	0,831	0,843	0,847	0,853	0,859	0,857	0,856
|C7| IG&Acc2 (I)		0,773	0,815	0,835	0,847	0,851	0,857	0,858	0,859	0,859
|C7| DF&Acc2 (I)		0,747	0,813	0,830	0,840	0,851	0,853	0,856	0,859	0,859
	MAX	0,785	0,824	0,839	0,847	0,851	0,857	0,859	0,859	0,859
	AVERAGE	0,747	0,819	0,831	0,839	0,845	0,849	0,855	0,856	0,856

Table 7.67. Macro F-measure results of the proposed combinations in local policy for Reuters dataset

Macro-F		10	30	50	100	200	500	1000	1500	2000
tf-idf (l)		**0,494**	0,512	0,519	0,508	0,514	0,493	0,495	0,491	0,492
CHI (l)		0,466	0,491	0,493	0,500	0,488	0,493	0,493	0,494	0,491
IG (l)		**0,494**	**0,530**	0,512	0,517	0,496	0,495	0,493	**0,496**	0,490
DF (l)		0,463	0,497	0,515	**0,539**	**0,532**	0,511	**0,500**	0,491	**0,493**
Acc2 (l)		0,492	0,525	**0,524**	0,527	0,515	**0,513**	**0,500**	0,492	0,489
	MAX	0,494	0,530	0,524	0,539	0,532	0,513	0,500	0,496	0,493
Macro-F		10	30	50	100	200	500	1000	1500	2000
Score Combination	MAX	**0,500**	**0,541**	**0,524**	0,536	**0,533**	0,512	0,499	**0,496**	0,492
	AVERAGE	0,490	0,528	0,515	0,523	0,513	0,503	0,496	0,491	0,489
Rank Combination	MAX	**0,511**	**0,533**	**0,531**	0,530	0,523	0,509	0,497	0,495	0,492
	AVERAGE	0,490	0,519	0,515	0,516	0,504	0,500	0,495	0,493	0,491
Macro-F	Combination 1	10	30	50	100	200	500	1000	1500	2000
|C1| tf-idf&CHI (l)		0,489	0,515	0,510	0,504	0,490	0,491	0,492	0,489	0,490
|C1| tf-idf&IG (l)		**0,503**	0,527	0,513	0,517	0,499	0,493	0,493	0,489	0,488
|C1| tf-idf&DF (l)		**0,496**	0,524	**0,531**	0,534	0,528	0,495	0,495	0,490	0,492
|C1| tf-idf&Acc2 (l)		0,492	**0,533**	**0,526**	0,529	0,517	0,496	0,494	0,490	0,489
|C1| CHI&IG (l)		0,490	0,510	0,495	0,504	0,485	0,492	0,493	0,494	0,491
|C1| CHI&DF (l)		0,492	0,5146	0,508	0,504	0,489	0,495	0,494	0,490	0,490
|C1| CHI&Acc2 (l)		**0,494**	0,513	0,503	0,502	0,489	0,502	0,496	0,492	0,491
|C1| IG&DF (l)		**0,497**	**0,530**	0,518	0,519	0,509	0,498	0,495	0,490	0,488
|C1| IG&Acc2 (l)		**0,499**	0,527	0,519	0,526	0,510	0,507	0,496	0,495	0,490
|C1| DF&Acc2 (l)		0,472	0,530	0,516	0,530	0,516	0,511	0,494	0,491	0,489
	MAX	0,503	0,533	0,531	0,534	0,528	0,511	0,496	0,495	0,492
	AVERAGE	0,493	0,522	0,514	0,517	0,503	0,498	0,494	0,491	0,490
Macro-F	Combination 2	10	30	50	100	200	500	1000	1500	2000
|C2| tf-idf&CHI (l)		0,483	0,515	0,510	0,506	0,498	0,494	0,495	0,490	0,488
|C2| tf-idf&IG (l)		**0,504**	0,515	0,523	0,512	0,499	0,499	0,494	0,490	0,488
|C2| tf-idf&DF (l)		**0,496**	**0,530**	0,522	0,530	**0,538**	0,496	0,495	0,490	**0,492**
|C2| tf-idf&Acc2 (l)		**0,507**	**0,539**	**0,534**	0,533	0,527	0,507	0,495	0,491	0,489
|C2| CHI&IG (l)		0,476	0,517	0,497	0,508	0,490	0,495	0,492	**0,496**	0,490
|C2| CHI&DF (l)		0,477	**0,539**	0,518	0,529	0,521	0,509	0,495	0,490	0,489
|C2| CHI&Acc2 (l)		0,478	0,527	**0,527**	0,527	0,517	0,510	**0,500**	0,492	0,489
|C2| IG&DF (l)		0,490	**0,531**	0,517	**0,537**	0,520	0,511	0,496	0,490	0,489
|C2| IG&Acc2 (l)		**0,500**	**0,533**	0,515	0,528	0,518	**0,513**	0,499	0,490	0,489
|C2| DF&Acc2 (l)		0,477	**0,536**	0,515	0,533	0,520	**0,513**	0,495	0,491	0,489
	MAX	0,507	0,539	**0,534**	**0,537**	**0,538**	**0,513**	**0,500**	**0,496**	**0,492**
	AVERAGE	0,489	0,528	0,518	0,524	0,515	0,505	0,496	0,491	0,489
Macro-F	Combination 3	10	30	50	100	200	500	1000	1500	2000
|C3| tf-idf&CHI (l)		0,486	0,488	0,516	0,500	0,492	0,497	0,493	0,494	0,490
|C3| tf-idf&IG (l)		**0,503**	0,502	0,515	0,504	0,494	0,498	0,495	0,496	0,492
|C3| tf-idf&DF (l)		0,491	0,522	**0,532**	0,529	0,523	0,493	0,495	0,490	0,492
|C3| tf-idf&Acc2 (l)		0,482	**0,531**	0,520	0,524	0,518	0,501	0,494	0,492	0,489
|C3| CHI&IG (l)		0,487	0,510	0,492	0,501	0,490	0,497	0,494	0,495	0,491
|C3| CHI&DF (l)		0,482	0,505	0,499	0,509	0,501	0,498	0,497	0,493	0,491
|C3| CHI&Acc2 (l)		0,485	0,521	0,511	0,522	0,503	0,502	0,497	0,493	0,491
|C3| IG&DF (l)		0,486	0,528	0,509	0,515	0,503	0,503	0,496	0,493	0,490
|C3| IG&Acc2 (l)		0,491	**0,533**	0,516	0,523	0,516	0,507	0,495	0,495	**0,492**
|C3| DF&Acc2 (l)		0,476	0,529	0,520	0,527	0,515	0,508	0,495	0,492	0,489
	MAX	0,503	0,533	**0,532**	0,529	0,523	0,508	0,497	0,496	**0,492**
	AVERAGE	0,487	0,517	0,513	0,515	0,505	0,500	0,495	0,493	0,491

Macro-F	Combination 4	10	30	50	100	200	500	1000	1500	2000
|C4| tf-idf&CHI (I)		0,490	0,512	0,503	0,504	0,498	0,492	0,493	0,488	0,489
|C4| tf-idf&IG (I)		0,505	0,508	0,510	0,513	0,498	0,496	0,495	0,491	0,487
|C4| tf-idf&DF (I)		0,495	0,522	0,530	0,534	0,527	0,494	0,496	0,490	0,492
|C4| tf-idf&Acc2 (I)		0,483	0,530	0,516	0,530	0,516	0,504	0,494	0,490	0,488
|C4| CHI&IG (I)		0,484	0,509	0,500	0,502	0,486	0,497	0,492	0,495	0,491
|C4| CHI&DF (I)		0,481	0,531	0,503	0,519	0,508	0,497	0,494	0,490	0,489
|C4| CHI&Acc2 (I)		0,484	0,521	0,507	0,513	0,503	0,501	0,497	0,492	0,491
|C4| IG&DF (I)		0,492	0,526	0,513	0,521	0,511	0,506	0,496	0,492	0,490
|C4| IG&Acc2 (I)		0,498	0,527	0,513	0,527	0,511	0,506	0,496	0,496	0,491
|C4| DF&Acc2 (I)		0,479	0,530	0,515	0,533	0,522	0,512	0,495	0,491	0,489
	MAX	0,505	0,531	0,530	0,534	0,527	0,512	0,497	0,496	0,492
	AVERAGE	0,489	0,522	0,511	0,519	0,508	0,501	0,495	0,491	0,490

Macro-F	Combination 5	10	30	50	100	200	500	1000	1500	2000
|C5| tf-idf&CHI (I)		0,491	0,508	0,504	0,494	0,494	0,492	0,494	0,489	0,489
|C5| tf-idf&IG (I)		0,503	0,511	0,507	0,505	0,498	0,499	0,496	0,496	0,489
|C5| tf-idf&DF (I)		0,493	0,522	0,532	0,536	0,525	0,492	0,496	0,490	0,492
|C5| tf-idf&Acc2 (I)		0,481	0,531	0,516	0,524	0,517	0,502	0,495	0,491	0,489
|C5| CHI&IG (I)		0,485	0,511	0,502	0,503	0,485	0,497	0,493	0,494	0,491
|C5| CHI&DF (I)		0,486	0,528	0,506	0,513	0,498	0,500	0,496	0,491	0,491
|C5| CHI&Acc2 (I)		0,486	0,525	0,506	0,524	0,503	0,502	0,497	0,491	0,490
|C5| IG&DF (I)		0,490	0,519	0,510	0,519	0,508	0,504	0,497	0,493	0,491
|C5| IG&Acc2 (I)		0,495	0,528	0,514	0,525	0,516	0,507	0,498	0,494	0,492
|C5| DF&Acc2 (I)		0,477	0,528	0,521	0,527	0,514	0,510	0,495	0,492	0,489
	MAX	0,503	0,531	0,532	0,536	0,525	0,510	0,498	0,496	0,492
	AVERAGE	0,489	0,521	0,512	0,517	0,506	0,500	0,496	0,492	0,490

Macro-F	Combination 6	10	30	50	100	200	500	1000	1500	2000
|C6| tf-idf&CHI (I)		0,497	0,504	0,508	0,502	0,491	0,489	0,492	0,491	0,488
|C6| tf-idf&IG (I)		0,504	0,515	0,511	0,507	0,497	0,493	0,493	0,491	0,488
|C6| tf-idf&DF (I)		0,493	0,516	0,533	0,534	0,527	0,495	0,496	0,490	0,492
|C6| tf-idf&Acc2 (I)		0,487	0,531	0,526	0,526	0,516	0,497	0,494	0,489	0,489
|C6| CHI&IG (I)		0,488	0,493	0,497	0,502	0,487	0,494	0,493	0,494	0,491
|C6| CHI&DF (I)		0,495	0,501	0,503	0,505	0,486	0,496	0,495	0,490	0,488
|C6| CHI&Acc2 (I)		0,496	0,522	0,502	0,502	0,490	0,503	0,497	0,493	0,491
|C6| IG&DF (I)		0,495	0,531	0,513	0,518	0,509	0,498	0,495	0,492	0,488
|C6| IG&Acc2 (I)		0,497	0,529	0,518	0,525	0,510	0,507	0,496	0,493	0,492
|C6| DF&Acc2 (I)		0,479	0,530	0,520	0,530	0,516	0,509	0,494	0,491	0,489
	MAX	0,504	0,531	0,533	0,534	0,527	0,509	0,497	0,494	0,492
	AVERAGE	0,493	0,517	0,513	0,515	0,503	0,498	0,494	0,491	0,490

Macro-F	Combination 7	10	30	50	100	200	500	1000	1500	2000
|C7| tf-idf&CHI (I)		0,501	0,510	0,513	0,499	0,488	0,491	0,492	0,491	0,487
|C7| tf-idf&IG (I)		0,503	0,510	0,518	0,503	0,493	0,497	0,494	0,489	0,488
|C7| tf-idf&DF (I)		0,490	0,523	0,534	0,534	0,523	0,523	0,494	0,490	0,492
|C7| tf-idf&Acc2 (I)		0,481	0,519	0,527	0,523	0,514	0,496	0,493	0,489	0,489
|C7| CHI&IG (I)		0,485	0,493	0,496	0,504	0,485	0,495	0,493	0,494	0,491
|C7| CHI&DF (I)		0,495	0,506	0,498	0,506	0,484	0,494	0,494	0,492	0,490
|C7| CHI&Acc2 (I)		0,497	0,527	0,498	0,504	0,494	0,503	0,497	0,494	0,490
|C7| IG&DF (I)		0,491	0,527	0,510	0,518	0,500	0,502	0,497	0,490	0,491
|C7| IG&Acc2 (I)		0,497	0,532	0,515	0,523	0,511	0,508	0,495	0,497	0,492
|C7| DF&Acc2 (I)		0,477	0,528	0,521	0,528	0,515	0,504	0,494	0,491	0,489
	MAX	0,503	0,532	0,534	0,534	0,523	0,523	0,497	0,497	0,492
	AVERAGE	0,492	0,518	0,513	0,514	0,501	0,501	0,494	0,492	0,490

7.4. Significance Test

Besides the micro and macro-averaged F-measures, the micro sign test (s-test) is also implemented to test the robustness of the results and measure the significance improvements for the proposed methods. In micro sign test, we compare the best individual metric with the proposed combinations that are better than it in micro-averaged F-measure for each dataset except Classic3. Since the results are very close to each other and almost all methods improve the best performance of the individual metric. Tables 7.69, 7.70, 7.71 and 7.72 demonstrate the comparison of the methods for other datasets.

The highest confidence levels for the proposed methods are achieved on the Reuters dataset. In Reuters, *tf-idf & Acc2* 2-combination is better than the highest individual metric *IG*. *tf-idf & Acc2* with rank combination achieves 98.51%, *tf-idf & Acc2* with C3 combination achieves 95.79% and *tf-idf & Acc2* with C7 combination achieves 95.79% confidence levels in global policy.

[Hitech]					
Method	keyword	Combination	keyword	z- value	conf
CHI (g)	1500	|S| IG&Acc2 (g)	2000	0,19	57.37%
CHI (g)	1500	|R| IG&Acc2 (g)	2000	0,55	70.81%
CHI (g)	1500	|C1| IG&Acc2 (g)	2000	0,47	68.25%
CHI (g)	1500	|C2| CHI&Acc2 (g)	2000	0,28	61.21%
CHI (g)	1500	|C3| IG&Acc2 (g)	2000	0,55	70.81%
CHI (g)	1500	|C4| tf-idf&CHI (g)	1500	0,45	67.53%
CHI (g)	1500	|C4| CHI&DF (g)	2000	0,38	64.73%
CHI (g)	1500	|C4| IG&Acc2 (g)	2000	0,75	77.32%
CHI (g)	1500	|C5| tf-idf&IG (g)	2000	0,09	53.62%
CHI (g)	1500	|C5| CHI&DF (g)	2000	0,28	61.21%
CHI (g)	1500	|C5| IG&Acc2 (g)	2000	0,55	70.96%
CHI (g)	1500	|C6| IG&Acc2 (g)	2000	0,56	71.13%
CHI (g)	1500	|C7| IG&Acc2 (g)	2000	0,54	70.65%
DF (l)	1500	|S| IG&Acc2 (l)	1000	0,64	73.95%
DF (l)	1500	|C1| DF&Acc2 (l)	1500	0,37	64.30%
DF (l)	1500	|C2| CHI&DF (l)	1500	0,65	74.25%
DF (l)	1500	|C2| IG&Acc2 (l)	1000	1,03	84.75%
DF (l)	1500	|C2| DF&Acc2 (l)	1500	1,48	93.10%
DF (l)	1500	|C4| tf-idf&Acc2 (l)	500	0,72	76.56%
DF (l)	1500	|C4| IG&DF (l)	1000	0,57	71.46%
DF (l)	1500	|C5| IG&DF (l)	500	-0,08	46.78%
DF (l)	1500	|C6| DF&Acc2 (l)	1500	0,46	67.68%
DF (l)	1500	|C7| IG&DF (l)	500	-0,24	40.48%

Table 7.68. Statistical comparison of the methods for Hitech Dataset

On the other hand, although the other proposed methods cannot achieve over %95 confidence level, they are also improve the performance of the existing metrics with a high confidence level such as DF & Acc2 with C2 combination in Hitech is better than the best metric DF with 93.10%, IG & Acc2 with C6 combination in Reuters is better than the best metric IG with 86.31% or tf-idf & CHI with C5 combination in LA1 is better than the best metric CHI with 86.24% confidence level.

[LA1]					
Method	keyword	Combination	keyword	z- value	conf
CHI (g)	2000	|S| tf-idf&CHI (g)	2000	0,47	68.10%
CHI (g)	2000	|R| tf-idf&CHI (g)	2000	0,33	62.76%
CHI (g)	2000	|C3| tf-idf&CHI (g)	2000	0,11	54.27%
CHI (g)	2000	|C5| tf-idf&CHI (g)	2000	1,09	86.24%
CHI (g)	2000	|C7| tf-idf&CHI (g)	2000	0,55	70.84%
Acc2 (l)	2000	|S| CHI&IG (l)	1000	-0,09	46.41%
Acc2 (l)	2000	|S| CHI&IG (l)	2000	0,11	54.27%
Acc2 (l)	2000	|S| CHI&Acc2 (l)	1500	0,49	68.88%
Acc2 (l)	2000	|S| CHI&Acc2 (l)	2000	0,63	73.56%
Acc2 (l)	2000	|S| IG&Acc2 (l)	2000	0,64	73.90%
Acc2 (l)	2000	|R| CHI&IG (l)	2000	0,31	62.09%
Acc2 (l)	2000	|R| CHI&Acc2 (l)	1500	0,21	58.35%
Acc2 (l)	2000	|R| CHI&Acc2 (l)	2000	0,00	50.00%
Acc2 (l)	2000	|R| IG&Acc2 (l)	2000	0,35	63.73%
Acc2 (l)	2000	|C1| CHI&Acc2 (l)	1500	0,45	67.26%
Acc2 (l)	2000	|C1| IG&Acc2 (l)	2000	0,50	69.15%
Acc2 (l)	2000	|C2| tf-idf&Acc2 (l)	1000	0,21	58.17%
Acc2 (l)	2000	|C2| CHI&Acc2 (l)	1500	0,74	76.99%
Acc2 (l)	2000	|C2| CHI&Acc2 (l)	2000	0,82	79.29%
Acc2 (l)	2000	|C3| CHI&IG (l)	2000	0,41	65.85%
Acc2 (l)	2000	|C3| CHI&Acc2 (l)	1500	0,11	54.22%
Acc2 (l)	2000	|C3| CHI&Acc2 (l)	2000	0,12	54.86%
Acc2 (l)	2000	|C3| IG&DF (l)	1500	-0,23	41.04%
Acc2 (l)	2000	|C3| IG&Acc2 (l)	2000	0,59	72.08%
Acc2 (l)	2000	|C4| tf-idf&IG (l)	2000	0,21	58.26%
Acc2 (l)	2000	|C4| CHI&IG (l)	2000	0,10	54,00%
Acc2 (l)	2000	|C4| CHI&DF (l)	2000	0,12	54.94%
Acc2 (l)	2000	|C4| CHI&Acc2 (l)	1500	0,80	78.75%
Acc2 (l)	2000	|C4| CHI&Acc2 (l)	2000	0,12	54.86%
Acc2 (l)	2000	|C4| IG&DF (l)	1500	0,36	64.10%
Acc2 (l)	2000	|C4| IG&DF (l)	2000	0,76	77.70%
Acc2 (l)	2000	|C4| IG&Acc2 (l)	2000	0,61	72.93%
Acc2 (l)	2000	|C5| CHI&IG (l)	1000	0,18	57.13%
Acc2 (l)	2000	|C5| CHI&Acc2 (l)	1500	0,32	62.61%
Acc2 (l)	2000	|C5| CHI&Acc2 (l)	2000	0,24	59.58%
Acc2 (l)	2000	|C5| IG&DF (l)	1500	0,12	54.66%
Acc2 (l)	2000	|C5| IG&Acc2 (l)	2000	0,36	63.91%
Acc2 (l)	2000	|C6| CHI&IG (l)	2000	0,41	66.00%
Acc2 (l)	2000	|C6| IG&DF (l)	2000	0,13	55.36%
Acc2 (l)	2000	|C7| CHI&IG (l)	2000	0,42	66.34%
Acc2 (l)	2000	|C7| CHI&Acc2 (l)	1500	0,10	54.17%
Acc2 (l)	2000	|C7| IG&Acc2 (l)	2000	0,83	79.69%

Table 7.69. Statistical comparison of the methods for LA1 Dataset

[Wap]					
Method	keyword	Combination	keyword	z- value	conf
DF (g)	2000	|S| tf-idf&Acc2 (g)	1500	-0,17	43.09%
DF (g)	2000	|S| IG&DF (g)	1500	0,60	72.57%
DF (g)	2000	|S| IG&Acc2 (g)	1000	-0,28	38.86%
DF (g)	2000	|S| IG&Acc2 (g)	2000	0,00	50.00%
DF (g)	2000	|S| DF&Acc2 (g)	2000	0,18	57.13%
DF (g)	2000	|R| CHI&Acc2 (g)	2000	-0,15	43.94%
DF (g)	2000	|R| IG&Acc2 (g)	1500	0,39	65.26%
DF (g)	2000	|R| DF&Acc2 (g)	1000	0,47	68.03%
DF (g)	2000	|R| DF&Acc2 (g)	1500	0,17	56.91%
DF (g)	2000	|R| DF&Acc2 (g)	2000	0,23	59.07%
DF (g)	2000	|C1| tf-idf&CHI (g)	1000	-0,16	43.64%
DF (g)	2000	|C1| tf-idf&Acc2 (g)	1000	0,00	50.00%
DF (g)	2000	|C1| CHI&DF (g)	1000	0,00	50.00%
DF (g)	2000	|C1| CHI&Acc2 (g)	2000	-0,17	43.29%
DF (g)	2000	|C1| IG&Acc2 (g)	2000	0,00	50.00%
DF (g)	2000	|C1| DF&Acc2 (g)	1500	0,37	64.25%
DF (g)	2000	|C1| DF&Acc2 (g)	2000	0,50	69.15%
DF (g)	2000	|C2| tf-idf&Acc2 (g)	1500	0,34	63.42%
DF (g)	2000	|C2| CHI&IG (g)	1500	0,00	50.00%
DF (g)	2000	|C2| CHI&Acc2 (g)	2000	-0,16	43.64%
DF (g)	2000	|C2| IG&DF (g)	1500	0,19	57.63%
DF (g)	2000	|C2| IG&DF (g)	2000	0,69	75.44%
DF (g)	2000	|C2| IG&Acc2 (g)	2000	0,37	64.25%
DF (g)	2000	|C3| CHI&Acc2 (g)	2000	0,00	50.00%
DF (g)	2000	|C3| IG&Acc2 (g)	1500	0,38	64.73%
DF (g)	2000	|C3| DF&Acc2 (g)	1500	0,34	63.42%
DF (g)	2000	|C4| tf-idf&Acc2 (g)	1000	0,30	61.85%
DF (g)	2000	|C4| tf-idf&Acc2 (g)	1500	0,00	50.00%
DF (g)	2000	|C4| CHI&Acc2 (g)	1500	-0,15	43.94%
DF (g)	2000	|C4| IG&DF (g)	1500	0,20	57.93%
DF (g)	2000	|C4| IG&Acc2 (g)	1500	0,17	56.91%
DF (g)	2000	|C4| IG&Acc2 (g)	2000	0,56	71.13%
DF (g)	2000	|C4| DF&Acc2 (g)	1500	0,17	56.71%
DF (g)	2000	|C5| tf-idf&DF (g)	1000	0,00	50.00%
DF (g)	2000	|C5| tf-idf&Acc2 (g)	1000	0,93	82.27%
DF (g)	2000	|C5| IG&DF (g)	1500	0,43	66.51%
DF (g)	2000	|C5| IG&Acc2 (g)	1500	0,78	78.36%
DF (g)	2000	|C5| IG&Acc2 (g)	2000	0,20	57.93%
DF (g)	2000	|C5| DF&Acc2 (g)	1000	0,15	55.93%
DF (g)	2000	|C5| DF&Acc2 (g)	1500	0,56	71.13%
DF (g)	2000	|C6| tf-idf&CHI (g)	1000	0,00	50.00%
DF (g)	2000	|C6| tf-idf&Acc2 (g)	1000	0,62	73.15%
DF (g)	2000	|C6| tf-idf&Acc2 (g)	1500	0,00	50.00%
DF (g)	2000	|C6| IG&Acc2 (g)	1500	0,56	71.13%
DF (g)	2000	|C6| IG&Acc2 (g)	2000	0,20	57.93%
DF (g)	2000	|C6| DF&Acc2 (g)	1500	0,76	77.52%
DF (g)	2000	|C6| DF&Acc2 (g)	2000	0,50	69.15%
DF (g)	2000	|C7| tf-idf&DF (g)	1500	0,45	67.26%
DF (g)	2000	|C7| tf-idf&Acc2 (g)	1000	0,62	73.15%
DF (g)	2000	|C7| CHI&DF (g)	500	0,25	59.87%
DF (g)	2000	|C7| IG&Acc2 (g)	1500	0,18	57.13%
DF (g)	2000	|C7| DF&Acc2 (g)	1500	0,19	57.63%
DF (g)	2000	|C7| DF&Acc2 (g)	2000	0,00	50.00%
DF (l)	200	|S| DF&Acc2 (l)	200	0,00	50.00%

Table 7.70. Statistical comparison of the methods for Wap Dataset

[Reuters]					
Method	keyword	Combination	keyword	z- value	conf
IG(g)	1500	|S| tf-idf&IG (g)	1500	0,63	73.65%
IG(g)	1500	|S| tf-idf&Acc2 (g)	2000	0,49	68.79%
IG(g)	1500	|S| CHI&DF (g)	2000	0,07	52.83%
IG(g)	1500	|R| tf-idf&CHI (g)	1000	0,00	50.00%
IG(g)	1500	|R| tf-idf&CHI (g)	1500	0,00	50.00%
IG(g)	1500	|R| tf-idf&Acc2 (g)	1500	0,50	69.27%
IG(g)	1500	|R| tf-idf&Acc2 (g)	2000	2,17	98.51%
IG(g)	1500	|R| CHI&Acc2 (g)	2000	0,21	58.14%
IG(g)	1500	|C1| tf-idf&Acc2 (g)	1500	0,74	76.95%
IG(g)	1500	|C1| tf-idf&Acc2 (g)	2000	0,86	80.57%
IG(g)	1500	|C1| IG&Acc2 (g)	2000	0,68	75.07%
IG(g)	1500	|C2| IG&DF (g)	1500	0,56	71.13%
IG(g)	1500	|C3| tf-idf&CHI (g)	1500	0,00	50.00%
IG(g)	1500	|C3| tf-idf&Acc2 (g)	1500	0,25	59.84%
IG(g)	1500	|C3| tf-idf&Acc2 (g)	2000	1,73	95.79%
IG(g)	1500	|C4| tf-idf&IG (g)	1500	0,78	78.16%
IG(g)	1500	|C4| tf-idf&Acc2 (g)	2000	0,79	78.42%
IG(g)	1500	|C4| DF&Acc2 (g)	2000	1,04	85.00%
IG(g)	1500	|C5| tf-idf&Acc2 (g)	1500	0,74	76.95%
IG(g)	1500	|C5| tf-idf&Acc2 (g)	2000	0,47	67.95%
IG(g)	1500	|C6| tf-idf&Acc2 (g)	1500	1,16	87.67%
IG(g)	1500	|C6| tf-idf&Acc2 (g)	2000	0,97	83.43%
IG(g)	1500	|C6| IG&DF (g)	2000	0,66	74.62%
IG(g)	1500	|C6| IG&Acc2 (g)	2000	1,09	86.31%
IG(g)	1500	|C7| tf-idf&Acc2 (g)	2000	1,45	92.63%
IG(g)	1500	|C7| CHI&Acc2 (g)	1500	0,36	64.20%
IG(g)	1500	|C7| DF&Acc2 (g)	2000	1,27	89.77%

Table 7.71. Statistical comparison of the methods for Reuters Dataset

Among the datasets, the lowest confidence levels are obtained in the Wap datasets, even the performance of the proposed methods are seen more successful in the case of micro-averaged F-measure.

These significance tests are also showed that the proposed methods can improve the success of the individual metrics in classification.

7.5. Summary of the Results

The experiments show that substantial improvements can be achieved in text categorization by combining feature selection methods. Although many feature selection methods exist in text categorization, it is hard to state one is in general superior to others since the success of the methods depends on various variables. It is more likely that combining different feature selection methods obtains more effective performance in text categorization.

Comparing the performance of the individual methods with the performance of the combination methods demonstrate that combining two feature selection methods can significantly improve the performance of the individual methods in all dataset. In addition, in general rank combination achieves better performance than score combination in the case of global policy but score combination significantly achieves better performance in the case of local policy.

One of the explicit observations is that there is a marked difference between micro- and macro-averaged F-measure values for skew and especially highly skew datasets due to the the poor performances over rare categories. The improvement of the macro-average F-measures is clearer than the improvement of the micro-averaged F-measures for the combination methods. From this observation we can say that the combination methods more apparently improve the performance of the classifier on rare categories in all dataset.

One of the important results of these experiments is that all combination methods improve the highest F-measure values of the individual metrics with almost all number of keywords from 10 to 2000. Especially, success of combining feature selection methods is more apparent when the keyword number is low. It is approaved that more successful performances are achieved by less number of keywords compared to individual metrics.

For individual metrics, the global policy is more successful than the local policy when the keyword number is high but it is outperformed by the local policy when the keyword number is low this assumpsion is still valied after combining feature selection methods.

Another notable conclusion for the study is that the success of the combination of the feature selection methods depends on the success of the individual feature selection methods.

In this study, we evaluate the performance of all possible binary combinations of the five popular feature selection methods. Among the possible combination pairs, the most successful results are achieved by *tf-idf & CHI, tf-idf & Acc2, CHI & IG, CHI & Acc2, IG & Acc2* 2-combinations in general but it should be noted that the successful combination pairs are changed according to the properties of the datasets.

Among the proposed combination methods, C2, C5, C6 and C7 show better performance than score and rank combinations which are better than the individual feature selection methods in many cases. In addition, one of the striking results among the all combinations C3 method significantly improves the success of the individual method in Reuters dataset.

The most successful combinations with the high levels confidence are obtained in the Reuters dataset. This is also a promosing result for the classification performance since the Reuters dataset is one of the most suited real-world document collections with the properties of skewness, sample size and the number of classes.

8. CONCLUSIONS AND FUTURE WORK

In this study, we present a comprehensive analysis of the comparison between the feature selection methods and their varied binary combinations for text categorization with a comparative discussion. Firstly we analyze the performance of five common feature selection methods on five standard datasets with varied skewness in both global and local policies, and then evaluate the performance of all possible binary score and rank combinations of these five feature selection methods with the same experimental settings in order to determine the most appropriate features for classification. Comparing the performance of the individual methods with the performance of the combination methods shows that combining two feature selection methods can significantly improve the performance of the individual methods. In addition, rank combination achieves better performance in the case of global policy while score combination significantly achieves better performance in the case of local policy.

In order to investigate the effectiveness of combining the individual metrics on the performances of text categorization and find the best combinations, secondly we propose new combination methods that some of them also outperform the success of the score and rank combinations. At the end of the study we can say that although many feature selection methods exist in text categorization, it is hard to state one is in general superior to others since the success of the methods depends on various variables. It is more likely that combining different feature selection methods obtains more effective performance in text categorization.

As a feature work, we will test the combination of the multiple feature selection metrics. In this study we only focus on binary combinations since the previous studies conclude that the best results achieved by combining two feature selection metrics but their experimental setting is very limited. Thus, it is necessity to see the results of the combination of more than two methods in order to make a clear conclusion. In addition we plan to extend the experiments with new feature selection methods. Further future plan is to propose a model that will provide insight into which feature selection methods achieve better performance when combined.

REFERENCES

1. Sebastiani, F., “Machine Learning in Automated Text Categorization”, *ACM Computing, Surveys*, Vol. 34, No. 1, pp. 1‑47, 2002.

2. Yang, Y. and J. O. Pedersen, “A Comparative Study on Feature Selection in Text Categorization”, *Proceedings of the Fourteenth International Conference on Machine Learning*, pp.412-420, July 08-12, 1997.

3. Forman, G., “An Extensive Empirical Study of Feature Selection Metrics for Text Classification”, *Journal of Machine Learning Research 3*, pp. 1289–1305, 2003.

4. Debole, F. And F. Sebastiani, “Supervised Term Weighting for Automated Text Categorization”, *Proceedings of SAC-03-18th ACM Symposium on Applied Computing*, ACM Press, pp. 784–788, 2003.

5. Ozgur, A., L. Ozgur and T. Gungor, “Text Categorization with Class-Based and Corpus-Based Keyword Selection”, *Proceedings of the 20th International Symposium on Computer and Information Sciences*, *Lecture Notes in Computer Science*, Vol.3733, pp.607-616, Springer-Verlag, 2005.

6. Ozgur, A. and T. Gungor, “Classification of Skewed and Homogenous Document Corpora with Class-Based and Corpus-Based Keywords”, *29th German Conference on Artificial Intelligence (KI 2006), Bremen - LNAI (Lecture Notes in Artificial Intelligence),* Vol.4314, pp.91-101, Springer-Verlag, 2006.

7. Tasci, S., “An evaluation of existing and new feature selection metrics in text categorization”, Computer Engineering, Bogaziçi University, 2006

8. Tasci, S. and T. Gungor “An evaluation of existing and new feature selection metrics in text categorization”, *Computer and Information Sciences ISCIS '08 23rd International Symposium on*, pp.1-6, 2008

9. Croft, W. B., “Combining Approaches to Information Retrieval”, *Advances in Information Retrieval*, pp. 1-36, 2000.

10. Olsson, J. S. and D. W. Oard, “Combining Feature Selectors for Text Classification”, *CIKM’06*, 2006.

11. Li, Y., D. F. Hsu and S. M. Chung “Combining Multiple Feature Selection Methods for Text Categorization by Using Rank-Score Characteristics”, *International Conference on Tools with Artificial Intelligence - ICTAI* , pp. 508-517, 2009

12. Tumer, K. and J. Ghosh, “Linear and order statistics combiners for pattern classification”, *Combining Artificial Neural Networks*, Springer Verlag, pp. 127–162, 1999.

13. Tax, D. M. J.; van Breukelen, M.; Duin, R. P. W. and Kittler, J., “Combining Multiple Classifiers by Averaging or by Multiplying”, *Pattern Recognition 33 (9)*, pp.1475 – 1485, 2000.

14. Dumais, S. T., J. Platt, D. Heckerman and M. Sahami, “Inductive learning algorithms and representations for text categorization”, *Submitted for publication*, 1998.

15. Yang, Y. and X. Liu, “A Re-examination of Text Categorization Methods”, *Proceedings of the Twenty-Second International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 42-49, 1999.

16. Joachims, T., “Text categorization with support vector machines: Learning with many relevant features”, *Proceedings 10^{th} European Conference on Machine Learning (ECML)*, Springer Verlag, 1998.

17. Liu, Y., H. T. Loh and A. Sun, “Imbalanced text classification: A term weighting approach”, *Expert Systems with Applications*, Volume**:** 36, Issue: 1, Elsevier Ltd, pp.690-701, 2009.

18. Salton, G., C. Yang, and A. Wong, "A Vector-Space Model for Automatic Indexing", *Communications of the ACM*, Vol. 18, No. 11, pp. 613‐620, 1975.

19. Han, E-H.S., Boley, D., Gini, M., Gross, R., Hastings, K., Karypis, G., Kumar, V., Mobasher, B., Moore, J., "WebAce: A Web Agent for Document Categorization and Exploration", *Proceedings of the 2nd International Conference on Autonomous Agents,* 1998.

20. Zhang, W., T. Yoshida, X. Tang, "Text classification based on multi-word with support vector machine", *Knowledge-Based Systems 21*, pp.879–886, 2008.

21. ftp://ftp.cs.cornell.edu/pub/smart/, 2004.

22. Porter, M. F., "An Algorithm for Suffix Stripping", *Program*, Vol. 14, pp. 130‐137, 1980.

23. http://tartarus.org/~martin/PorterStemmer/, 2004.

24. Salton, G. and C. Buckley, "Term Weighting Approaches in Automatic Text Retrieval", *Information Processing and Management 24 no. 5*, pp. 513–523, 1988.

25. Zobel, J. and A. Moffat, "Exploring the similarity space", *SIGIR Forum,* 32(1), pp. 18–34, 1998.

26. Wanas, N., D. A. Said, N. H. Hegazy and N. M. Darwish, "A Study of Local and Global Thresholding Techniques in Text Categorization", *Proc. Fifth Australasian Data Mining Conference*, 2006

27. Calvo, R. and H. Ceccatto, 'Intelligent document classification', *Intelligent Data Analysis 4*, pp. 411–420, 2000

28. Fox, E. and J. Shaw, “Combination of multiple searches”, *Proceedings of the 2nd Text Retrieval Conference (TREC-2)*, National Institute of Standards and Technology Special Publication 500-215, pp. 243–252, 1994

29. Hull, D., J.Pedersen and H. Schutze, “Method combination for document filtering”, *Proceedings of the 19th ACM SIGIR Conference on Researchand Development in Information Retrieval*, pp.279–287, 1996.

30. TREC. Text REtrieval conference. http://trec.nist.gov, 1999.

31. Yizhang, G., “Methods for Pattern Classification”, *New Advances in Machine Learning*, InTech Press, pp. 49-74, 2010

32. Leopold, E. and J. Kindermann, “Text Categorization with Support Vector Machines. How to Represent Texts in Input Space? ”, *Machine Learning*, 46, 423–444, 2002

33. Rokach, L. and O. Maimon, *Data Mining with Decision Trees: Theory and Applications,* Word Scientific Publishing, 2008

34. Kivinen, J., M. K. Warmuth and P. Auer, “The Perceptron algorithm vs. Winnow: linear vs. logarithmic mistake bounds when few input variables are relevant”, 1997

35. Manning, C., P. Raghavan and H. Schütze, “Introduction to Information Retrieval”, Cambridge University Press, 2008.

36. Stranieri, A. and J. Zeleznikow, *Knowledge Discovery from Legal Databases*, Springer, 2005

37. Cooley, R., “Classification of news stories using support vector machines”, IJCAI'99 *Workshop on Text Mining*, Stockholm, Sweden, 1999.

38. Zhao, Y. and G. Karypis, "Criterion Functions for Document Clustering Experiments and Analysis", University of Minnesota, Minneapolis, 2002

39. Tin, J. and Y. Kwok, “Automated Text Categorization Using Support Vector Machine”, *International Conference on Neural Information Processing (ICONIP),* 1998

40. Tang, N., “Text Categorization using Support Vector Machines”, 2001

41. Cortes, C. and V. Vapnik. “Support-vector networks”, *Machine Learning*, 20:273- 297, 1995.

42. Vapnik, V., “Statistical Learning Theory”, AT&T Labs-Research, London University, Wiley, 1998.

43. Z. Bodo, Z. Minier and L. Csato, "Text Categorization Experiments Using Wikipedia", *Proceedings of the International Text Wikipedia Conference on Knowledge Engineering*, Principles and Techniques, pp. 66-72, 2007.

44. Sra. S., A. Banerjee, J. Ghosh and I. Dhillon, “Text Clustering with Mixture of von Mises-Fisher Distributions”, *Text Mining Classification, Clustering, and Applications*, Chapman and Hall/CRC, pp.121–161, 2009.

45. Zhao, Y. and G. Karypis, “Evaluation of Hierarchical Clustering Algorithms for Document Datasets”, *Conference on Information and Knowledge Management CIKM’02*, 2002.

46. Xiong, H., M. Steinbach, A. Ruslim and V. Kumar, "Characterizing pattern preserving clustering", *Knowledge and Information Systems*, Vol. 19, No. 3., pp. 311-336, 2009.

47. Ertz, L., M. Steinbach and V. Kumar, "Finding Topics in Collections of Documents: A Shared Nearest Neighbor Approach", *Proceedings of Text Mine’01, First SIAM International Conference on Data Mining*, 2001.

48. Kittler, J., M. Hatef and R.P.W. Duin, “Combining classifiers”, *Proceedings of ICPR'96*, pp.897-901, 1996.

49. Lee, J., “Combining multiple evidence from different properties of weighting schemes”, *Proceedings of the 18th ACMSIGIR Conference on Research and Development in Information Retrieval*, pp. 180–188, 1995.

50. Lee, J., “Analyses of multiple evidence combination”, *Proceedings of* the 20th ACM SIGIR Conference on Research and Development in Information *Retrieval*, pp. 267–276, 1997.

APPENDIX A: Stopword list

a
a's
able
about
above
according
accordingly
across
actually
after
afterwards
again
against
ain't
all
allow
allows
almost
alone
along
already
also
although
always
am
among
amongst
an
and
another
any
anybody
anyhow
anyone
anything
anyway
anyways
anywhere
apart
appear
appreciate
appropriate
are
aren't
around
as
aside
ask
asking
associated
at
available
away
awfully
b
be
became
because
become
becomes
becoming
been
before
beforehand
behind
being
believe
below
beside
besides
best
better
between
beyond
both
brief
but
by
c
c'mon
c's
came
can
can't
cannot
cant
cause
causes
certain
certainly
changes
clearly
co
com
come
comes
concerning
consequently
consider
considering
contain
containing
contains
corresponding
could
couldn't
course
currently
d
definitely
described
despite
did
didn't
different
do
does
doesn't
doing
don't
done
down
downwards
during
e
each
edu
eg
eight
either
else
elsewhere
enough
entirely
especially
et
etc
even
ever
every
everybody
everyone
everything
everywhere
ex
exactly
example
except
f
far
few
fifth
first
five
followed
following
follows
for
former
formerly
forth
four
from
further
furthermore
g
get
gets
getting
given
gives
go
goes
going
gone
got
gotten
greetings
h
had
hadn't
happens
hardly
has
hasn't
have
haven't
having
he
he's
hello
help
hence
her
here
here's
hereafter
hereby
herein
hereupon
hers
herself
hi
him
himself
his
hither
hopefully
how
howbeit
however
i
i'd
i'll
i'm
i've
ie
if
ignored
immediate
in
inasmuch
inc
indeed
indicate
indicated
indicates
inner
insofar
instead
into
inward
is
isn't
it
it'd
it'll
it's
its
itself
j
just
k
keep
keeps
kept
know
knows
known
l
last
lately
later
latter
latterly
least
less
lest
let
let's
like
liked
likely
little
look
looking
looks
ltd
m
mainly
many
may
maybe
me
mean
meanwhile
merely
might
more
moreover
most
mostly
much
must
my
myself
n
name
namely
nd
near
nearly
necessary
need
needs
neither
never
nevertheless
new
next
nine
no
nobody
non
none
noone
nor
normally
not
nothing
novel
now
nowhere
o
obviously
of
off
often
oh
ok
okay
old
on
once
one
ones
only
onto
or
other
others
otherwise
ought
our
ours
ourselves
out
outside
over
overall
own
p
particular
particularly
per
perhaps
placed
please
plus
possible
presumably
probably
provides
q
que
quite
qv
r
rather
rd
re
really
reasonably
regarding
regardless
regards
relatively
respectively
right
s
said
same
saw
say
saying
says
second
secondly
see
seeing
seem
seemed
seeming
seems
seen
self
selves
sensible
sent
serious
seriously
seven
several
shall
she
should
shouldn't
since
six
so
some
somebody
somehow
someone
something
sometime
sometimes
somewhat
somewhere
soon
sorry
specified
specify
specifying
still
sub
such
sup
sure
t
t's
take
taken
tell
tends
th
than
thank
thanks
thanx
that
that's
thats
the
their
theirs
them
themselves
then
thence
there
there's
thereafter
thereby
therefore
therein
theres
thereupon
these
they
they'd
they'll
they're
they've
think
third
this
thorough
thoroughly
those
though
three
through
throughout
thru
thus
to
together
too
took
toward
towards
tried
tries
truly
try
trying
twice
two
u
un
under
unfortunately
unless
unlikely
until
unto
up
upon
us
use
used
useful
uses
using
usually
uucp
v
value
various
very
via
viz
vs
w
want
wants
was
wasn't
way
we
we'd
we'll
we're
we've
welcome
well
went
were
weren't
what
what's
whatever
when
whence
whenever
where
where's
whereafter
whereas
whereby
wherein
whereupon
wherever
whether
which
while
whither
who
who's
whoever
whole
whom
whose
why
will
willing
wish
with
within
without
won't
wonder
would
would
wouldn't
x
y
yes
yet
you
you'd
you'll
you're
you've
your
yours
yourself
yourselves
z
zero