

Machine Learning (CSE 446): Introduction

Sham M Kakade

© 2018

University of Washington
`skakade@cs.washington.edu`

Jan 3, 2018

Learning and Machine Learning?

- ▶ Broadly, what is “learning”?

Wikipedia, “ Learning is the process of acquiring new or modifying existing knowledge, behaviors, skills, values, or preferences. Evidences that learning has occurred may be seen in changes in behavior from simple to complex.”

- ▶ What is “machine learning”?

An AI centric viewpoint: ML is about getting computers to do the types of things people are good at.

- ▶ How is it...

- ▶ different from statistics?
- ▶ different from AI?

(When people say “AI” they almost always mean “ML.”)

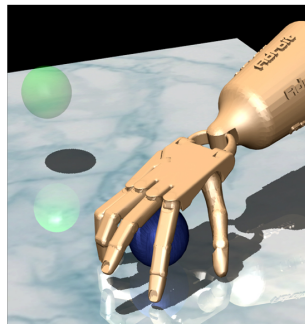
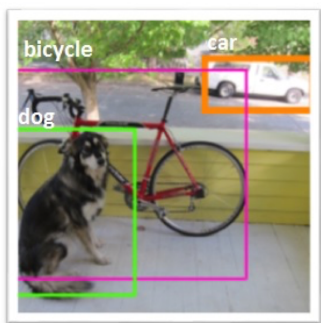
What is ML about?

- ▶ Easy for a computer: $(42384 * 3421.82)^{1/3}$
- ▶ Easy for a child:
 - ▶ speech recognition
 - ▶ object recognition
 - ▶ question/answering (“what color is the sky?”)
- ▶ Computers are designed to execute mathematically precise computational primitives (and they have become much faster!).
- ▶ **This class:** The algorithmic and statistical thinking (and techniques) for how we *train* computers to get better at these more 'easy-for-human' tasks.

ML is starting to work....

ImageNet: 10^6 images.
 10^3 categories

- ▶ No longer just an academic pursuit...
- ▶ Almost “overnight” impacts to society:
(threshold) improvements in performance translate into societal impact



Today, ML is begin used for:

- ▶ Video and image processing
- ▶ Speech and language processing
- ▶ Search engines
- ▶ Robot control
- ▶ Medical and health analysis
- ▶ not just “AI-ish” problems:
sensor networks, traffic navigation, medical imaging, computational biology,
finance

Is it Magic?

- ▶ “sort of, yes”: why is the future (and *never-before-seen* instances) predictable from the past?
“inductive bias” is critical for learning.
- ▶ “in practice, no”: we will examine the algorithmic tools and statistical methods appropriately.
- ▶ “responsibly, NO”: there are consequences and limitations.

Course logistics

Your Instructors

- ▶ Sham Kakade (instructor)

Research interests:

- ▶ theory: rigorous algorithmic and statistical analysis of these methods
- ▶ practice: understanding how to advance the state of the art (robotics, music +comp. vision, NLP)

- ▶ TAs:

Kousuke Ariga, Benjamin Evans, Xingfan Huang, Sean Jaffe, Vardhman Mehta, Patrick Spieker, Jeannette Yu, Kaiyu Zheng.

Info

Course website: <https://courses.cs.washington.edu/courses/cse446/18wi/>


Contact: cse446-staff@cs.washington.edu

Please only use this email for course related questions (unless privacy is needed).

Canvas: https://canvas.uw.edu/courses/1124156/discussion_topics

Office hours: TBA.

Textbooks

- ▶ “A Course in Machine Learning”, Hal Daume.
 - ▶ “Machine Learning: A Probabilistic Perspective”, Kevin Murphy.
- 

Outline of CSE 446

- ▶ Problem formulations: classification, regression
- ▶ Techniques: decision trees, nearest neighbors, perceptron, linear models, probabilistic models, neural networks, kernel methods, clustering
- ▶ “Meta-techniques”: ensembles, expectation-maximization
- ▶ Understanding ML: limits of learning, practical issues, bias & fairness
- ▶ Recurring themes: (stochastic) gradient descent, the “scope” of ML, overfitting

optimization

Grading

- ▶ Assignments (40%)
 - ▶ 5 in total
 - ▶ both mathematics pencil and paper, mostly programming
 - ▶ Graded based on attempt and correctness
 - ▶ Late policy: 33% off for (up to) one day late; 66% off for (up to) two days late; ...
- ▶ Midterm (20%)
- ▶ Final exam (40%)
- ▶ Caveat: Your grade may go up or down in extreme cases.
(down) Failure to hand in all the HW, (up) very strong exam scores
- ▶ You **MUST** make the exam dates (unless you have an exception based on UW policies). Do not enroll in the course otherwise.

“Can I Take The Class?”

- ▶ Short answer: if you are qualified and can register, yes
 - ▶ Math prerequisites: probability, statistics, algorithms, and linear algebra background.
 - ▶ Programming prereqs: strong programmer (e.g. comfortable in python)
- ▶ We will move fast; lectures will focus on concepts and mathematics
- ▶ work hard, do the readings, etc...

To-Do List

- ▶ Quiz section meetings start tomorrow. **Bring your laptop! Python review**
- ▶ Readings (do them, before the class)
- ▶ Academic integrity statement: on the course web page. ultimately, it is up to you to carry yourself with integrity.
- ▶ Gender and diversity statement (an acknowledgement): please try to act appropriately, knowing that.

Integrity

- ▶ Academic integrity policy: on the course web page. ultimately, it is up to you to carry yourself with integrity.
- ▶ Gender and diversity statement: (an acknowledgement) the current state is not balanced in any reasonable way; please try to act appropriately. people can surprise you...

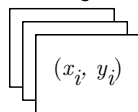
The Standard Learning Framework

"Inductive" Supervised Machine Learning

- ▶ **Training:** a learning algorithm takes a set of example input-output pairs, $\{(x_1, y_1), \dots, (x_N, y_N)\}$, and returns a function f (the 'hypothesis'); the goal is for $f(x)$ to recover the true label y , for each example, and on future examples
- ▶ **Testing:** we check how well f predicts on a set of test examples, $\{(x'_1, y'_1), \dots, (x'_M, y'_M)\}$, by measuring how well $f(x')$ matches y .

$$f(x) = y$$

training data



learning algorithm



f

x



$f(x)$

y

Inputs and Output

- ▶ x can be pretty much anything we can represent
 - ▶ To start, we'll think of x as a vector (or bundle) of attribute-value pairs, e.g.,
 $\phi(x) = \vec{v}$.
- ▶ y can be
 - ▶ a real value (regression)
 - ▶ a label (classification)
 - ▶ an ordering (ranking)
 - ▶ a vector (multivariate regression)
 - ▶ a sequence/tree/graph (structured prediction)
 - ▶ ...

$$\phi(x) \in \{0, \dots, 255\} \begin{matrix} \uparrow & \uparrow & \uparrow \\ 120 \times 240 \times 3 \\ \text{height} & \text{width} & \text{rgb.} \end{matrix}$$

-



-



- How are you? → Como estas?



More Examples:

- ▶ Regression:
Predict the depth of an object (e.g. a pedestrian) in an image.
- ▶ Ranking:
What order of ads should be displayed?