

Arzucan Özgür

Dept. of Computer Engineering
Boğaziçi University
34342 Bebek, Istanbul, Turkey

Phone +90 (212) 359-7226
E-mail arzucan.ozgur@boun.edu.tr
Url <http://www.cmpe.boun.edu.tr/~ozgur/>

RESEARCH INTERESTS

Bioinformatics and Natural Language Processing

EDUCATION

University of Michigan, Ann Arbor, Michigan September 2006 – September 2010
Ph.D., Computer Science and Engineering
Advisor: Dragomir R. Radev
Thesis: “Text and Network Mining for Literature-Based Scientific Discovery in Biomedicine”

Boğaziçi University, İstanbul, Turkey September 2002 – September 2004
M.S., Computer Engineering
Advisor: Ethem Alpaydın
Thesis: “Supervised and Unsupervised Machine Learning Techniques for Text Document Categorization”

Boğaziçi University, İstanbul, Turkey September 1998 – July 2002
B.S., Computer Engineering

PROFESSIONAL EXPERIENCE

Boğaziçi University, İstanbul, Turkey March 2017 – Present
Associate Professor, Department of Computer Engineering

Boğaziçi University, İstanbul, Turkey September 2011 – March 2017
Assistant Professor, Department of Computer Engineering

İstanbul Technical University, İstanbul, Turkey October 2010 – August 2011
Instructor, Faculty of Computer and Informatics

Avaya Labs Research, Westminster, Colorado June 2008 – August 2008
Research Intern, Multimedia Technologies Research Department

University of Michigan, Ann Arbor, Michigan September 2006 – September 2010
Research Assistant,
Department of Electrical Engineering and Computer Science
National Center for Integrative Biomedical Informatics

Boğaziçi University, İstanbul, Turkey September 2002 – July 2006
Teaching Assistant, Department of Computer Engineering

HONORS AND AWARDS

- Excellence in Research Award, Boğaziçi University Foundation (BÜVAK), Turkey, 2022.
- Outstanding Reviewer Recognition, Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020.
- The Turkish Science Academy Young Scientist Award (TÜBA-GEBIP), Turkey, 2019.
- The Science Academy Young Scientist Award (BAGEP), Turkey, 2016.
- European Commission FP7 Marie Curie Actions Career Integration Grant (CIG), 2012-2016.

PUBLICATIONS

Patents

- George Erhart, Valentine Matula, Arzucan Ozgur, and David Skiba. System and Method for Generating Queries, United States Patent and Trademark Office (USPTO), US 8301619. Assigned to Avaya Inc., Basking Ridge, N.J. (US), October 2012.

Thesis

- Arzucan Özgür. “Text and Network Mining for Literature-Based Scientific Discovery in Biomedicine”, Ph.D. Thesis, Computer Science and Engineering, University of Michigan, Ann Arbor, MI, USA, September 2010. Thesis Advisor: Dragomir R. Radev.
- Arzucan Özgür. “Supervised and Unsupervised Machine Learning Techniques for Text Document Categorization”, M.S. Thesis, Computer Engineering, Boğaziçi University, İstanbul, Turkey, July 2004. Thesis Advisor: Ethem Alpaydın.

Refereed International Journal Papers

- [1] Gokce Uludogan, Elif Ozkirimli, Kutlu Ulgen, Nilgun Karali, Arzucan Ozgur. “Exploiting pre-trained biochemical language models for targeted drug design.” *Bioinformatics*, Volume 38, Issue Supplement_2, Pages ii155–ii161, 2022.
- [2] Arda Celebi and Arzucan Ozgur. “Cluster-based mention typing for named entity disambiguation.” *Natural Language Engineering*, 28(1), 1-37, 2022.
- [3] Saziye Betul Ozates, Arzucan Ozgur, Tunga Gungor, Balkiz Ozturk Basaran. “A Hybrid Deep Dependency Parsing Approach Enhanced With Rules and Morphology: A Case Study for Turkish.” *IEEE Access*, vol. 10, pp. 93867-93886, 2022.
- [4] Utku Turk, Furkan Atmaca, Saziye Betul Ozates, Gozde Berk, Seyyit Talha Bedir, Abdullatif Koksall, Balkiz Ozturk, Tunga Gungor, Arzucan Ozgur. “Resources for Turkish dependency parsing: Introducing the BOUN treebank and the BoAT annotation tool.” *Language Resources and Evaluation*, 56 (1), pp. 259-307, 2022.
- [5] Gonul Ayci, Murat Sensoy, Arzucan Ozgur, and Pinar Yolum. “Uncertainty-aware Personal Assistant for Making Personalized Privacy Decisions.” *ACM Trans. Internet Technol.*, 2022.
- [6] Artur Yakimovich, Arzucan Ozgur, Tunca Dogan, and Elif Ozkirimli. “Editorial: Machine Learning Methodologies to Study Molecular Interactions.” *Front Mol Biosci*, 8:806474, 2021.
- [7] Riza Ozcelik, Hakime Ozturk, Arzucan Ozgur, Elif Ozkirimli. “ChemBoost: A Chemical Language Based Approach for Protein–Ligand Binding Affinity Prediction.” *Molecular Informatics*, Volume 40, Issue 5, Pages 2000212, 2021.
- [8] N. Ozlem Ozcan Simsek, Arzucan Ozgur, Fikret Gurgun. “A novel gene selection method for gene expression data for the task of cancer type classification.” *Biology Direct*, Volume 16, Issue 1, Pages 1–5, 2021.
- [9] Hakime Ozturk, Arzucan Ozgur, Philippe Schwaller, Teodoro Laino, Elif Ozkirimli. “Exploring chemical space using natural language processing methodologies for drug discovery.” *Drug*

Discovery Today, Volume 25, Issue 4, Pages 689-705, 2020.

- [10] Mert Tiftikci, Arzucan Ozgur, Yongqun He, Junguk Hur. "Machine learning-based identification and rule-based normalization of adverse drug reactions in drug labels." *BMC Bioinformatics*, 20:21, 2019.
- [11] N. Ozlem Ozcan Simsek, Arzucan Ozgur, Fikret Gurgen. "Statistical representation models for mutation information within genomic data." *BMC Bioinformatics*, 20:324, 2019.
- [12] Ilknur Karadeniz, Arzucan Ozgur. "Linking Entities through an Ontology using Word Embeddings and Syntactic Re-ranking". *BMC Bioinformatics*, 20:156, 2019.
- [13] Rezarta Islamaj Dogan, Sun Kim, Andrew Chatr-Aryamontri, Chih-Hsuan Wei, Donald Comeau, Rui Antunes, Sergio Matos, Qingyu Chen, Aparna Elangovan, Nagesh C Panyam, Karin Verspoor, Hongfang Liu, Yanshan Wang, Zhuang Liu, Berna Altinel, Zehra Melce Husunbeyi, Arzucan Ozgur, Aris Fergadis, Chen-Kai Wang, Hong-Jie Dai, Tung Tran, Ramakanth Kavuluru, Ling Luo, Albert Steppi, Jinfeng Zhang, Jinchan Qu, Zhiyong Lu. "Overview of the BioCreative VI Precision Medicine Track: mining protein interactions and mutations for precision medicine". *Database*, Volume 2019, bay147, 1 January 2019.
- [14] Hakime Ozturk, Arzucan Ozgur, and Elif Ozkirimli. "DeepDTA: deep drug-target binding affinity prediction". *Bioinformatics*, Volume 34, Issue 17, Pages i821-i829, 2018.
- [15] Hakime Ozturk, Elif Ozkirimli, and Arzucan Ozgur. "A novel methodology on distributed representations of proteins using their interacting ligands". *Bioinformatics*, Volume 34, Issue 13, Pages i295-i303, 2018.
- [16] Arda Celebi and Arzucan Ozgur. "Segmenting hashtags and analyzing their grammatical structure". *Journal of the Association for Information Science and Technology (JASIST)*. doi:10.1002/asi.23989, 69(5):675-686, 2018.
- [17] Caner Derici, Yigit Aydin, Cigdem Yenialaca, Nihal Yagmur Aydin, Gunizi Kartal, Arzucan Ozgur, Tunga Gungor. "A closed-domain question answering framework using reliable resources to assist students". *Natural Language Engineering*, Volume 24, Issue 5, pp. 725-762, 2018.
- [18] Junguk Hur, Arzucan Ozgur, and Yongqun He. "Ontology-based literature mining and class effect analysis of adverse drug reactions associated with neuropathy-inducing drugs". *Journal of Biomedical Semantics*, 9:17, 2018.
- [19] Gizem Sogancioglu, Hakime Ozturk, and Arzucan Ozgur. "BIOSSES: A Semantic Sentence Similarity Estimation System for the Biomedical Domain". *Bioinformatics*, 33 (14): i49-i58, 2017.
- [20] Ferhat Aydin, Zehra Melce Husunbeyi, and Arzucan Ozgur. "Automatic query generation using word embeddings for retrieving passages describing experimental methods". *Database*, 2017: baw166, doi: 10.1093/database/baw166, 2017.
- [21] Junguk Hur, Arzucan Ozgur, and Yongqun He. "Ontology-based literature mining of E. coli vaccine-associated gene interaction networks". *Journal of Biomedical Semantics*, 8:12, 2017.

- [22] Arzucan Ozgur, Junguk Hur, and Yongqun He. "The Interaction Network Ontology-supported modeling and mining of complex interactions represented with multiple keywords in biomedical literature". *BioData Mining*, 9:41, doi: 10.1186/s13040-016-0118-0, 2016.
- [23] Erinc Gokdeniz, Arzucan Ozgur, and Resit Canbeyli. "Automated Neuroanatomical Relation Extraction: A Linguistically Motivated Approach with a PVT Connectivity Graph Case Study". *Frontiers in Neuroinformatics*, 110:39. doi: 10.3389/fninf.2016.00039, 2016.
- [24] Sun Kim, Rezarta Islamaj Dogan, Andrew Chatr-aryamontri, Christie S. Chang, Rose Oughtred, Jennifer M. Rust, Riza Batista-Navarro, Jacob Carter, Sophia Ananiadou, Sergio Matos, Andr Santos, David Campos, Jos Lus Oliveira, Onkar Singh, Jitendra Jonnagaddala, Hong-Jie Dai, Emily Chia-Yu Su, Yung-Chun Chang, Yu-Chen Su, Chun-Han Chu, Chien Chin Chen, Wen-Lian Hsu, Yifan Peng, Cecilia N. Arighi, Cathy H. Wu, K. Vijay-Shanker, Ferhat Aydin, Zehra Melce Hsnbeyi, Arzucan zgr, Soo-Yong Shin, Dongseop Kwon, Kara Dolinski, Mike Tyers, W. John Wilbur, Donald C. Comeau. "BioCreative V BioC track overview: collaborative biocurator assistant task for BioGRID". *Database*, 2016 : baw121. doi: 10.1093/database/baw121, 2016.
- [25] Hakime Ozturk, Elif Ozkirimli, and Arzucan Ozgur. "A comparative study of SMILES-based compound similarity functions for drug-target interaction prediction". *BMC Bioinformatics*, 17:128, 2016.
- [26] Ahmet Yildirim, Suzan Uskudarli, and Arzucan Ozgur. "Identifying Topics in Microblogs Using Wikipedia". *PLoS ONE*, 11(3): e0151885, 2016.
- [27] Ilknur Karadeniz, Junguk Hur, Yongqun He, and Arzucan Ozgur. "Literature Mining and Ontology based Analysis of Host-Brucella Gene-Gene Interaction Network". *Frontiers in Microbiology*, 6:1386, 2015.
- [28] Ilknur Karadeniz and Arzucan Ozgur. "Detection and Categorization of Bacteria Habitats using Shallow Linguistic Analysis". *BMC Bioinformatics*, 16(Suppl 10):S5, 2015.
- [29] Wallace K. B. Chan, Hongjiu Zhang, Jianyi Yang, Jeffrey R. Brender, Junguk Hur, Arzucan Ozgur, and Yang Zhang. "GLASS: a comprehensive database for experimentally-validated GPCR-ligand associations". *Bioinformatics*, Volume 31, Issue 18, Pp. 3035-3042, 2015.
- [30] Saliha Durmus, Tunahan Cakir, Arzucan Ozgur, Reinhard Guthke. "A Review on Computational Systems Biology of Pathogen-Host Interactions". *Frontiers in Microbiology*, 6, 235, 2015.
- [31] Hakime Ozturk, Elif Ozkirimli, Arzucan Ozgur. "Classification of Beta-Lactamases and Penicillin Binding Proteins Using Ligand-Centric Network Models". *PLoS ONE*, 10(2): e0117874, 2015.
- [32] Junguk Hur*, Arzucan Ozgur*, Zuoshuang Xiang, and Yongqun He. "Development and Application of an Interaction Network Ontology for Literature Mining of Vaccine-associated Gene-Gene Interactions", *Journal of Biomedical Semantics*, 6:2, 2015. (* Equal contribution)
- [33] Melike Korucuoglu, Senol Isci, Arzucan Ozgur, Hasan H. Otu. "Bayesian Pathway Analysis of Cancer Microarray Data". *PLoS ONE*, 9(7): e102803, 2014.
- [34] Saliha Durmus Tekir, Tunahan Cakir, Emre Ardic, Ali Semih Sayilirbas, Gokhan Konuk, Mithat Konuk, Hasret Sariyer, Azat Ugurlu, Ilknur Karadeniz, Arzucan Ozgur, F. Erdogan Sevilgen, and Kutlu O. Ulgen. "PHISTO: Pathogen-Host Interaction Search Tool". *Bioinformatics*, Vol. 29 no. 10, pages 1357-1358, 2013.

- [35] Junguk Hur*, Arzucan Ozgur*, Zuoshuang Xiang, and Yongqun He. "Identification of fever and vaccine-associated gene interaction networks using ontology-based literature mining". *Journal of Biomedical Semantics*, 3:18, 2012. (* Equal contribution)
- [36] Y. Kano, J. Bjerne, F. Ginter, E. Buyko, U. Hahn, K. B. Cohen, K. Verspoor, C. Roeder, L. E. Hunter, H. Kilicoglu, S. Bergler, S. V. Landeghem, T. V. Parys, Y. V. de Peer, M. Miwa, S. Ananiadou, M. Neves, A. Pascual-Montano, A. Özgür, D. R. Radev, S. Riedel, R. Satre, H-W. Chun, J-D. Kim, S. Pyysalo, T. Ohta, and J. Tsujii. "U-Compare Bio-Event Meta-Service: Compatible BioNLP Event Extraction Services ", *BMC Bioinformatics*, 12:481, 2011.
- [37] Arzucan Özgür, Zuoshuang Xiang, Dragomir R. Radev, and Yongqun He. "Mining of vaccine-associated IFN-gamma gene interaction networks using the Vaccine Ontology", *Journal of Biomedical Semantics*, 2(Suppl 2):S8, 2011.
- [38] Arzucan Özgür, Zuoshuang Xiang, Dragomir R. Radev, and Yongqun He. "Literature-based discovery of IFN-gamma and vaccine-mediated gene interaction networks", *BioMed Research International*, vol. 2010, Article ID 426479, 2010.
- [39] V. Glenn Tarcea, Terry Weymouth, Alex Ade, Jing Gao, Vasudeva Mahavisno, Zach Wright, Adriane Chapman, Magesh Jayapandian, Arzucan Özgür, Yuanyuan Tian, Jim Cavalcoli, Barbara Mirel, Jignesh Patel, Dragomir Radev, Brian Athey, David States, and H.V. Jagadish. "Michigan Molecular Interactions (MiMI) R2: From Interacting Proteins to Pathways", *Nucleic Acids Research*, 37 (Database Issue), D642-D646, 2009.
- [40] Arzucan Özgür, Thuy Vu, Güneş Erkan, and Dragomir R. Radev. "Identifying gene-disease associations using centrality on a literature mined gene interaction network", *Bioinformatics*, Volume 24, Number 13, pp. i277-i285, 2008.
- [41] F. Leitner, M. Krallinger, C. Rodriguez-Penagos, J. Hakenberg, C. Plake, C-J. Kuo, C-N. Hsu, R. T-H. Tasi, H-C. Hung, W. W. Lau, C. A. Johnson, R. Satre, K. Yoshida, Y. H. Chen, S. Kim, S-Y. Shin, B-T. Zhang, W. A. Baumgartner Jr., L. Hunter, B. Haddow, M. Matthew, X. Wang, P. Ruch, F. Ehrler, A. Özgür, G. Erkan, D. R. Radev, M. Krauthammer, T. Luong, R. Hoffmann, C. Sander, and A. Valencia. "Introducing Meta-Services for Biomedical Information Extraction", *Genome Biology*, 9(S2):S6, 2008.
- [42] Arzucan Özgür, Burak Çetin, and Haluk Bingol. "Co-occurrence Network of Reuters News", *International Journal of Modern Physics C (IJMPC)*, Volume 19, Issue 5, pp. 689-702, 2008.
- [43] Arzucan Özgür and Taflan İ. Gündem. "Efficient Indexing Technique for XML-Based Electronic Product Catalogs", *Electronic Commerce Research and Applications*, Volume 5, Issue 1, pp. 66-77, Elsevier Science, 2006.

Refereed International Conference and Workshop Papers

- [44] Merve Unlu Menevse, Yusufcan Manav, Ebru Arisoy, Arzucan Ozgur. "A Framework for Automatic Generation of Spoken Question-Answering Data." *Findings of the Association for Computational Linguistics: EMNLP 2022* (Accepted).
- [45] Saziye Betul Ozates, Arzucan Ozgur, Tunga Gungor, Ozlem Cetinoglu. "Improving Code-Switching Dependency Parsing with Semi-Supervised Auxiliary Tasks." *Findings of the Association for Computational Linguistics: NAACL 2022*, pp. 1159-1171, 2022.

- [46] Mustafa Melih Mutlu and Arzucan Ozgur. “A Dataset and BERT-based Models for Targeted Sentiment Analysis on Turkish Texts.” In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop, pages 467–472, Dublin, Ireland, 2022. Association for Computational Linguistics.
- [47] Zehra Melce Husunbeyi, Didar Akar, and Arzucan Ozgur. “Identifying Hate Speech Using Neural Networks and Discourse Analysis Techniques.” In Proceedings of the First Workshop on Language Technology and Resources for a Fair, Inclusive, and Safe Society (LATERAISSE) within the 13th Language Resources and Evaluation Conference, pages 32–41, Marseille, France, 2022. European Language Resources Association.
- [48] Yi Huang, Buse Giledereli, Abdullatif Koksall, Arzucan Ozgur, Elif Ozkirimli. “Balancing Methods for Multi-label Text Classification with Long-Tailed Class Distribution”. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 8153–8161, Online and Punta Cana, Dominican Republic, 2021. Association for Computational Linguistics.
- [49] Yi Huang, Buse Giledereli, Abdullatif Koksall, Arzucan Ozgur, and Elif Ozkirimli. “PIDNA at BioASQ MESINESP: Hybrid Semantic Indexing for Biomedical Articles in Spanish.” CLEF (Working Notes), 241-246, 2021.
- [50] Abdullatif Koksall, Yusuf Yuksel, Bekir Yildirim, and Arzucan Ozgur. “BOUN at SemEval-2021 Task 9: Text Augmentation Techniques for Fact Verification in Tabular Data.” In Proceedings of the 15th International Workshop on Semantic Evaluation (SemEval-2021), pages 431–437, Online, 2021. Association for Computational Linguistics.
- [51] Talha Bedir, Karahan Sahin, Onur Gungor, Suzan Uskudarli, Arzucan Ozgur, Tunga Gungor, Balkiz Ozturk. “Overcoming the challenges in morphological annotation of Turkish in universal dependencies framework”. In Proceedings of the Joint 15th Linguistic Annotation Workshop (LAW) and 3rd Designing Meaning Representations (DMR) Workshop, pages 112–122, Punta Cana, Dominican Republic, 2021. Association for Computational Linguistics.
- [52] Abdullatif Koksall, Arzucan Ozgur. “The RELX Dataset and Matching the Multilingual Blanks for Cross-lingual Relation Classification”. In Findings of the Association for Computational Linguistics: EMNLP 2020, pages 340–350, 2020. Association for Computational Linguistics.
- [53] Abdullatif Koksall, Hilal Donmez, Riza Ozcelik, Elif Ozkirimli, Arzucan Ozgur. “Vapur: A Search Engine to Find Related Protein - Compound Pairs in COVID-19 Literature”. In Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020, Online, 2020. Association for Computational Linguistics.
- [54] Berfu Buyukoz, Ali Hurriyetoglu, Arzucan Ozgur. “Analyzing ELMo and DistilBERT on socio-political news classification”. In Proceedings of the Workshop on Automated Extraction of Socio-political Events from News, pages 9–18, Marseille, France, 2020. European Language Resources Association (ELRA).
- [55] Hilal Donmez, Abdullatif Koksall, Elif Ozkirimli, Arzucan Ozgur. “BOUN-REX at CLEF-2020 ChEMU Task 2: Evaluating Pretrained Transformers for Event Extraction”. In Proceedings of CLEF (Conference and Labs of the Evaluation Forum) 2020 Working Notes, Thessaloniki, Greece, September 22-25, 2020.
- [56] Jiayuan He, Dat Quoc Nguyen, Saber A. Akhondi, Christian Druckenbrodt, Camilo Thorne, Ralph Hoessel, Zubair Afzal, Zenan Zhai, Biaoyan Fang, Hiyori Yoshikawa, Ameer Albahem,

Jingqi Wang, Yuankai Ren, Zhi Zhang, Yaoyun Zhang, Mai Hoang Dao, Pedro Ruas, Andre Lamurias, Francisco M. Couto, Jenny Copara, Nona Naderi, Julien Knafou, Patrick Ruch, Douglas Teodoro, Daniel M. Lowe, John Mayfield, Abdullatif Koksall, Hilal Donmez, Elif Ozkirimli, Arzucan Ozgur, Darshini Mahendran, Gabrielle Gurdin, Nastassja Lewinski, Christina Tang, Bridget T. McInnes, C. S. Malarkodi, Pattabhi Rk Rao, Sobha Lalitha Devi, Lawrence Cavedon, Trevor Cohn, Timothy Baldwin, Karin Verspoor. "An Extended Overview of the CLEF 2020 ChEMU Lab: Information Extraction of Chemical Reactions from Patents". In Proceedings of CLEF (Conference and Labs of the Evaluation Forum) 2020 Working Notes, Thessaloniki, Greece, September 22-25, 2020.

- [57] Atif Emre Yuksel, Yasar Alim Turkmen, Arzucan Ozgur, Ayse Berna Altinel. "Turkish Tweet Classification with Transformer Encoder". In Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019), pp. 1380-1387, 2019.
- [58] Ilknur Karadeniz, Omer Faruk Tuna, Arzucan Ozgur. "BOUN-ISIK Participation: An Unsupervised Approach for the Named Entity Normalization and Relation Extraction of Bacteria Biotopes", Proceedings of The 5th Workshop on BioNLP Open Shared Tasks, pages 150-157, Hong Kong, China, November 4, 2019.
- [59] Utku Turk, Furkan Atmaca, Saziye Betul Ozates, Balkiz Ozturk Basaran, Tunga Gungor, Arzucan Ozgur. "Improving the Annotations in the Turkish Universal Dependency Treebank", Proceedings of the Third Workshop on Universal Dependencies (UDW, SyntaxFest 2019), Paris, France, 26 August, 2019.
- [60] Utku Turk, Furkan Atmaca, Saziye Betul Ozates, Abdullatif Koksall, Balkiz Ozturk Basaran, Tunga Gungor, Arzucan Ozgur. "Turkish Treebanking: Unifying and Constructing Efforts", Proceedings of the 13th Linguistic Annotation Workshop (LAW), pages 166-177, Florence, Italy, 1 August, 2019.
- [61] Saziye Betul Ozates, Arzucan Ozgur, Tunga Gungor, Balkiz Ozturk. "A Morphology-based Representation Model for LSTM-based Dependency Parsing of Agglutinative Languages", Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies, pages 238-247, Brussels, Belgium, 2018.
- [62] Atakan Yuksel, Hakime Ozturk, Elif Ozkirimli, Arzucan Ozgur. "CNN-based Chemical-Protein Interactions Classification", Proceedings of the BioCreative VI Workshop, pages 184-186, Bethesda, MD, USA, 2017.
- [63] Berna Altinel, Zehra Melce Husunbeyi, Arzucan Ozgur. "Text classification using ontology and semantic values of terms for mining protein interactions and mutations", Proceedings of the BioCreative VI Workshop, pages 110-114, Bethesda, MD, USA, 2017.
- [64] Deger Ayata, Murat Saraclar, and Arzucan Ozgur. "BUSEM at SemEval-2017 Task 4A Sentiment Analysis with Word Embedding and Long Short Term Memory RNN Approaches", Proceedings of the 11th International Workshop on Semantic Evaluations (SemEval-2017), pages 777-783, Vancouver, Canada, August 3 - 4, 2017.
- [65] Arzucan Ozgur, Junghuk Hur, Zuoshuang Xiang, Edison Ong, Dragomir R. Radev, and Yongqun He. "Iagnet: A centrality and INO-based web system for analyzing and visualizing literature-mined networks". ICBO/BioCreative, 2016.
- [66] Mert Tiftikci, Hakan Sahin, Berfu Buyukoz, Alper Yayikci, Arzucan Ozgur. "Ontology-based Categorization of Bacteria and Habitat Entities using Information Retrieval Techniques". Proceedings of the 4th BioNLP Shared Task Workshop, pages 56-63, Berlin, Germany, August 13,

2016. Association for Computational Linguistics.

- [67] Cagil Sonmez, Arzucan Ozgur, Erdem Yoruk. “Towards Building a Political Protest Database to Explain Changes in the Welfare State”. Proceedings of the 10th SIGHUM Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH), pages 106–110, Berlin, Germany, August 11, 2016. Association for Computational Linguistics.
- [68] Arda Celebi and Arzucan Ozgur. “Segmenting Hashtags using Automatically Created Training Data”. In Proceedings of the 10th Language Resources and Evaluation Conference (LREC 2016), 23-28 May, 2016.
- [69] Saziye Betul Ozates, Arzucan Ozgur, and Dragomir Radev. “Sentence Similarity based on Dependency Tree Kernels for Multi-document Summarization”. In Proceedings of the 10th Language Resources and Evaluation Conference (LREC 2016), 23-28 May, 2016.
- [70] Eda Okur, Hakan Demir, and Arzucan Ozgur. “Named Entity Recognition on Twitter for Turkish using Semi-supervised Learning with Word Embeddings”. In Proceedings of the 10th Language Resources and Evaluation Conference (LREC 2016), 23-28 May, 2016.
- [71] Arzucan Ozgur, Junguk Hur, and Yongqun He. “Extension of the Interaction Network Ontology for literature mining of gene-gene interaction networks from sentences with multiple interaction keywords”. Proceedings of the International Workshop on Biomedical Data Mining, Modeling, and Semantic Integration: A Promising Approach to Solving Unmet Medical Needs (BDM2I 2015), The International Semantic Web Conference (ISWC 2015), Bethlehem, PA, USA, October 11-15, 2015.
- [72] Ferhat Aydin, Zehra Melce Husunbeyi, and Arzucan Ozgur. “Retrieving Passages Describing Experimental Methods using Ontology and Term Relevance based Query Matching”. Proceedings of the Fifth BioCreative Challenge Evaluation Workshop, pp. 42-50, Sevilla, Spain, September 9-11, 2015.
- [73] Caner Dericci, Kerem Celik, Ekrem Kutbay, Yigit Aydin, Tunga Gungor, Arzucan Ozgur, Gunizi Kartal. “Question Analysis for a Closed Domain Question Answering System”. Proceedings of the Conference on Intelligent Text Processing and Computational Linguistics (CICLing), LNCS 9042, pp. 468–482, Springer, 2015.
- [74] Hakan Demir and Arzucan Ozgur. “Improving Named Entity Recognition for Morphologically Rich Languages using Word Embeddings”. The 13th International Conference on Machine Learning and Applications (ICMLA’14), Detroit, Michigan, USA, December, 2014.
- [75] Cagil Sonmez and Arzucan Ozgur. “A Graph-based Approach for Contextual Text Normalization”. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 313–324, Doha, Qatar, October, 2014. Association for Computational Linguistics.
- [76] Muhammed Yavuz Nuzumlali and Arzucan Ozgur. “Analyzing Stemming Approaches for Turkish Multi-Document Summarization”. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 702–706, Doha, Qatar, October, 2014. Association for Computational Linguistics.
- [77] Burak Aydin and Arzucan Ozgur. “Expanding Machine Translation Training Data with an Out-of-Domain Corpus using Language Modeling based Vocabulary Saturation”. Conference of

the Association for Machine Translation in the Americas (AMTA), Vancouver, Canada, October, 2014.

- [78] Arda Celebi and Arzucan Ozgur. "Self-training a Constituency Parser using N-gram Trees". In Proceedings of the 9th Language Resources and Evaluation Conference (LREC 2014), Reykjavik, Iceland, May, 2014.
- [79] Ilknur Karadeniz and Arzucan Ozgur. "Bacteria Biotope Detection, Ontology-based Normalization, and Relation Extraction using Syntactic Rules". In Proceedings of BioNLP Shared Task 2013 Workshop, Sofia, Bulgaria, August, 2013. Association for Computational Linguistics.
- [80] Nadin Kokciyan, Arda Celebi, Arzucan Ozgur, and Suzan Uskudarli. "BOUNCE: Sentiment Classification in Twitter using Rich Feature Sets". In Proceedings of the International Workshop on Semantic Evaluation, SemEval'13, Atlanta, Georgia, June, 2013.
- [81] Cuneyd Murad Ozsert and Arzucan Ozgur. "Word Polarity Detection using a Multilingual Approach". Proceedings of the Conference on Intelligent Text Processing and Computational Linguistics (CICLing), Samos, Greece, LNCS 7817, pp. 75-82, Springer, 2013.
- [82] Arda Celebi and Arzucan Ozgur. "N-gram Parsing for Jointly Training a Discriminative Constituency Parser". Proceedings of the Conference on Intelligent Text Processing and Computational Linguistics (CICLing), Samos, Greece, 2013.
- [83] Vahed Qazvinian, Dragomir R. Radev, and Arzucan Özgür. "Citation Summarization Through Keyphrase Extraction", In Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010), Beijing, China, August 23-27, 2010.
- [84] Arzucan Özgür and Dragomir R. Radev. "Detecting Speculations and their Scopes in Scientific Text", In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP'09), pages 1398-1407, Singapore, August 2009.
- [85] Arzucan Özgür and Dragomir R. Radev. "Supervised classification for extracting biomedical events", In Proceedings of the BioNLP'09 Workshop Shared Task on Event Extraction at NAACL-HLT, Boulder, Colorado, USA, pages 111-114, June 2009.
- [86] Güneş Erkan, Arzucan Özgür, and Dragomir R. Radev. "Semi-supervised classification for extracting protein interaction sentences using dependency parsing", In Proceedings of the Conference of Empirical Methods in Natural Language Processing (EMNLP '07), Prague, Czech Republic, June 28-30 2007.
- [87] Güneş Erkan, Arzucan Özgür, and Dragomir R. Radev. "Extracting interacting protein pairs and evidence sentences by using dependency parsing and machine learning techniques", In Proceedings of the Second BioCreAtIvE Challenge Workshop - Critical Assessment of Information Extraction in Molecular Biology, Madrid, Spain, ISBN 84-933255-6-2, April 23-25 2007.
- [88] Arzucan Özgür and Tunga Güngör. "Classification of Skewed and Homogenous Document Corpora with Class-Based and Corpus-Based Keywords", Lecture Notes in Artificial Intelligence, Vol. 4314, (Proc. of the 29th German Conference on Artificial Intelligence), pp. 91-101, Springer-Verlag, 2006.
- [89] Arzucan Özgür, Levent Özgür, and Tunga Güngör. "Text Categorization with Class-Based and Corpus-Based Keyword Selection", Lecture Notes in Computer Science, Vol.3733, (Proc. of the 20th International Symposium on Computer and Information Sciences), pp. 607-616,

Springer-Verlag, 2005.

- [90] Arzucan Özgür and Haluk Bingol. “Social Network of Co-occurrence in News Articles”, Lecture Notes in Computer Science, Vol. 3280, (Proc. of the 19th International Symposium on Computer and Information Sciences), pp. 688-697, Springer Verlag, 2004.

Refereed National Conference Papers

- [91] Abdullatif Koxsal and Arzucan Ozgur. “Twitter Dataset and Evaluation of Transformers for Turkish Sentiment Analysis.” 29th Signal Processing and Communications Applications Conference (SIU), pp. 1-4, 2021. IEEE.
- [92] Melik Masarifoglu, Umit Tigrak, Sefa Hakyemez, Guven Gul, Erdal Bozan, Ali Hakan Buyuklu, and Arzucan Ozgur. “Sentiment Analysis of Customer Comments in Banking using BERT-based Approaches.” 29th Signal Processing and Communications Applications Conference (SIU), pp. 1-4, 2021. IEEE.
- [93] Melike Esmâ İlker, Lale Akarun, Arzucan Ozgur. “Identifying Image Related Sentences in News Articles”, SIU 2019.
- [94] Deger Ayata, Murat Saraclar, Arzucan Ozgur. “Turkish Tweet Sentiment Analysis with Word Embedding and Machine Learning”, SIU 2017.
- [95] Deger Ayata, Murat Saraclar, Arzucan Ozgur. “Political Opinion/Sentiment Prediction via Long Short Term Memory Recurrent Neural Networks on Twitter”, SIU 2017.
- [96] Caner Dericci, Kerem Celik, Ekrem Kutbay, Yigit Aydin, Arzucan Ozgur, Gunizi Kartal, Tunga Gungor. “Rule-Based Focus Extraction in Turkish Question Answering Systems”, 2014 IEEE 22nd Signal Processing and Communications Applications Conference (SIU 2014).

FUNDED PROJECTS

- | | |
|---|-------------|
| European Commission - EuropeAid | 2021 - 2024 |
| Project name: Utilizing Digital Technology for Social Cohesion, Positive Messaging and Peace by Boosting Collaboration, Exchange and Solidarity | |
| Co-investigator | |
| The Scientific and Technological Research Council of Turkey, TUBITAK 1001 | 2020 - 2023 |
| Project name: Chemical Language Processing for Target-based Drug Design | |
| Principal Investigator | |
| The Scientific and Technological Research Council of Turkey, TUBITAK 1005 | 2018 - 2019 |
| Project name: A Deep Learning based Turkish Dependency Parser | |
| Principal Investigator | |
| European Commission, FP7 Marie Curie Career Integration Grant | 2012-2016 |
| Project name: Contextual Text Mining from the Biomedical Scientific Literature | |
| Principal Investigator | |

Bogazici University Research Fund, BAP Project name: Named Entity Disambiguation using Sequence-to-Sequence Models Principal Investigator	2018-2019
Bogazici University Research Fund, BAP Project name: Protein-Ligand Database Construction Principal Investigator	2016-2018
Bogazici University Research Fund, BAP Project name: Named Entity Recognition and Hashtag Segmentation in Social Media Text Principal Investigator	2015-2017
Bogazici University Research Fund, BAP Project name: Jointly Self-trained Parsers Principal Investigator	2012-2013

CONTRIBUTED SOFTWARE PROJECTS

Vapur: A Search Engine to Find Related Protein – Compound Pairs in COVID-19 Literature

<https://tabilab.cmpe.boun.edu.tr/vapur/>

WarmMolGen (A Tool for Protein Target-based Chemical Molecule Generation)

<https://github.com/boun-tabi/biochemical-lms-for-drug-design>

TULAP (The Turkish Language Processing Platform)

<https://tulap.cmpe.boun.edu.tr>

BOUN-Pars: A Dependency Parser for Turkish

<https://tabilab.cmpe.boun.edu.tr/BOUN-PARS/index.html>

BIOSSES: Biomedical Semantic Similarity Estimation System

<http://tabilab.cmpe.boun.edu.tr/BIOSSES/>

Hashtag Segmentation Tool

http://tabilab.cmpe.boun.edu.tr/projects/hashtag_segmentation/

PHISTO (Pathogen-Host Interaction Search Tool)

<http://www.phisto.org>

Ignet (Integrated gene network)

<https://ignet.org>

MIMI (Michigan Molecular Interactions)

<http://mimi.ncibi.org>

TEACHING EXPERIENCE

Boğaziçi University, İstanbul, Turkey Fall 2011-Present
Faculty Member, Department of Computer Engineering

- *CMPE 150 – Introduction to Computing (C)* (Spring 2013, Fall 2013, Spring 2014, Fall 2014, Spring 2015, Fall 2015, Spring 2016, Fall 2016, Spring 2017, Spring 2019, Fall 2019, Spring 2020, Fall 2020, Spring 2021)
- *CMPE 150 – Introduction to Computing (Java)* (Fall 2011, Spring 2012, Fall 2012)
- *CMPE 220 – Discrete Computational Structures* (Fall 2017)
- *CMPE 321 – Introduction to Database Systems* (Spring 2021, Spring 2022)
- *CMPE 493 – Introduction to Information Retrieval* (Fall 2014, Fall 2015, Spring 2017, Spring 2018, Spring 2019, Spring 2020, Fall 2020, Fall 2021)
- *CMPE 549 – Bioinformatics* (Fall 2019, Fall 2022)
- *CMPE 561 – Natural Language Processing* (Fall 2013, Spring 2016)

- *CMPE 59H – Bioinformatics* (Fall 2011, Fall 2012, Spring 2014, Spring 2015, Fall 2016, Fall 2017, Fall 2018)
- *CMPE 59K – Information Retrieval* (Spring 2013)

Boğaziçi University, İstanbul, Turkey

Summer 2011

Visiting Instructor, Department of Computer Engineering

- *CMPE 493 – Introduction to Information Retrieval* (Summer 2011)

İstanbul Technical University, İstanbul, Turkey

Spring 2011-Sumer 2011

Instructor, Faculty of Computer and Informatics

- *BLG 381E – Advanced Data Structures* (Summer 2011)
- *BLG 252 – Object Oriented Programming* (Spring 2011)
- *BLG 540E – Text Retrieval Systems* (Spring 2011)

PROFESSIONAL ACTIVITIES

- Organizing Committee Member, RECOMB 2023.
- Vice Chair for the IAPR Technical Committee on Pattern Recognition for Bioinformatics and Digital Health, 2022-2024.
- Research Topic Editor: *Frontiers in Molecular Biosciences*, Special Issue on Machine Learning Methodologies To Study Molecular Interactions, 2020-2021.
- Editorial Board Member: *Turkish Journal of Electrical Engineering and Computer Sciences* (2019-2020).
- Editorial Advisory Board Member: *Heliyon Journal*, Elsevier, 2015-2018.
- Research Topic Editor: *Frontiers in Microbiology*, Special Issue on Computational Systems Biology of Pathogen-Host Interactions, 2015.
- Program Committee Member: Served for several conferences including ISMB, ECCB, AAI, IJCAI, EMNLP, ACL, ECAI, COLING, HIBIT, SIU.
- Reviewer: Served for several journals including *Bioinformatics*, *BMC Bioinformatics*, *PLOS Computational Biology*, *PLOS ONE*, *Scientific Reports*, *Computer Speech and Language*, *Natural Language Engineering*, *Turkish Journal of Electrical Engineering and Computer Sciences*, *Journal of Bioinformatics and Computational Biology*, *Neural Networks*.