

A host of prediction problems

Topics we'll cover

- ① Machine learning versus algorithms
- ② A taxonomy of prediction problems
- ③ Roadmap for the course

Machine learning versus Algorithms

A central goal of both fields:

develop procedures that exhibit a desired input-output behavior.

- **Algorithms:** input-output mapping can be precisely defined.

Input: Graph G , two nodes u, v in the graph.

Output: Shortest path from u to v in G

- **Machine learning:** mapping cannot easily be made precise.

Input: Picture of an animal. \rightarrow *x/y example*

Output: Name of the animal.

*build model from data
of input/output pairs
make prediction from model
machine figures out a mapping between input & output*

Instead, provide examples of (input,output) pairs.

Ask the machine to *learn* a suitable mapping itself.

Prediction problems: inputs and outputs

Basic terminology:

- The input space, \mathcal{X} .
E.g. 32×32 RGB images of animals.
- The output space, \mathcal{Y} .
E.g. Names of 100 animals.

x:



y: "bear"

After seeing a bunch of examples (x, y) , pick a mapping

$$f : \mathcal{X} \rightarrow \mathcal{Y}$$

that accurately recovers the input-output pattern of the examples.

Categorize prediction problems by the type of **output space**:

(1) discrete, (2) continuous, or (3) probability values

Discrete output space: classification

Binary classification

E.g., Spam detection

$\mathcal{X} = \{\text{email messages}\}$

$\mathcal{Y} = \{\text{spam, not spam}\}$

Multiclass

E.g., News article classification

$\mathcal{X} = \{\text{news articles}\}$

$\mathcal{Y} = \{\text{politics, business, sports, ...}\}$

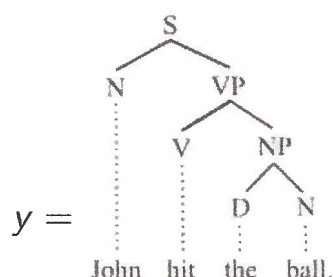
Structured outputs

E.g., Parsing

$\mathcal{X} = \{\text{sentences}\}$

$\mathcal{Y} = \{\text{parse trees}\}$

$x = \text{"John hit the ball"}$



Continuous output space: regression

- Pollution level prediction

Predict tomorrow's air quality index in my neighborhood

$\mathcal{Y} = [0, \infty)$ (< 100 : okay, > 200 : dangerous)

- Insurance company calculations

What is the expected life expectancy of this person?

$\mathcal{Y} = [0, 120]$

What are suitable predictor variables (\mathcal{X}) in each case?

Probability estimation

$\mathcal{Y} = [0, 1]$ represents **probabilities**

Example: Credit card transactions

- x = details of a transaction
- y = probability this transaction is fraudulent

Why not just treat this as a binary classification problem?

Roadmap for the course

- 1 Solving prediction problems
Classification, regression, probability estimation
- 2 Representation learning
Clustering, projection, dictionary learning, autoencoders
- 3 Deep learning