Course Overview

Professor Ameet Talwalkar

Slide Credit: Professor Fei Sha

Outline

What is machine learning?

One possible definition[1]

a set of methods that can automatically *detect patterns* in data, and then use the uncovered patterns to *predict future data*, or to perform other kinds of decision making *under uncertainty*

cf. Murphy's book

Example: detect patterns

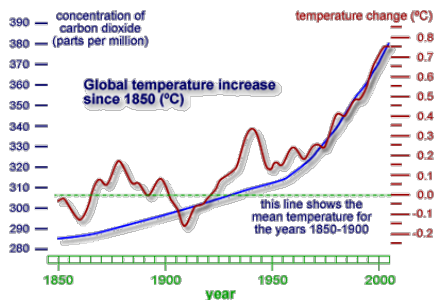**How the temperature has been changing in the last 140 years?**
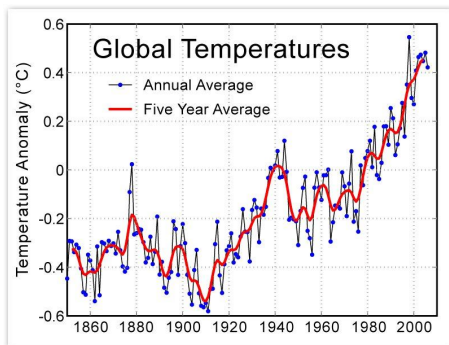
Patterns

- We see repeated periods of fluctuation
- General trend is that temperatures are rising

How do we describe the pattern?

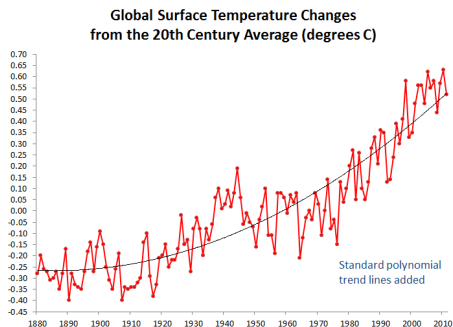**Build a model: fit the data with a polynomial function**

- The model is not accurate for individual years
- But overall, the model captures the major trend

Predicting future

What is temperature of 2010?

**Global Surface Temperature Changes
from the 20th Century Average (degrees C)**

- This particular polynomial model is not exactly accurate for that specific year, but it is pretty close

What we have learned from this example?

Key ingredients in the machine learning task

- Data: collected from past observations (*training data*)
- Modeling: devised to capture the patterns in the data
  - ▸ The model does not have to be true — as long as it is close, it is useful
  - ▸ We should tolerate randomness and mistakes — many interesting things are stochastic by nature.
- Prediction: apply the model to forecast what is going to happen in future

A rich history of applying statistical learning methods
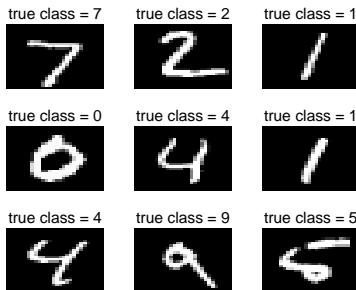
Recognizing flowers (by R. Fisher, 1936)
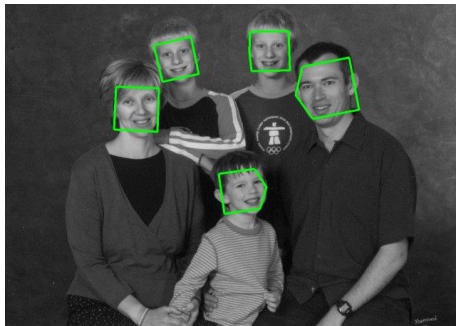Types of Iris: setosa, versicolor, and virginica

Huge success 20 years ago

Recognizing handwritten zipcodes and checks (AT&T Labs, circa late 1990s)

More modern ones, in your social life

Recognizing your friends on Facebook

Learn your preferences

Recommending what you might like

Why is machine learning so hot?

- Flood of data leads to several high-impact applications

- Consumer applications:
  - ▶ speech recognition, information retrieval and search, email and document classification, stock price prediction, object recognition, product recommendation, ⋯
  - ▶ Highly desirable expertise from industry: Google, Facebook, Microsoft, Yahoo, Twitter, IBM, LinkedIn, Amazon, ⋯

- Scientific applications:
  - ▶ Biology and genetics: identify disease-causing genes and gene networks
  - ▶ Climate science: predicting global warming trends
  - ▶ Social science: social network analysis; social media analysis
  - ▶ Business and finance: marketing, operation research
  - ▶ Emerging ones: healthcare, energy, ⋯

What is in machine learning?

Different flavors of learning problems

- Supervised learning: make prediction given labeled training observations, e.g., Spam detection, Iris
- Unsupervised learning: Discover hidden and latent patterns in data; data exploration, e.g., topic modelling in text data
- Reinforcement learning: act optimally (or at least well) under uncertainty, e.g., defining a robot's behavior with the world based on the feedback (rewards / punishments) it receives from each action
- Many other paradigms

The focus and goal of this course

- Supervised learning (before midterm)
- Unsupervised learning (after midterm)

Outline

1 Overview of machine learning

2 About this Course
  • Administration

3 Review of basic math

Quick Polls

- Undergrad / Masters / PhD?
- Computer Science / Other Engineering / Other?
- Registered / Waiting List / Hoping to Register / Auditing?

Registration

- Course is currently full, and we cannot increase the size
- I expect several students will drop the course
- If spots open up, priority will be given to wait list, then to Computer Science students
- Priority will also be given to students based on their scores on the first problem set

Course Logistics

Let's go to the course website:
http://cs.ucla.edu/~ameet/teaching/fall15/cs260/index.html

Any questions?

Outline