Why “Learn”?

- Machine learning is programming computers to optimize a performance criterion using example data or past experience.
- There is no need to “learn” to calculate payroll.
- Learning is used when:
  - Human expertise does not exist (navigating on Mars),
  - Humans are unable to explain their expertise (speech recognition),
  - Solution changes in time (routing on a computer network),
  - Solution needs to be adapted to particular cases (user biometrics).
What We Talk About When We Talk About “Learning”

- Learning general models from a data of particular examples
- Data is cheap and abundant (data warehouses, data marts); knowledge is expensive and scarce.
- Example in retail: Customer transactions to consumer behavior:
  
  People who bought “Blink” also bought “Outliers” (www.amazon.com)

- Build a model that is a good and useful approximation to the data.
Data Mining

- **Retail:** Market basket analysis, Customer relationship management (CRM)
- **Finance:** Credit scoring, fraud detection
- **Manufacturing:** Control, robotics, troubleshooting
- **Medicine:** Medical diagnosis
- **Telecommunications:** Spam filters, intrusion detection
- **Bioinformatics:** Motifs, alignment
- **Web mining:** Search engines
- ...
What is Machine Learning?

- Optimize a performance criterion using example data or past experience.
- Role of Statistics: Inference from a sample
- Role of Computer science: Efficient algorithms to
  - Solve the optimization problem
  - Representing and evaluating the model for inference
Applications

- Association
- Supervised Learning
  - Classification
  - Regression
- Unsupervised Learning
- Reinforcement Learning
Learning Associations

- Basket analysis:
  \[ P(Y \mid X) \] probability that somebody who buys \( X \) also buys \( Y \) where \( X \) and \( Y \) are products/services.

Example: \( P(\text{chips} \mid \text{beer}) = 0.7 \)
Classification

- Example: Credit scoring
- Differentiating between low-risk and high-risk customers from their income and savings

**Discriminant:** IF income > $\theta_1$ AND savings > $\theta_2$ THEN low-risk ELSE high-risk
Classification: Applications

- Aka Pattern recognition
- Face recognition: Pose, lighting, occlusion (glasses, beard), make-up, hair style
- Character recognition: Different handwriting styles.
- Speech recognition: Temporal dependency.
- Medical diagnosis: From symptoms to illnesses
- Biometrics: Recognition/authentication using physical and/or behavioral characteristics: Face, iris, signature, etc
- ...

Lecture Notes for E Alpaydın 2010 Introduction to Machine Learning 2e © The MIT Press (V1.0)
Face Recognition

Training examples of a person

Test images

ORL dataset,
AT&T Laboratories, Cambridge UK
Regression

- Example: Price of a used car
- $x$: car attributes
- $y$: price
  
  $y = g(x | \theta)$

$g(\cdot)$ model,

$\theta$ parameters
Regression Applications

- Navigating a car: Angle of the steering
- Kinematics of a robot arm
- Response surface design

\[ \alpha_1 = g_1(x, y) \]
\[ \alpha_2 = g_2(x, y) \]
Supervised Learning: Uses

- **Prediction of future cases**: Use the rule to predict the output for future inputs
- **Knowledge extraction**: The rule is easy to understand
- **Compression**: The rule is simpler than the data it explains
- **Outlier detection**: Exceptions that are not covered by the rule, e.g., fraud
Unsupervised Learning

- Learning “what normally happens”
- No output
- Clustering: Grouping similar instances
- Example applications
  - Customer segmentation in CRM
  - Image compression: Color quantization
  - Bioinformatics: Learning motifs
Reinforcement Learning

- Learning a policy: A sequence of outputs
- No supervised output but delayed reward
- Credit assignment problem
- Game playing
- Robot in a maze
- Multiple agents, partial observability, ...
Resources: Datasets

- Statlib: [http://lib.stat.cmu.edu/](http://lib.stat.cmu.edu/)
Resources: Journals

- Journal of Machine Learning Research [www.jmlr.org](http://www.jmlr.org)
- Machine Learning
- Neural Computation
- Neural Networks
- IEEE Transactions on Neural Networks
- IEEE Transactions on Pattern Analysis and Machine Intelligence
- Annals of Statistics
- Journal of the American Statistical Association
- ...
Resources: Conferences

- International Conference on Machine Learning (ICML)
- European Conference on Machine Learning (ECML)
- Neural Information Processing Systems (NIPS)
- Uncertainty in Artificial Intelligence (UAI)
- Computational Learning Theory (COLT)
- International Conference on Artificial Neural Networks (ICANN)
- International Conference on AI & Statistics (AISTATS)
- International Conference on Pattern Recognition (ICPR)
- ...