https://xkcd.com/894/

1

INTRODUCTION TO MACHINE LEARNING

David Kauchak
CS 158 – Spring 2022

2

Introductions

Dr. | Prof | Professor
Dave | Kauchak

Pronouns: he/him/his

3

Why are you here?

What is Machine Learning?

Why are you taking this course?

What topics would you like to see covered?

4

Machine Learning is…

Machine learning is a subfield of computer science that evolved from the study of pattern recognition and computational learning theory in artificial intelligence.

5

Machine Learning is…

Machine learning is programming computers to optimize a performance criterion using example data or past experience.

-- Ethem Alpaydin

The goal of machine learning is to develop methods that can automatically detect patterns in data, and then to use the uncovered patterns to predict future data or other outcomes of interest.

-- Kevin P. Murphy

The field of pattern recognition is concerned with the automatic discovery of regularities in data through the use of computer algorithms and with the use of these regularities to take actions.

-- Christopher M. Bishop

6

Machine Learning is…

Machine learning is about predicting the future based on the past.

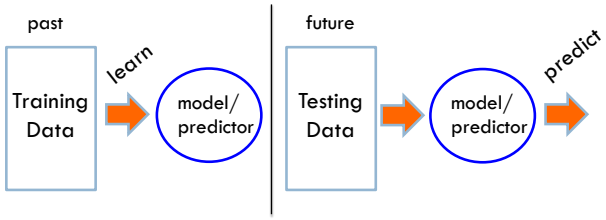
-- Hal Daume III

7

Machine Learning is…

Machine learning is about predicting the future based on the past.

-- Hal Daume III



8

Machine Learning, aka

data mining: data analysis, not prediction, though often involves some shared techniques

inference and/or *estimation* in statistics

pattern recognition in engineering

signal processing in electrical engineering

induction

optimization

9

Goals of the course: learn about…

Different machine learning problems

Common techniques/tools used
- theoretical understanding
- practical implementation

Proper experimentation and evaluation

Dealing with large (huge) data sets
- Parallelization frameworks
- Programming tools

10

Goals of the course

Be able to laugh at these signs
(or at least know why one might…)

11

Administrative

Course page:
http://www.cs.pomona.edu/classes/cs158/

Assignments
- Weekly
- Mostly programming (Java, mostly)
- Some written/write-up
- Generally due Sunday evenings

Two "midterm" exams and one final (all time limited take home)

Late Policy

Collaboration

12

Course expectations

Plan to stay busy!

Applied class, so lots of programming

Machine learning involves math

13

Other things to note

Videos before class

Lots of class participation!

Read the book (it's good)

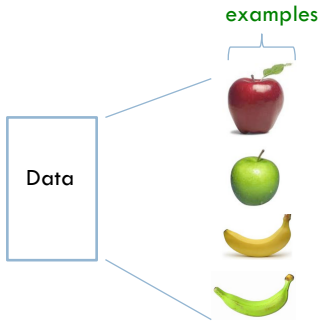
14

Machine learning problems

What high-level machine learning problems have you seen or heard of before?

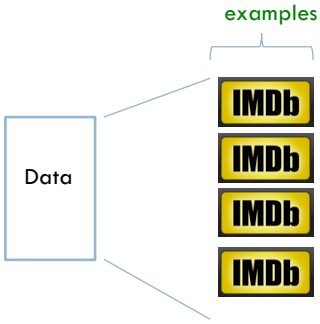
15

Data



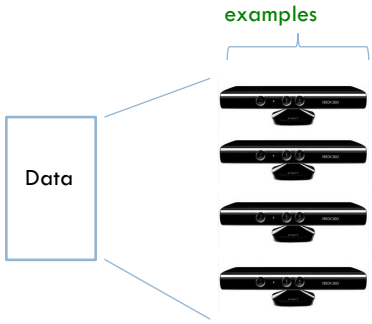
16

Data



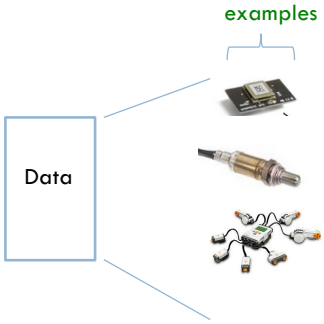
17

Data



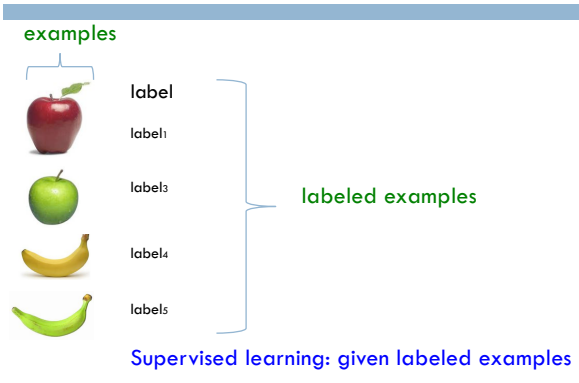
18

Data



19

Supervised learning



20

Supervised learning

label

label1

label3

label4

label5

model/ predictor

Supervised learning: given labeled examples

21

Supervised learning

model/ predictor

predicted label

Supervised learning: learn to predict new example

22

Supervised learning: classification

label

apple

apple

banana

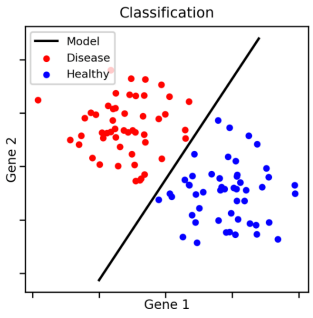
banana

Classification: a finite set of labels

Supervised learning: given labeled examples

23

Classification Example



24

Classification Applications

Face recognition

Character recognition

Spam detection

Medical diagnosis: From symptoms to illnesses

Biometrics: Recognition/authentication using physical and/or behavioral characteristics: Face, iris, signature, etc

…

25

Supervised learning: regression

label

-4.5

10.1

3.2

4.3

Regression: label is real-valued

Supervised learning: given labeled examples

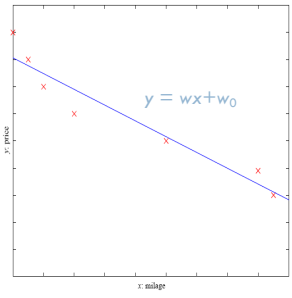
26

Regression Example

Price of a used car

x : car attributes
(e.g. mileage)

y : price



27

Regression Applications

Economics/Finance: predict the value of a stock

Epidemiology

Car/plane navigation: angle of the steering wheel, acceleration, …

Temporal trends: weather over time

…

28

Supervised learning: ranking

label

1

4

2

3

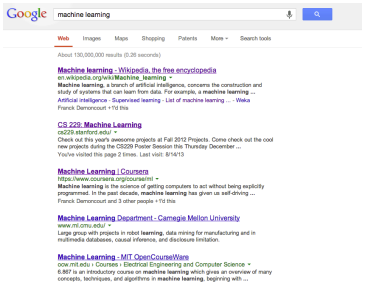
Ranking: label is a ranking

Supervised learning: given labeled examples

29

Ranking example

Given a query and a set of web pages, rank them according to relevance



30

Ranking Applications

User preference, e.g. movie ranking
iTunes

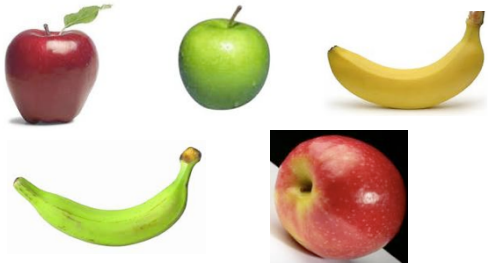
flight search (search in general)

reranking N-best output lists

…

31

Unsupervised learning



Unupervised learning: given data, i.e. examples, but no labels

32

Unsupervised learning applications

learn clusters/groups without any label

customer segmentation (i.e. grouping)

image compression

bioinformatics: learn motifs

…

33

Reinforcement learning

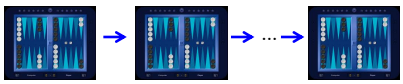
left, right, straight, left, left, left, straight	GOOD
left, straight, straight, left, right, straight, straight	BAD
left, right, straight, left, left, left, straight	18.5
left, straight, straight, left, right, straight, straight	-3

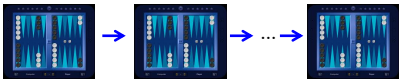
Given a *sequence* of examples/states and a *reward* after completing that sequence, learn to predict the action to take for an individual example/state

34

Reinforcement learning example

Backgammon

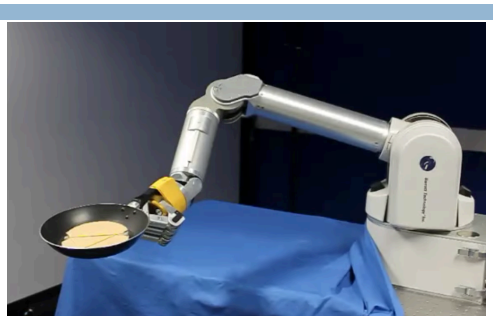




Given sequences of moves and whether or not the player won at the end, learn to make good moves

35

Reinforcement learning example



https://www.youtube.com/watch?v=W_gxLKSsSIE

36

Other learning variations

What data is available:

- Supervised, unsupervised, reinforcement learning
- semi-supervised, active learning, …

How are we getting the data:

- online vs. offline learning

Type of model:

- generative vs. discriminative
- parametric vs. non-parametric

37