CSC 665 Section 2: Machine Learning Theory

Chicheng Zhang
University of Arizona

Logistics

- Instructor: Chicheng Zhang (chichengz@cs.arizona.edu)
- Office Hours: Mon 1-2pm, Gould-Simpson 720
- Website: https://zcc1307.github.io/csc665/

Machine learning (ML)

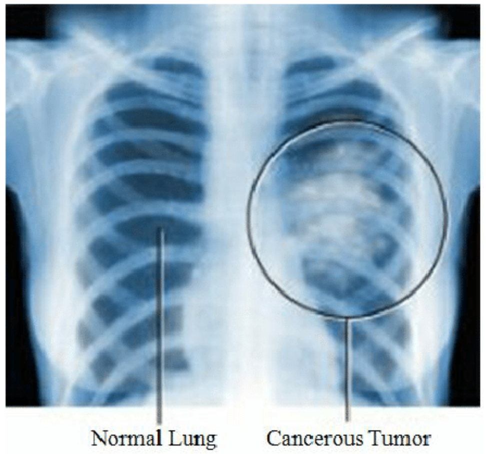
- Algorithms build models that reflect patterns in the data, and use them for decision making
- Examples:
  - Spam classification
  - Image recognition (face, handwriting)
  - Medical applications (diagnosis, treatment assignment)
  - Online advertising
  - Game playing
  - (the list grows continually..)

Settings of ML

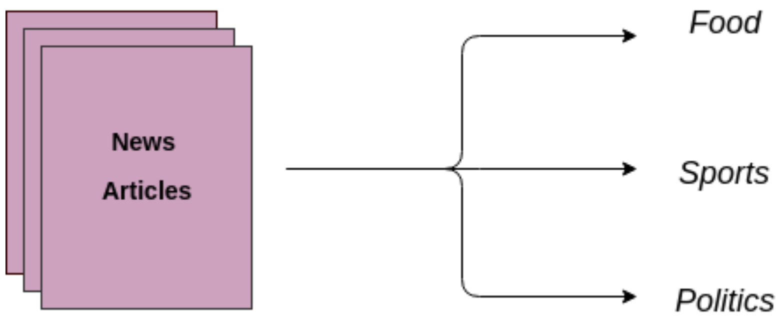
- Supervised learning
- Unsupervised learning
- Interactive learning

Supervised learning

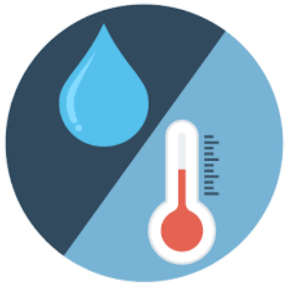
- Classification
  - Binary: image classification



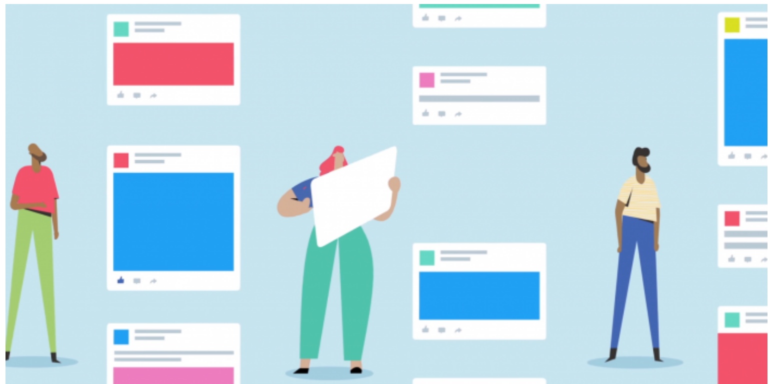
  - Multiclass: text classification



- Regression:
  - Weather forecasting (e.g. temperature / humidity)

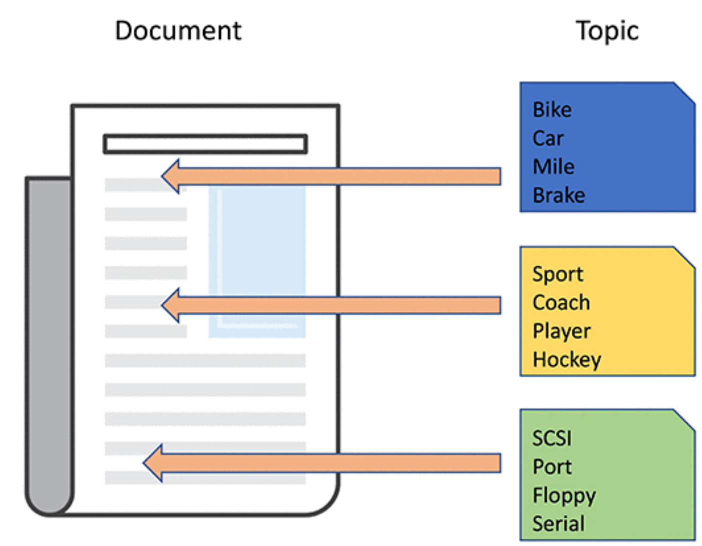


- Rich output:
  - Social media feed ranking

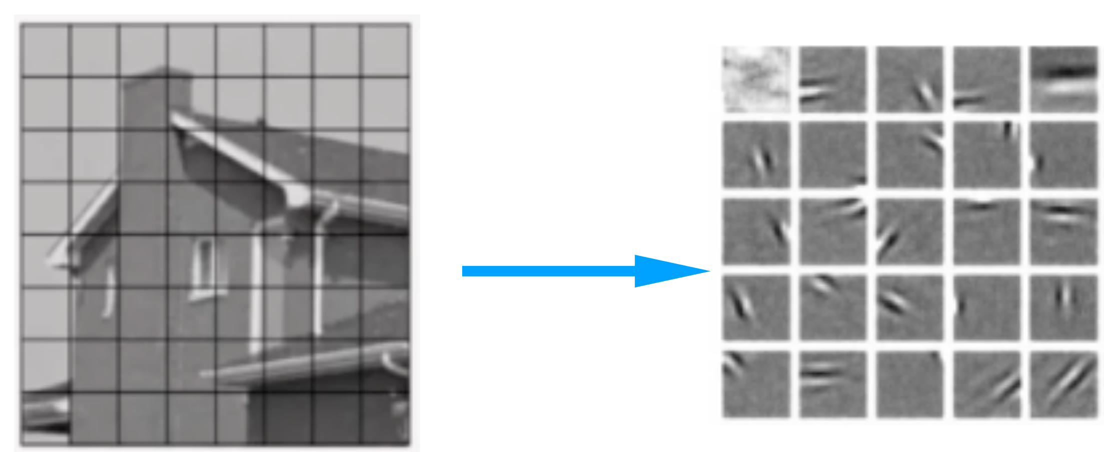


Unsupervised learning

- Clustering
  - Topic modeling



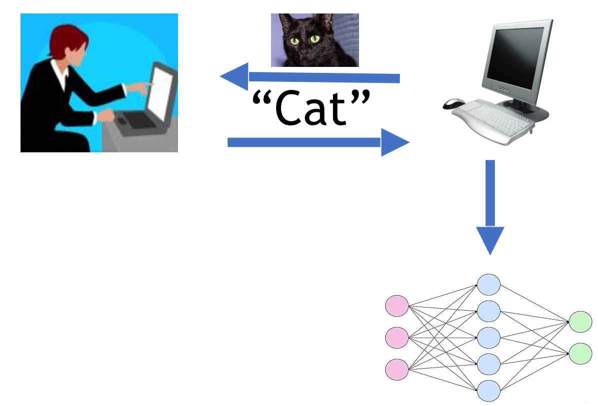
- Feature extraction
  - Image compression



- Change-point detection
  - Trend detection in social media

Interactive learning

- Online learning
  - Spam classification
  - News recommendation
- Active learning
- Reinforcement learning
- Imitation learning
- …



What this course is about

- A rigorous performance analysis on machine learning algorithms
- Performance measure:
  - Sample complexity: #samples used
  - Computational complexity: time/space consumed
- First half: statistical learning (mainly supervised learning)
- Second half: online learning

Example: spam classification

- Email: Free Software! Offer lasts till 2/20 -- spam
- Email: CS Colloquium Lecture Series -- not spam
- Data:

	free	offer	lecture	cs	Spam?
Email 1	1	1	0	0	+1
Email 2	0	0	1	1	-1

- Prediction rule (Linear classifier): sign(9*free + 1*offer – 3.1*lecture)
- How many samples are needed to find a good linear classifier (say, with error < 0.01)?

Course preview

- Part 1: statistical learning
  - The probably approximately correct (PAC) learning model
  - Practical algorithms motivated from theory - SVM, boosting
  - Analysis of (regularized) empirical risk minimization
- Part 2: online learning
  - Online optimization, the regret model
  - Online-to-batch conversion
  - Online learning with partial feedback

Why learning theory?

- Have guiding principles in mind when developing learning algorithms
  - e.g. regularization (early stopping, weight decay, adagrad)
- Motivate new practical learning algorithms
  - e.g. boosting originates in a purely theoretical question
- Understanding the success of practical method (and potentially resulting in better empirical approaches)
  - e.g. explaining the success of learning overparametrized neural nets

Prerequisites

- (Multivariate) Calculus
- Linear algebra
- Probability
- Basic programming
- Additional background (helpful, not strictly required):
  - Numerical optimization, algorithm design and analysis
- Calibration Homework (Due Sep 3 in class)

Evaluation

- Homework (50%)
    - weekly assignments, 1-3 problems / assignment
- Project (15%)
- Midterm (15%): (tentatively Oct. 10)
- Final (20%)

Project

- Groups of 2-4
- Can be one of the following styles:
    - Literature survey (be critical!)
    - Implementation (do your experiments echo the theoretical results? If not, why?)
    - Research
        - Advances state-of-the-art on existing learning models
        - Study new learning paradigm
- Project ideas will be up by mid-September

Project timelines

- Proposal: Oct 24
- Progress report: Nov 14
- Final presentations: Dec 5 and Dec 10, in class
- Final report (4 pages): Dec 11

Thank you!
Questions?

Machine learning courses at UA

- Fall 2019
  - CSC 665 Section 1: Advanced Topics in Probabilistic Graphical Models by Jason Pacheco
  - Probabilistic modeling perspective
- Spring 2020
  - CSC 665: Topics in Online Learning and Bandits by Kwang-Sung Jun
  - Focuses on the online perspective
- Other courses (see course website)