CS 622 Advanced Machine Learning

Dr. Waseem Shahzad
Professor
Department of Computer Science

Reading Materials

- **Recommended Supplementary Texts:**

- Ian H. Witten & Eibe Frank, *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*, Morgan Kaufmann, 2013.
- T. Hastie, R. Tibshirani, & J. H. Friedman, *The Elements of Statistical Learning : Data Mining, Inference, and Prediction* Springer Verlag, 2001.
- Richard O. Duda, Peter E. Hart, & David G. Stork, *Pattern Classification* Wiley-Interscience, 2000.
- P. Langley, *Elements of Machine Learning*, Morgan Kaufman Publishers, San Fancisco, CA, 1995.
- S. M. Weiss & C. A. Kulikowski, *Computer Systems that Learn*, Morgan Kaufman Publishers, San Fancisco, CA, 1991.
- J. W. Shavlik & T. G. Dietterich (Eds.), *Readings in Machine Learning*, Morgan Kaufman Publishers, San Fancisco, CA, 1990.

Evaluation

Grading

Project	15
Student Presentations	5
Assignments + Quiz	10
Sessional Exams (2)	30
Final Exam	40

	100

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.

What is machine learning

- Making the computers perform the intelligent tasks you performed is Artificial Intelligence
- Learning to perform these tasks using existing data is called Machine Learning
- In short, Machine Learning is learning from DATA

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

Machine Learning

- 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
- …

Applications

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

Learning Associations

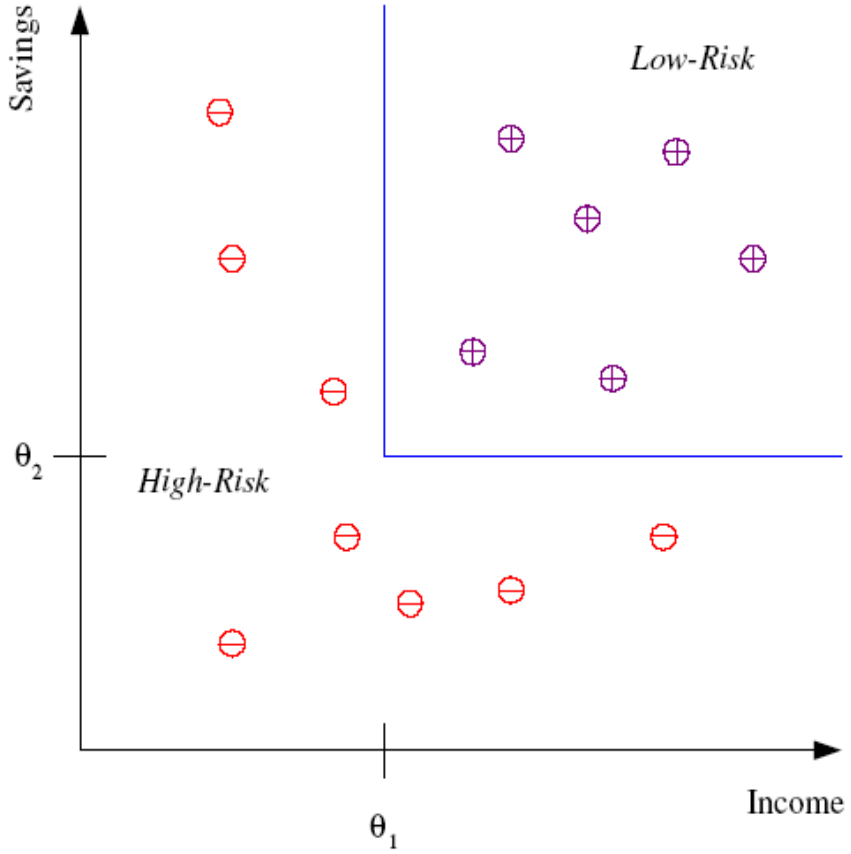
- Basket analysis:

  $P(Y \mid X)$ probability that somebody who buys X also buys Y where X and Y are products/services.

  Example: $P(\text{chips} \mid \text{carbonated\_drink}) = 0.7$

Classification

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



Discriminant: IF *income* > θ_1 AND *savings* > θ_2
THEN low-risk ELSE high-risk

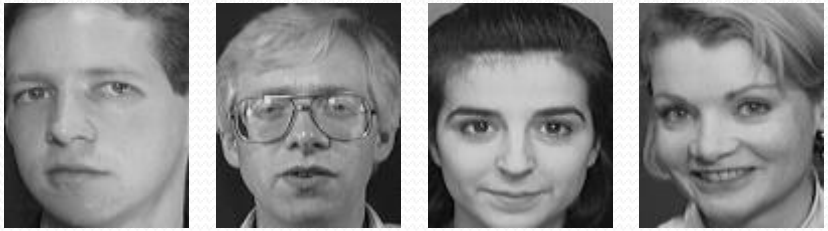
Classification: Applications

- 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
- ...

Face Recognition

Training examples of a person

Test images



ORL dataset,
AT&T Laboratories, Cambridge UK

Regression

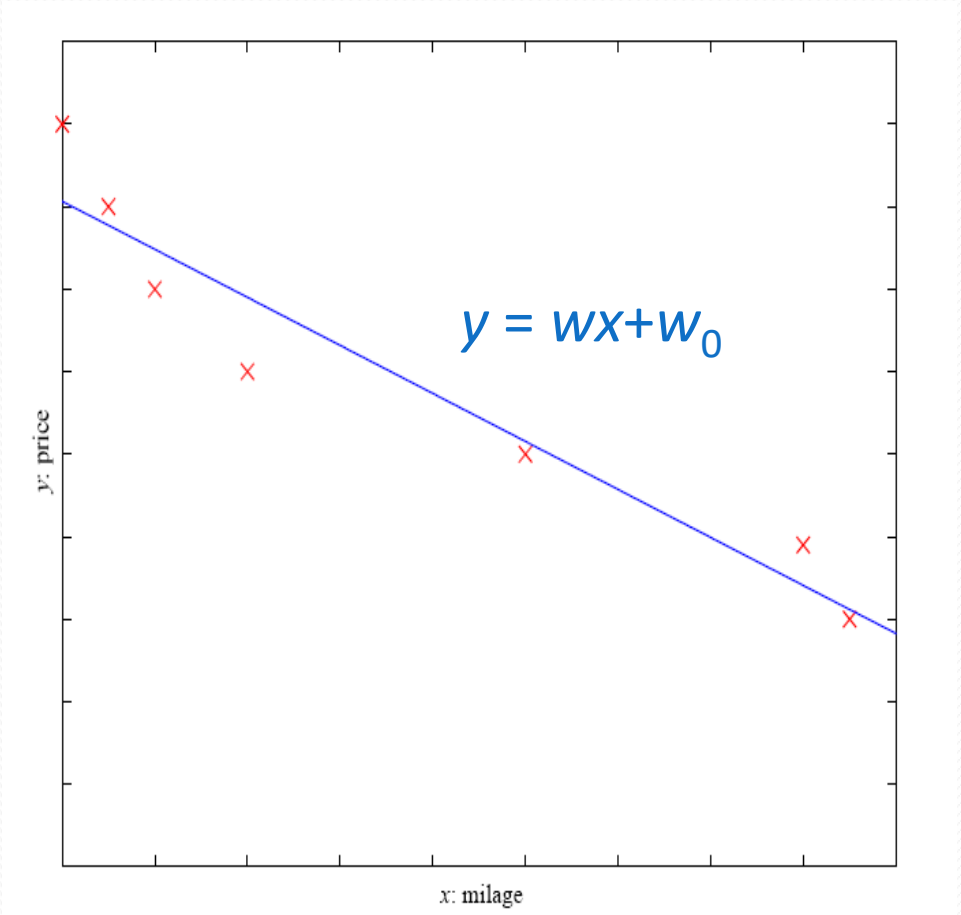
- Example: Price of a used car
- x : car attributes

  y : price

  $$y = g(x \mid \theta)$$

  $g(\ )$ model,

  θ parameters



Supervised Learning: Uses

- Prediction of future cases:
  - Forecasting: Predicting stock prices, weather conditions, or sales trends.
  - Healthcare: Predicting patient outcomes, such as disease progression or treatment effectiveness.
  - Customer Behavior: Anticipating churn, purchase likelihood, or lifetime value of a customer.
  - Supply Chain: Predicting demand to optimize inventory levels.
  - Financial Services: Risk assessment for loan approvals or credit scoring.
- Knowledge extraction:
  - Decision Support: Generating rules or patterns that help doctors, lawyers, or analysts make decisions (e.g., identifying symptoms linked to a disease).
  - Business Intelligence: Uncovering actionable insights from customer reviews, feedback, or sales data.
  - Education: Understanding factors contributing to student success or dropout rates.
  - Agriculture: Identifying optimal farming practices based on weather, soil, and crop yield data.

Supervised Learning: Uses

- Compression:
  - Data Simplification: Creating summary representations of large datasets, such as dimensionality reduction with techniques like PCA or t-SNE.
  - Image and Video Compression: Using machine learning to compress multimedia content while retaining quality.
  - Model Simplification: Building simpler models that encapsulate complex datasets for easier interpretation.
  - Natural Language Processing (NLP): Summarizing large text datasets or documents into concise representations.
- Outlier detection
  - Fraud Detection: Identifying fraudulent transactions in banking, e-commerce, or insurance claims.
  - Anomaly Detection: Spotting unusual patterns in manufacturing processes, like equipment failures.
  - Cybersecurity: Detecting network intrusions or malicious activity.
  - Public Health: Identifying unusual outbreaks or rare disease cases in epidemiology.

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

- UCI Repository: http://www.ics.uci.edu/~mlearn/MLRepository.html
- UCI KDD Archive: http://kdd.ics.uci.edu/summary.data.application.html
- Statlib: http://lib.stat.cmu.edu/
- Delve: http://www.cs.utoronto.ca/~delve/

Resources: Journals

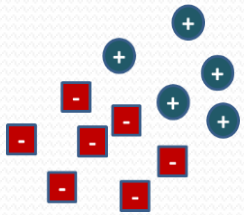
- Journal of Machine Learning Research www.jmlr.org
- Machine Learning
- Neural Computation
- Neural Networks
- IEEE Transactions on Neural Networks
- IEEE Transactions on Pattern Analysis and Machine Intelligence
- Annals of Statistics
- Journal of the American Statistical Association
- ...

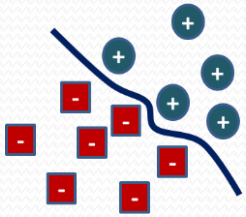
Resources: Conferences

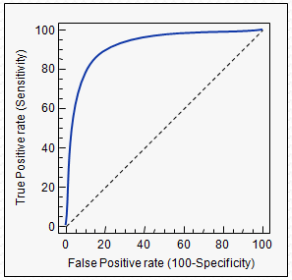
- International Conference on Machine Learning (ICML)
- European Conference on Machine Learning (ECML)
- Neural Information Processing Systems (NIPS)
- Uncertainty in Artificial Intelligence (UAI)
- Computational Learning Theory (COLT)
- International Conference on Artificial Neural Networks (ICANN)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

A typical machine learning system

Identify a Problem → Data collection and feature extraction → Model Training → Evaluation → Deployment







Webserver
Software Package