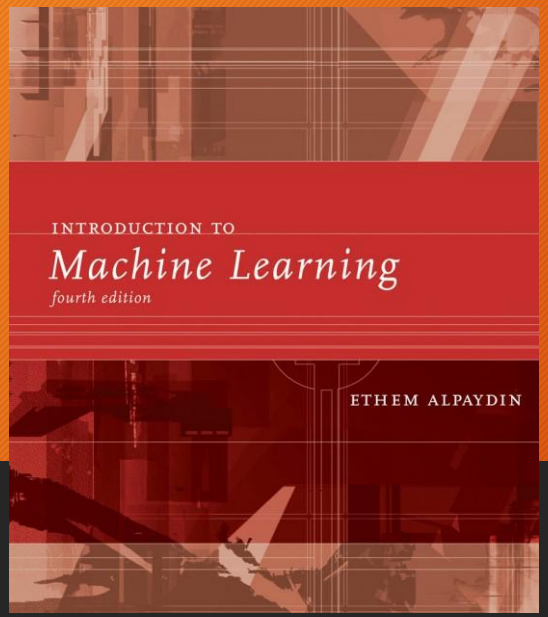


Lecture Slides for

INTRODUCTION TO Machine Learning 4e

ETHEM ALPAYDIN
The MIT Press, 2020

ethem.alpaydin@gmail.com

Troy University Resources

CHAPTER 1:

Introduction

Big Data

- Widespread use of personal computers and wireless communication leads to “big data”
- We are both producers and consumers of data
- Data is not random, it has structure, e.g., customer behavior
- We need “big theory” to extract that structure from data for

  (a) Understanding the process

  (b) Making predictions for the future
- Cheaper computational power (e.g., GPUs).

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

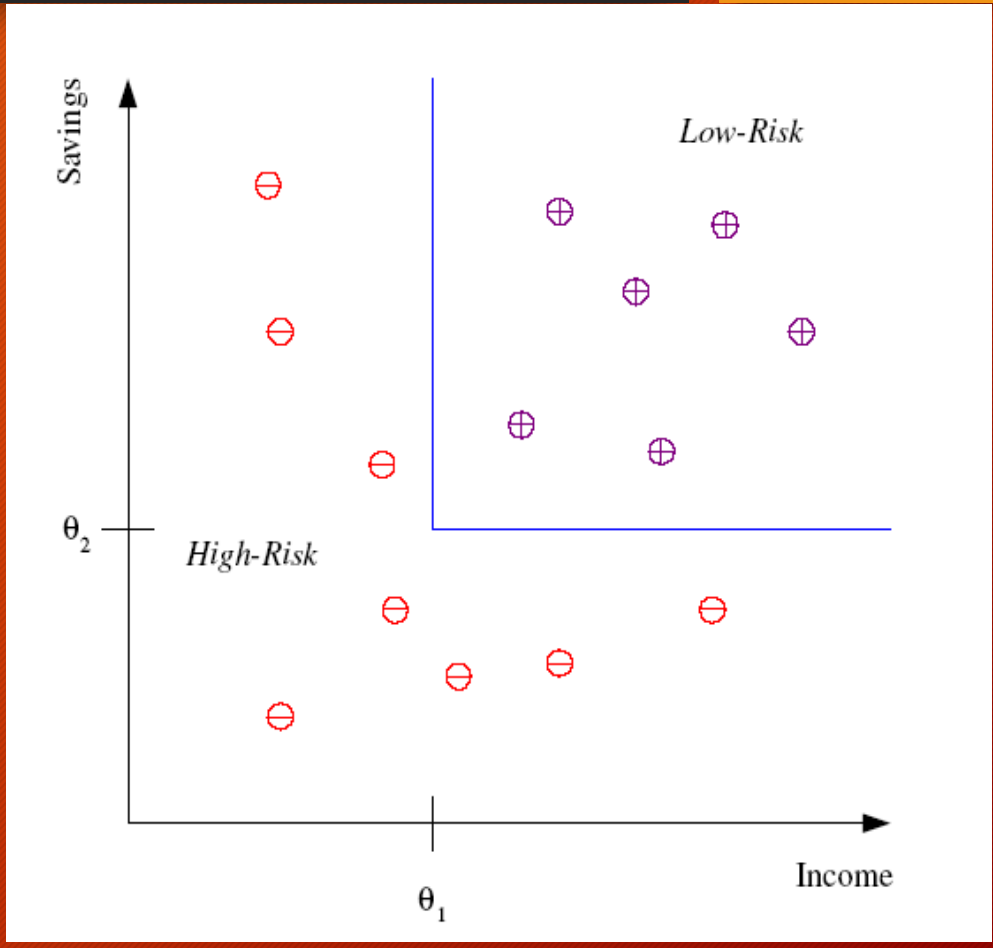
- Basket analysis:

  $P(Y \mid X)$ 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* > 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
- Outlier/novelty detection: Detecting faults

Face Recognition

Training examples of a person



Test images



ORL dataset,
AT&T Laboratories, Cambridge UK

Regression

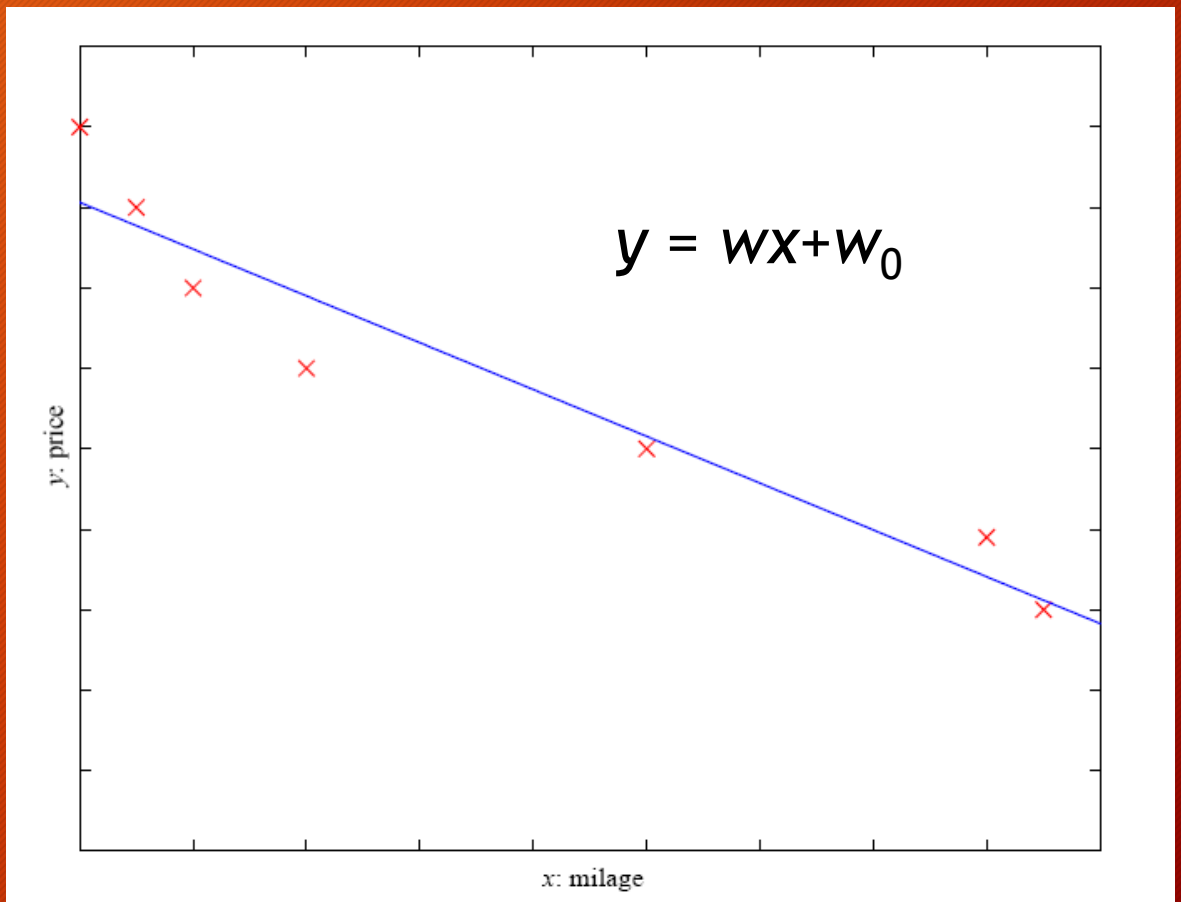
- Example: Price of a used car
- x : car attributes

  y : price

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

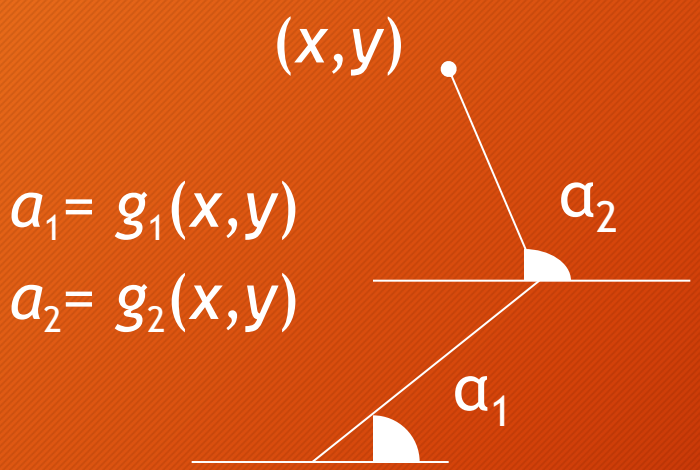
  $g()$, model (fixed)

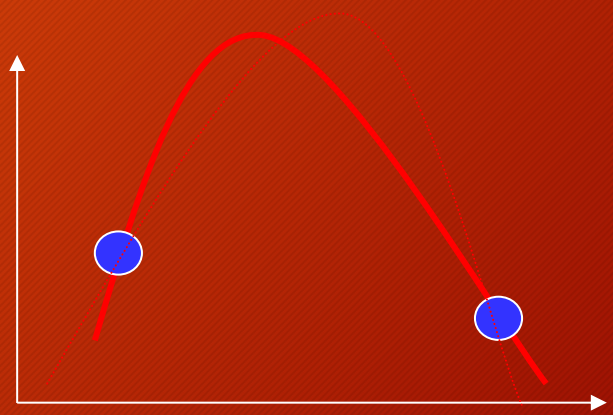
  θ, parameters (modifiable->learned)



Regression Applications

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





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

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, ...

Related Topics

- High-performance computing:
  - Parallel processing, GPUs
- Data privacy and security:
  - Sanitization/anonymization of personal data
  - privacy-preserving ML
- Model interpretability and trust:
  - The importance of testing
  - Bias in training data, fairness of decisions
  - Adversarial examples
  - We do not want black-box but explainable models

Data Science

- A new field composed of
  1. Machine learning (theoretical component)
  2. High-performance computation (engineering component); both hardware and software aspects
  3. Ethical/legal aspects of data collection, analysis, and data-driven decision making (social component)