

Lecture Slides for

INTRODUCTION TO

# Machine Learning

ETHEM ALPAYDIN  
© The MIT Press, 2004

[alpaydin@boun.edu.tr](mailto:alpaydin@boun.edu.tr)  
<http://www.cmpe.boun.edu.tr/~ethem/i2ml>

CHAPTER 1:

## Introduction

### 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 “Da Vinci Code” also bought “The Five People You Meet in Heaven” (www.amazon.com)*
- Build a model that is a *good and useful approximation* to the 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

5

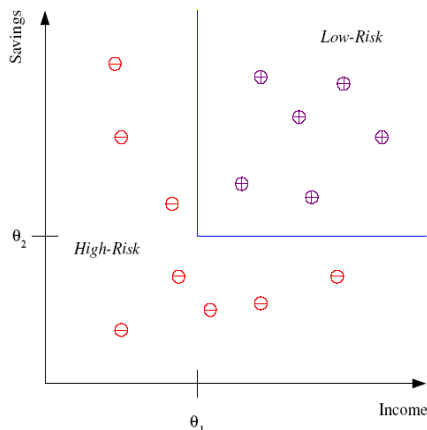
# Applications

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

6

## 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**

7

## 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.
  - Use of a dictionary or the syntax of the language.
  - Sensor fusion: Combine multiple modalities; eg, visual (lip image) and acoustic for speech
- Medical diagnosis: From symptoms to illnesses
- ...

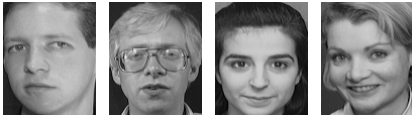
8

# Face Recognition

Training examples of a person



Test images

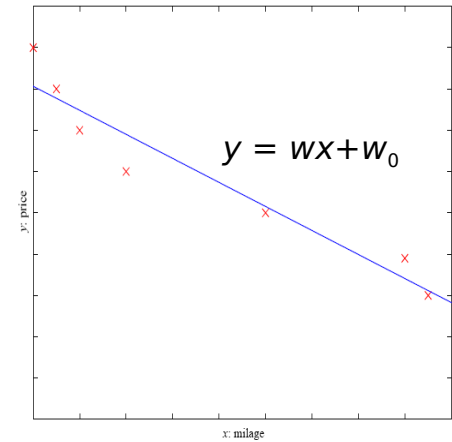


AT&T Laboratories, Cambridge UK  
<http://www.uk.research.att.com/facedatabase.html>

9

# Regression

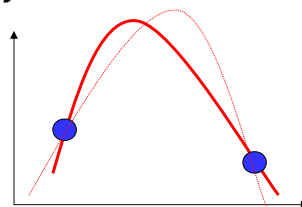
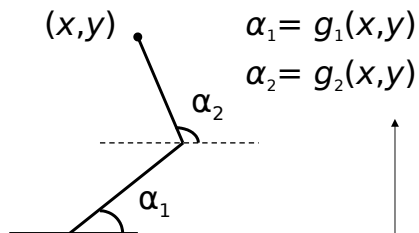
- Example: Price of a used car
- $x$  : car attributes
- $y$  : price
- $y = g(x | \theta)$
- $g(\cdot)$  model,
- $\theta$  parameters



10

# Regression Applications

- Navigating a car: Angle of the steering wheel (CMU NavLab)
- Kinematics of a robot arm



- Response surface design

11

# 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

12

## 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

13

## 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, ...

14

## Resources: Datasets

- UCI Repository: <http://www.ics.uci.edu/~mllearn/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/>

15

## 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
- ...

16

# *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 Joint Conference on Artificial Intelligence (IJCAI)
- International Conference on Neural Networks (IJCNN)
- International Conference on Pattern Recognition (ICPR)