Introduction

Suman Kumar

Department of computer Science

Troy University

Slides mostly follow "Introduction to Machine Learning, 3rd edition Ethen Alpaydin

Some parts of the material can be found in Machine learning with Scikit-Learn & tensor flow by Aurelien Geron

Machine Learning

- Field of study that gives computers the ability to learn without being explicitly programmed. – ***Arthur Samuel, 1959***

- A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience – ***Tom Mitchell, 1997***

Traditional Approach vs Machine Learning

- Write rules-> evaluate -> analyze -> study-> write rules
- Train ML(data)-> evaluate-> analyze-> study->trainML(Data)
- Examples: Spam filter, Speech to text, text identification etc.
- Humans can learn from Machine Learning.

Machine 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 do we mean by learning ?

- Learning general models from a data of particular examples
- 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

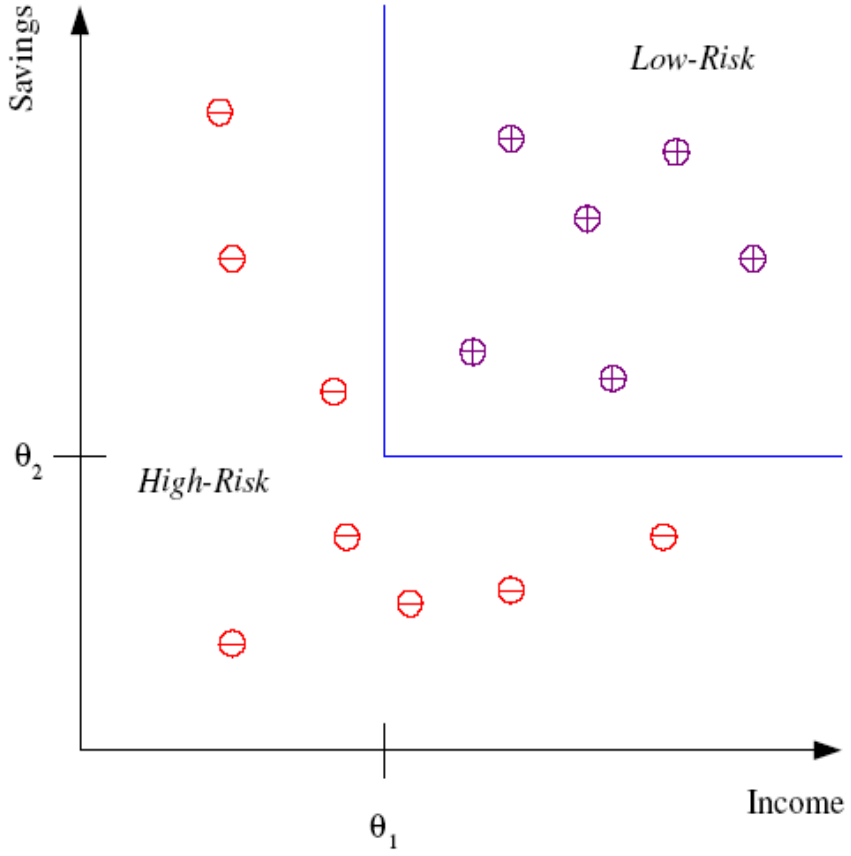
- Retail: Market basket analysis, Customer relationship management (CRM)
- Finance: Credit scoring, fraud detection
- Manufacturing: Control, robotics, troubleshooting
- Medicine: Medical diagnosis
- Telecommunications: Spam filters, intrusion detection
- Web mining: Search engines
- …

Different types of learnings

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

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
- ...

Regression

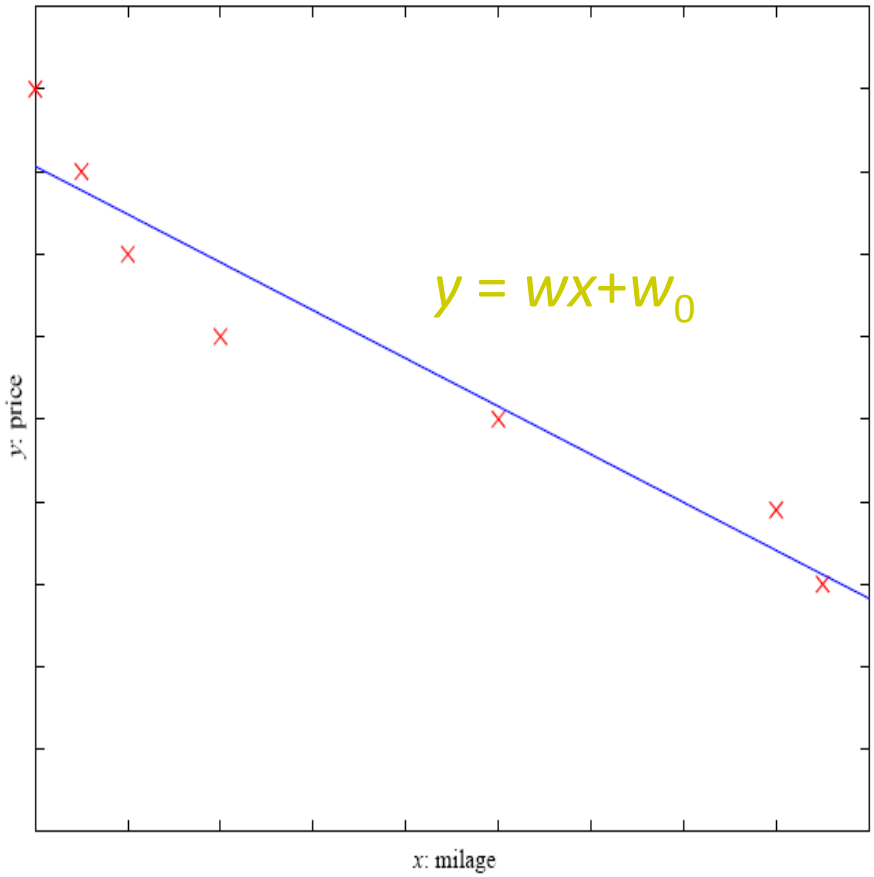
- Example: Price of a used car
- x : car attributes

  y : price

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

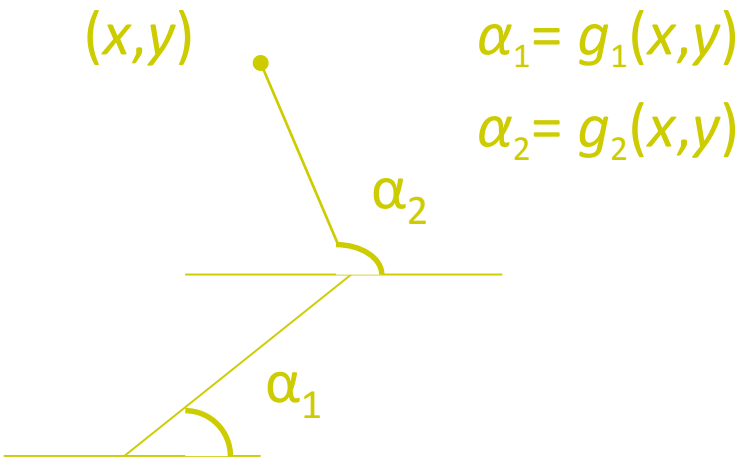
  $g(\ )$ model,

  θ parameters



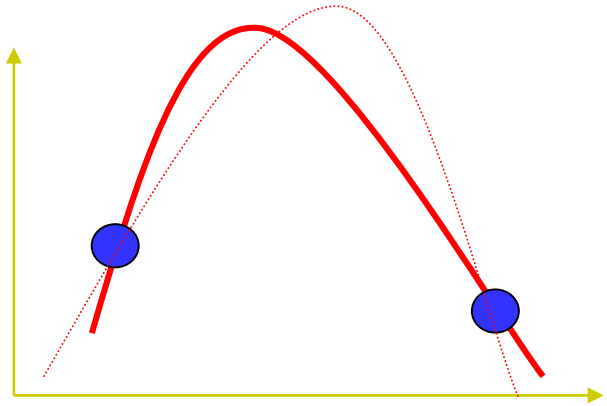
Regression Applications

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



$\alpha_1 = g_1(x,y)$

$\alpha_2 = g_2(x,y)$



- Response surface design

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

Supervised Learning: Algorithms

- K-nearest neighbors
- Linear regression
- Logistic regression
- Support Vector Machines (SVMs)
- Decision Trees and Random Forests
- Neural Networks

Unsupervised Learning

- Learning “what normally happens”
- No output
- Clustering: Grouping similar instances
- Example applications
  - Segmentation
  - Anomaly detection
  - Dimensionality reduction

Unsupervised Learning: algorithms

- Clustering
  - K-Means
  - Hierarchical Cluster Analysis
  - Expectation maximization
- Visualization and dimensionality reduction
  - Principal component analysis (PCA)
  - Kernel PCA
  - t-distributed stochastic neighbor embedding (t-SNE)
- Association rule learning
  - Apriori
  - Eclat

Semisupervised Learning

- Algorithms for partially labeled data
- Example application: Photo hosting services
- Most algorithms are combination of supervised and unsupervised learning algorithm

Reinforcement Learning

- Learning a policy: A sequence of outputs
- No supervised output but delayed reward
- Systems are called agents
- Observe->select action from policty-> action-> get reward/penalty-> update policy .. repeate
- Example applications:
  - Credit assignment problem
  - Game playing
  - Robot in a maze
  - Multiple agents, partial observability, ...

Other ML system classifications..

- Batch vs online
  - Batch: must use all the available data
  - Online: system learn incrementally by using data instances sequentially
- Instance vs Model based
  - Instance: Learn the examples by heart, then generalize to new cases using similarity measures
  - Model: Use the model to generalize/predict

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)
- …