10-601 Introduction to Machine Learning

Machine Learning Department
School of Computer Science
Carnegie Mellon University

Course Overview

Readings:

Matt Gormley
Lecture 1
January 18, 2016

Outline

- Prerequisites
  - What they are
  - Where they are used
  - Who is taking 601
- Course Policies
  - Syllabus on website
  - Background Test / Background Exercises
- How to succeed in 601
  - How to NOT succeed in 601
  - Reminders
- Schedule Overview

PREREQUISITES

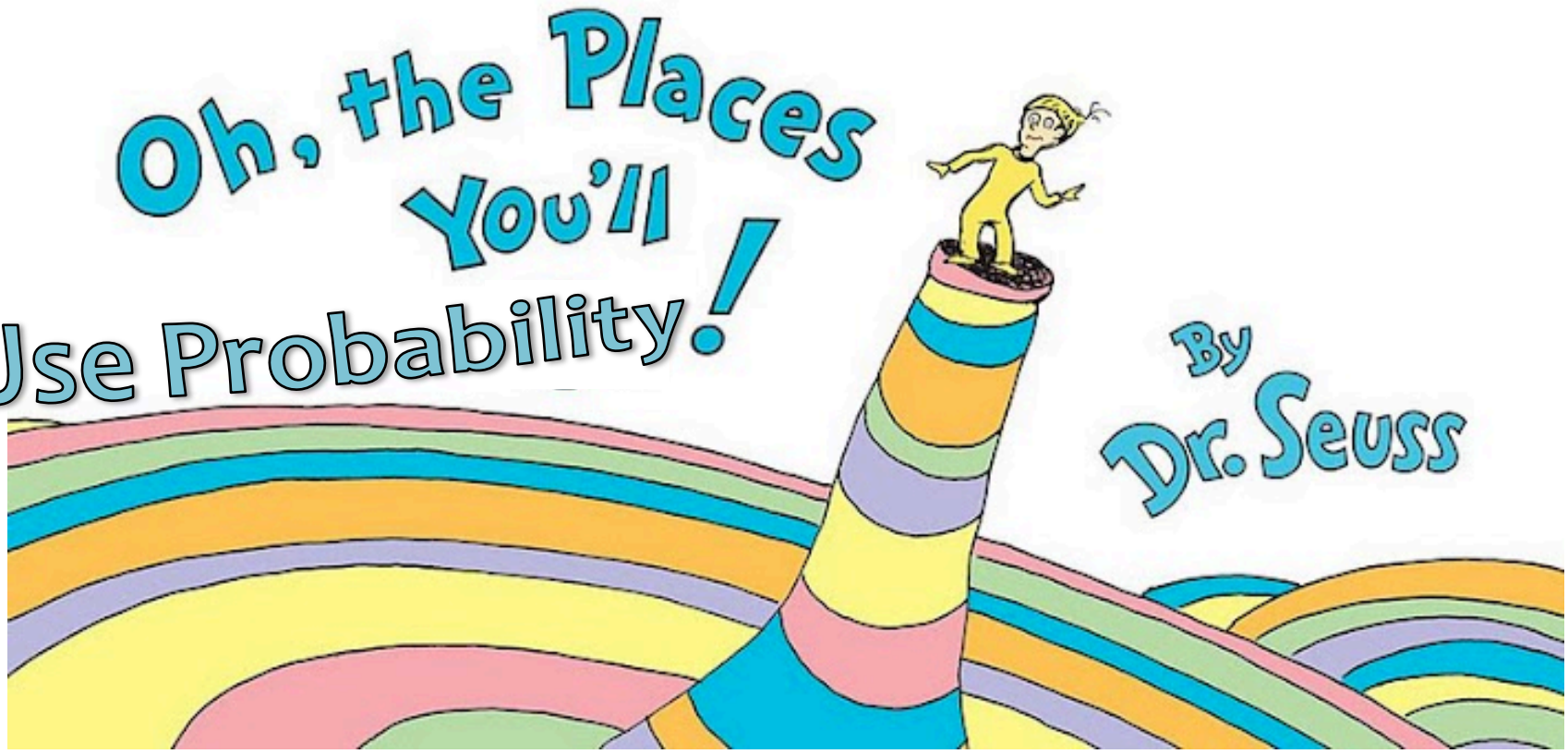
Prerequisites

What they are:

- Significant programming experience (15-122)
  - Written programs of 100s of lines of code
  - Comfortable learning a new language (if not a Python or Octave expert)
- Probability and statistics (36-217, 36-225, etc.)
- Mathematical maturity: discrete mathematics (21-127, 15-151), linear algebra, and calculus

Prerequisites

What if I don't meet them and plan to take the course anyway?


Oh, the Places You'll Use Probability!
By Dr. Seuss

Oh, the Places You'll Use Probability!

Supervised Classification

- Naïve Bayes

$$p(y|x_1, x_2, \ldots, x_n) = \frac{1}{Z} p(y) \prod_{i=1}^{n} p(x_i|y)$$

- Logistic regression

$$P(Y = y|X = x; \boldsymbol{\theta}) = p(y|x; \boldsymbol{\theta})$$
$$= \frac{\exp(\boldsymbol{\theta}_y \cdot \mathbf{f}(x))}{\sum_{y'} \exp(\boldsymbol{\theta}_{y'} \cdot \mathbf{f}(x)}$$

Note: This is just motivation – we'll cover these topics later!

Oh, the Places You'll Use Probability!

ML Theory

(Example: Sample Complexity)

- Goal: h has small error over D.

  True error: $err_D(h) = \Pr_{x \sim D}(h(x) \neq c^*(x))$

  How often $h(x) \neq c^*(x)$ over future instances drawn at random from D

- But, can only measure:

  Training error: $err_S(h) = \frac{1}{m}\sum_i I\big(h(x_i) \neq c^*(x_i)\big)$

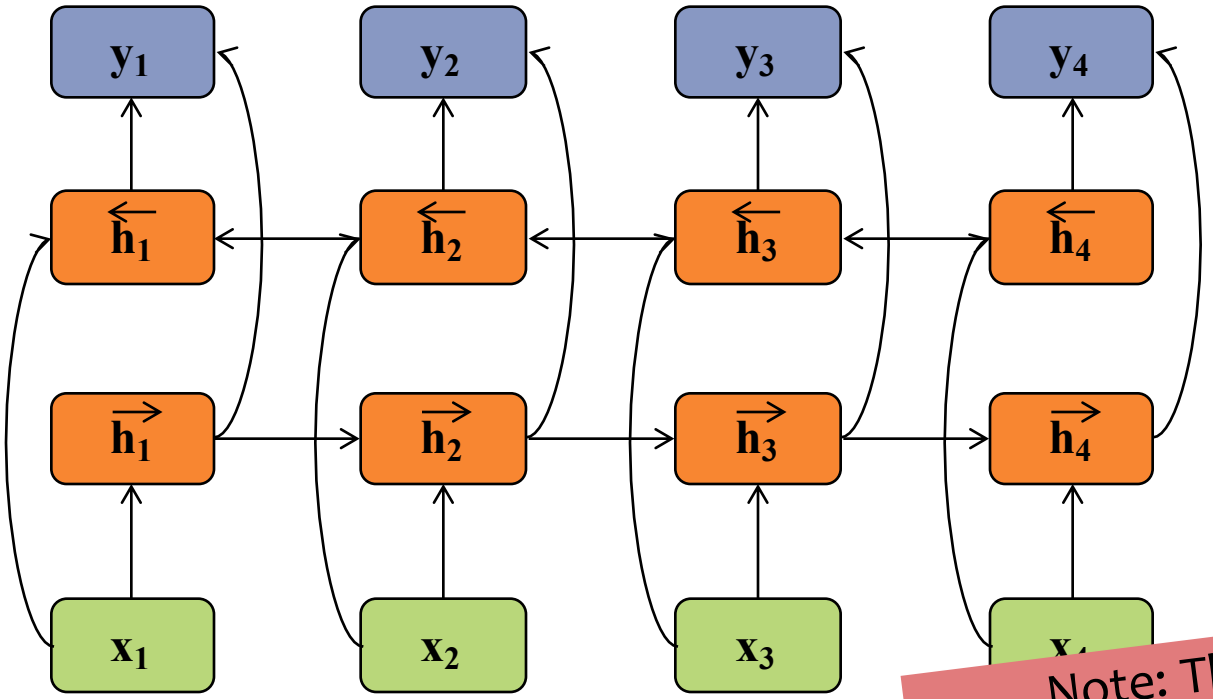
  How often $h(x) \neq c^*(x)$ over training instances

Sample complexity: bound $err_D(h)$ in terms of $err_S(h)$

Note: This is just motivation – we'll cover these topics later!

Oh, the Places You'll Use Probability!

Deep Learning
(Example: Deep Bi-directional RNN)

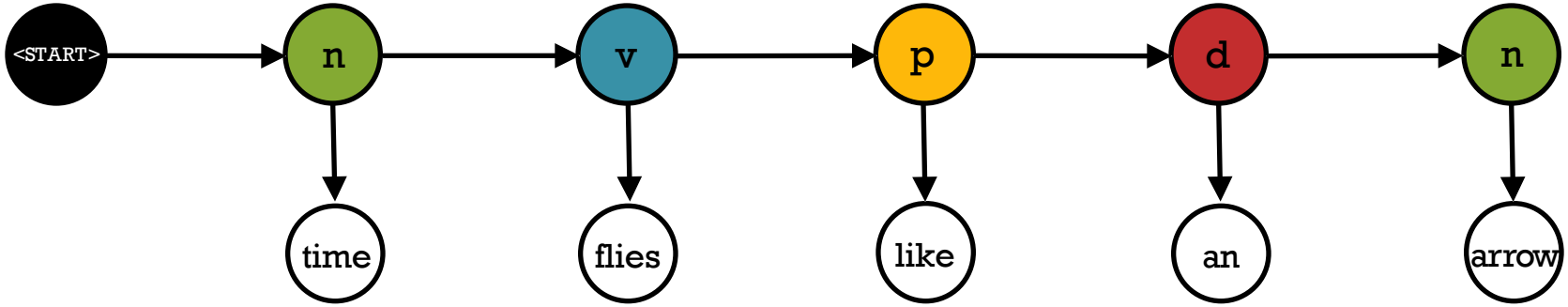


Note: This is just motivation – we'll cover these topics later!

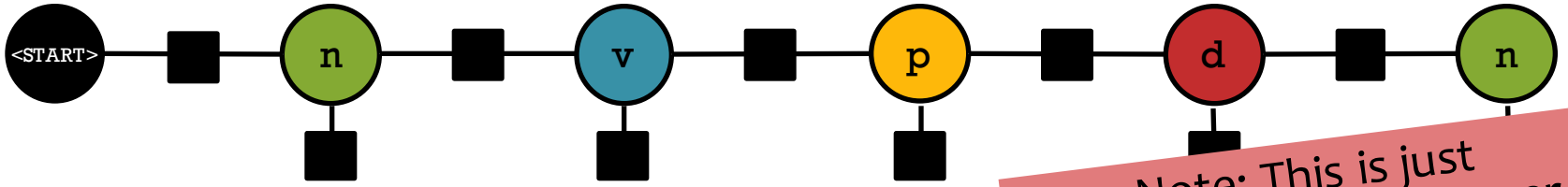
Oh, the Places You'll Use Probability!

Graphical Models

- Hidden Markov Model (HMM)



- Conditional Random Field (CRF)



Note: This is just motivation – we'll cover these topics later!

Prerequisites

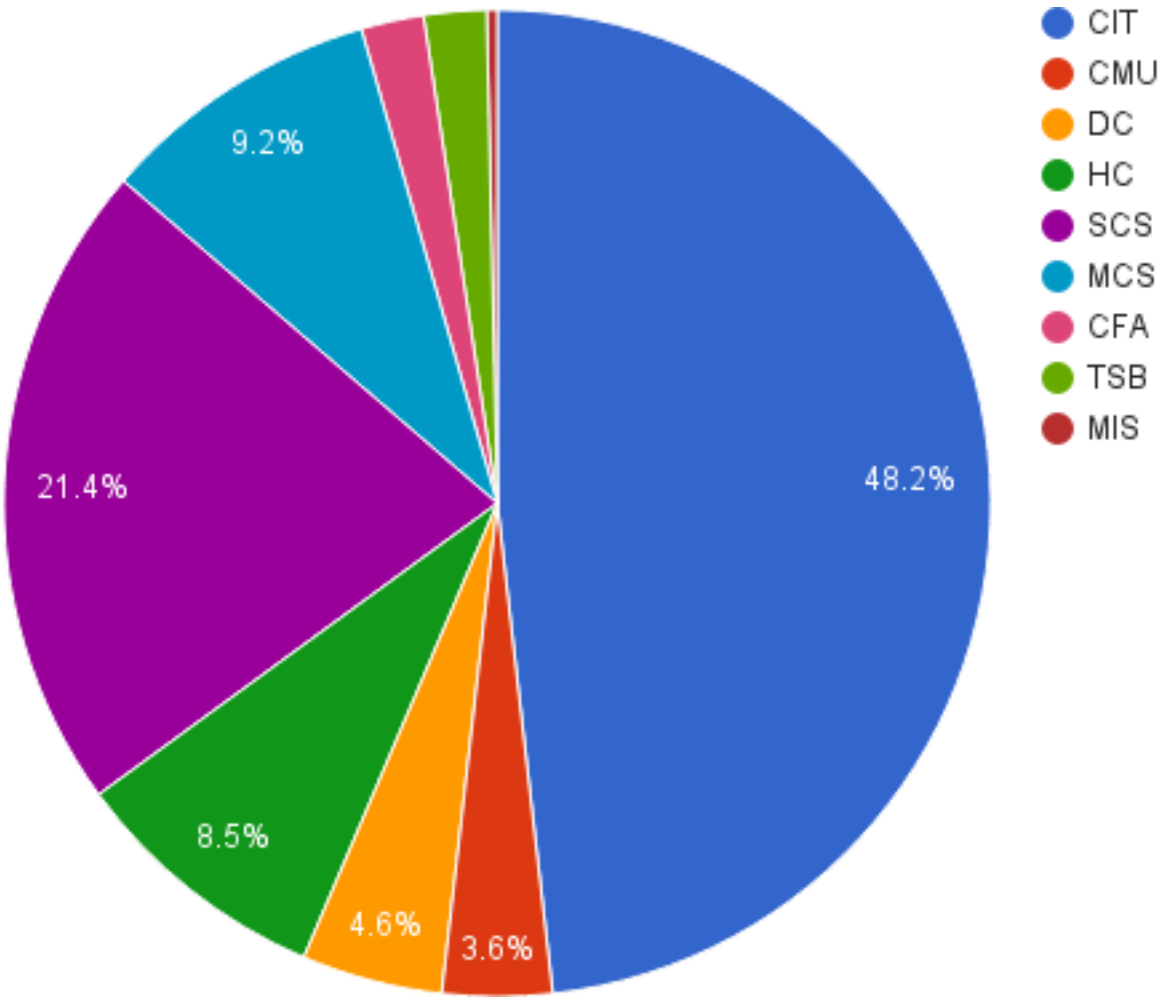
What if I'm not sure whether I meet them?

- Don't worry: we're not sure either
- However, we've designed a way to assess your background knowledge so that you know what to study!

  (see discussion of Background Test on website)

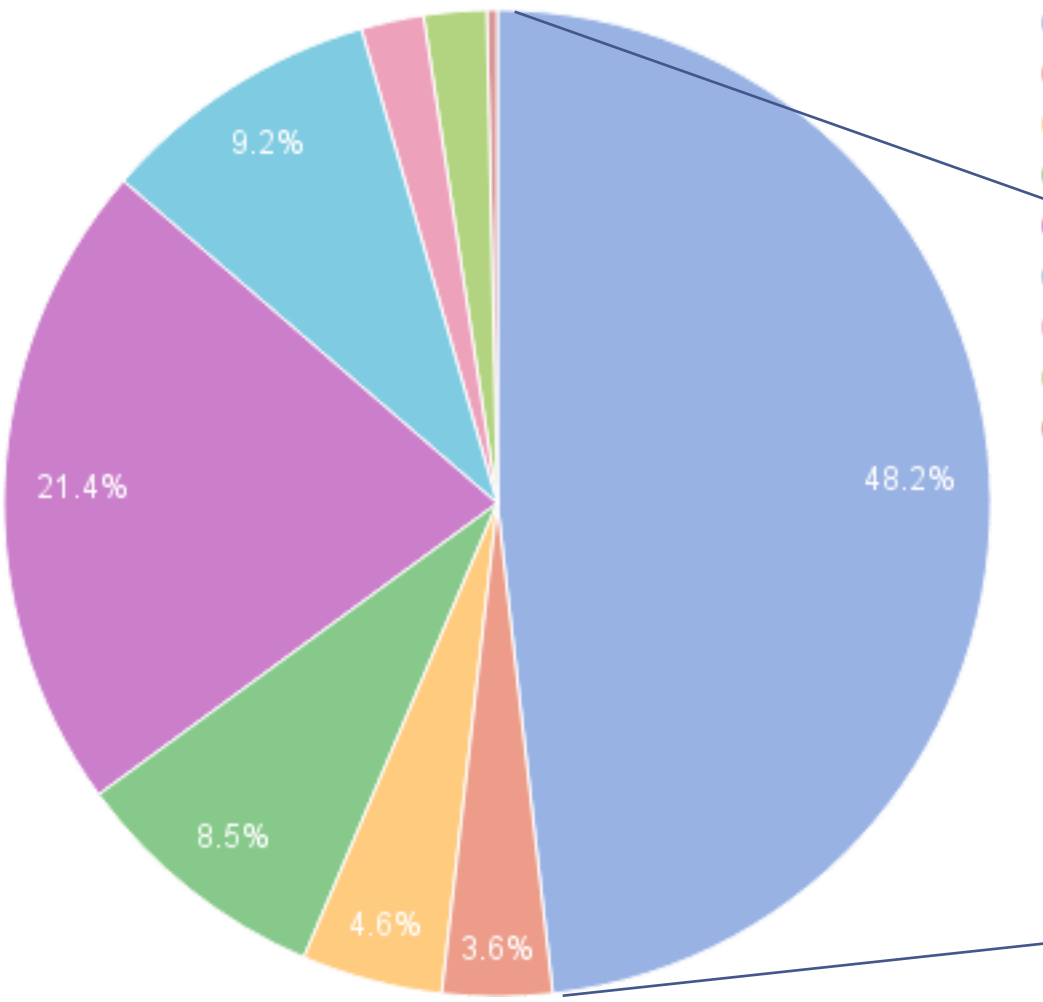
Who is taking 601?

Percent by College

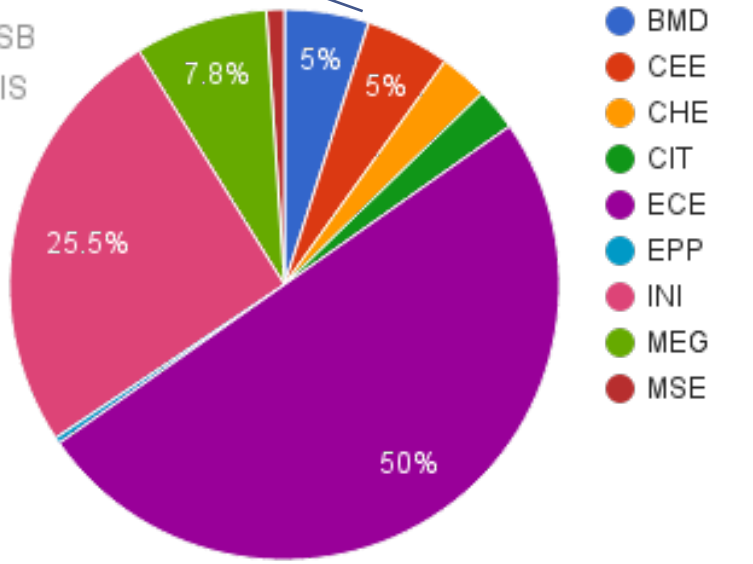


Who is taking 601?

Percent by College

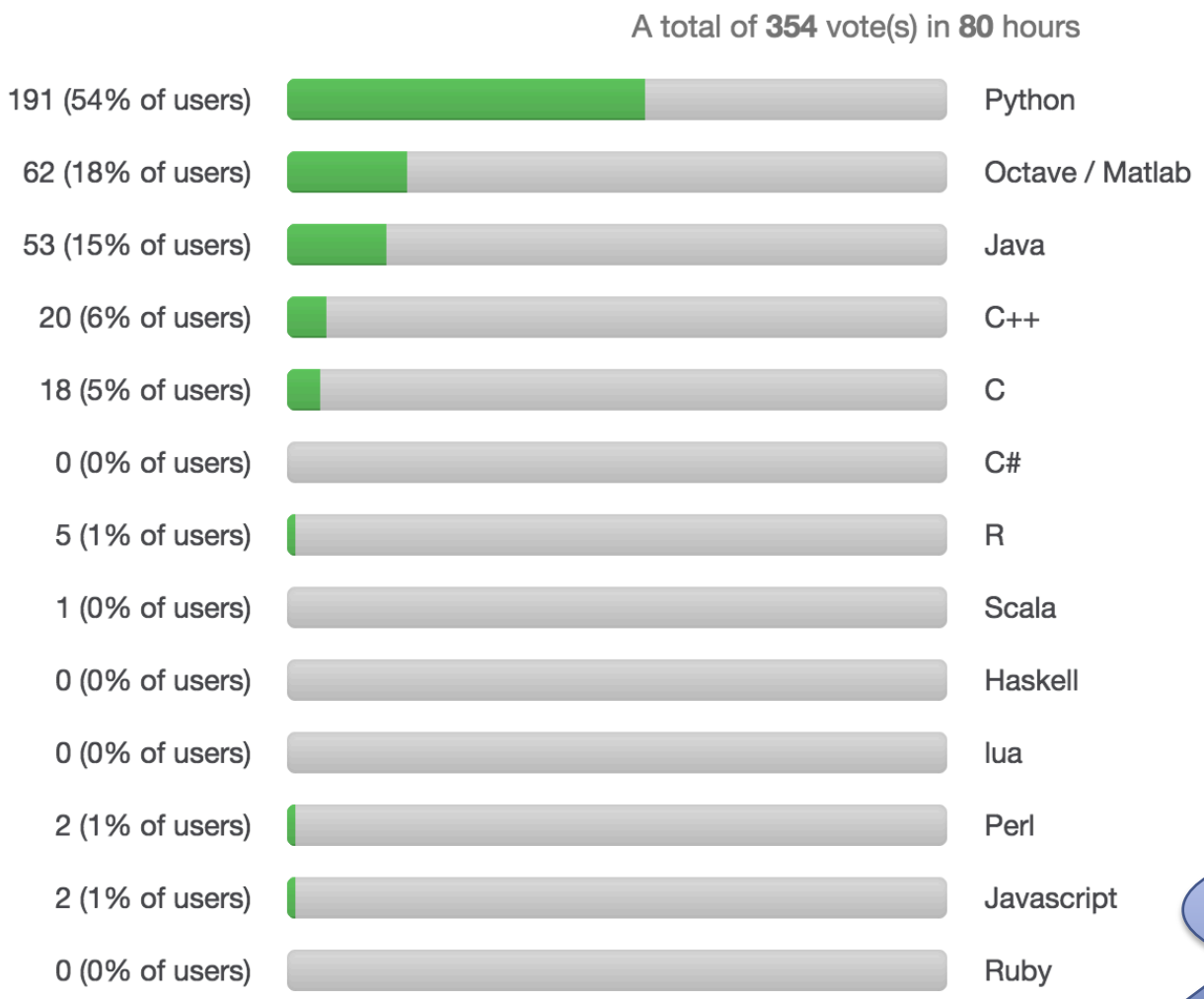


Percent by CIT Dept



Who is taking 601?

Programming Language Preferences



GoLang!

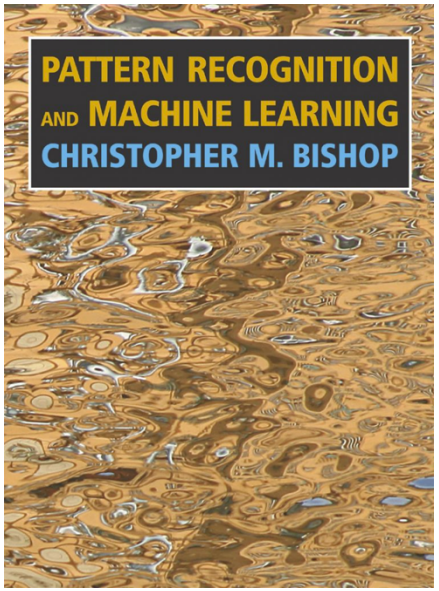
COURSE POLICIES

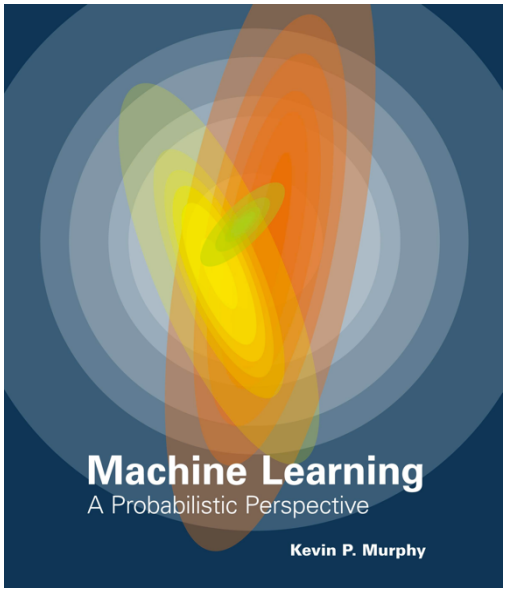
Lectures

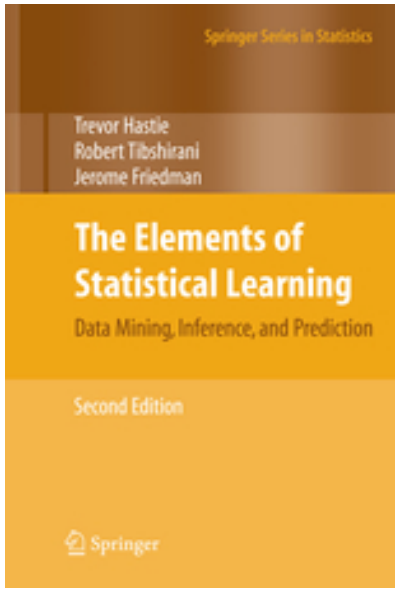
- You should ask lots of questions
  - Interrupting (by raising a hand) to ask your question is strongly encouraged
  - Asking questions later (or in real time) on Piazza is also great
- When I ask a question…
  - I want you to answer
  - Even if you don't answer, think it through as though I'm about to call on you
- Interaction improves learning (both in-class and at my office hours)

Textbooks

You are not *required* to read a textbook, but it will help immensely!







Syllabus

The syllabus is located on the course webpage:

http://www.cs.cmu.edu/~mgormley/courses/10601-s17

The **course policies** are **required** reading.

HOW TO SUCCEED IN 10-601

How to Succeed in 601

- Start the homework assignments *early*
- Ask the course staff for help (e.g. office hours)
- Don't skip lectures
- Read a Machine Learning textbook (seriously, pick one and read it!)

How to NOT Succeed in 601

- Ignore the prerequisites
- Stop submitting homework assignments
- Get overwhelmed by the fast pace of the course, but tell no one
- Copy code from a friend
- Cheat on an exam

Teaching Assistants

- 3-4 TAs will author each homework assignment
- Those are the best TAs to query about that assignment in office hours
- It is not the TAs job to teach you a programming language or to debug your program

Reminders

- Background Test
  - Tue, Jan. 24 at 6:30pm
- Background Exercises (Homework 1)
  - Released: Tue, Jan. 24 after the test
  - Due: Mon, Jan. 30 at 5:30pm

SCHEDULE OVERVIEW

Frameworks for Learning

- Function approximation
- Probabilistic approach
- Information theoretic approach

Learning Settings

- Supervised learning
- Unsupervised learning
- Semi-supervised learning
- Feature learning
- Structured prediction*
- Active learning
- Reinforcement learning
- Collaborative filtering

Theoretical Questions

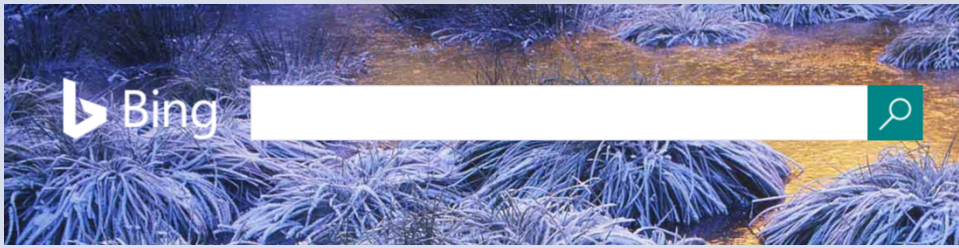
- How many examples do we need to learn?
- How do we quantify our ability to generalize to unseen data?
- Which algorithms are better suited to specific learning settings?

Machine Learning & Ethics

What ethical responsibilities do we have as machine learning experts?

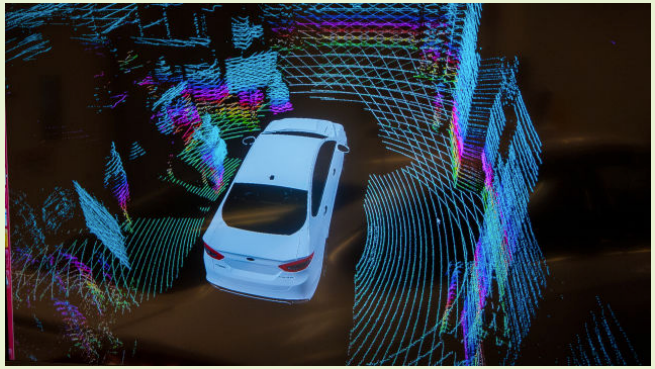
Some topics that we won't cover are probably deserve an entire course

If our search results for news are optimized for ad revenue, might they reflect gender / racial / socio-economic biases?



http://bing.com/

http://arstechnica.com/



How do autonomous vehicles make decisions when all of the outcomes are likely to be negative?

Should restrictions be placed on intelligent agents that are capable of interacting with the world?

http://vizdoom.cs.put.edu.pl/