

1. Introduction

- What is machine learning?
- Syllabus and logistics

Course details

CSE 512 Machine Learning (Fall 2020)

- Class time: 2:40pm-4pm Eastern standard time
- Location: Online (Zoom)

Teaching staff

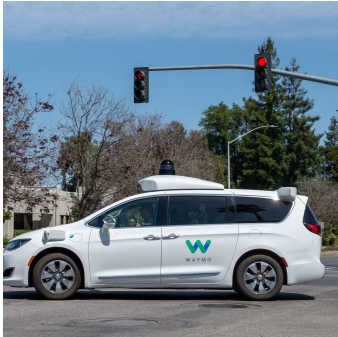
- Instructor: Yifan Sun (email yifan.sun@stonybrook.edu)
- TAs: TBA, Office hours: TBA

Important resources

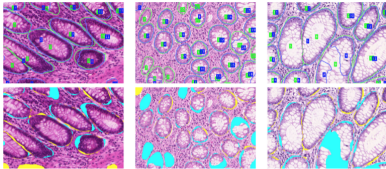
- Blackboard <https://blackboard.stonybrook.edu/webapps>
- Piazza <https://piazza.com/> working on website... (public facing)

Let me know ASAP if you don't have access to either one

What is machine learning?



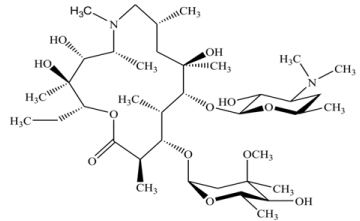
What is machine learning?



(d) benign

(e) benign

(f) malignant



What is machine learning?



People

• Jolie in a PEOPLE cover story

Want more stories like this?

Sign up for our newsletter and other special offers:

[SIGN ME UP](#)

MORE ON: [Angelina Jolie](#), [Brad Pitt](#)

Share this story:


[f](#) Share [t](#) Tweet


Your reaction: 47 votes


[WOW](#) [LOL](#) [LOVE](#) [Sad](#) [Angry](#)


1 2 37 1 4

AROUND THE WEB

 The 12 Most Stunning Celebrity Brides
WeddingWire

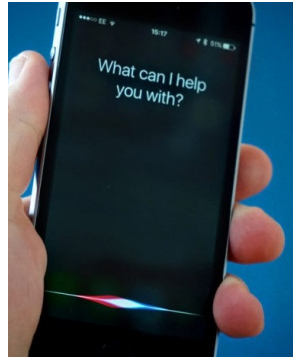
 Where Did They Go? 10 Award-Winning Actors Who Don't Act Anymore
Wall St. Cheat Sheet

 A Complete Look At The Wedding The Osbuns Helped Make Possible
Bustle

 What You Need to Know Before Shopping for a Wedding Dress
WeddingWire

[more info](#)

What is machine learning?



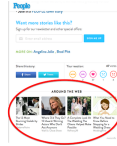
What is machine learning?



Facebook



Boston dynamic



everybody



Waymo

Some things are too

burdensome,
(document extraction)

dangerous,
(bomb threats)

boring,
(image search)

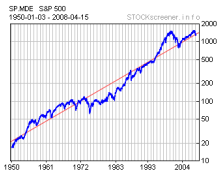
subjective
(stock trends)

for a human

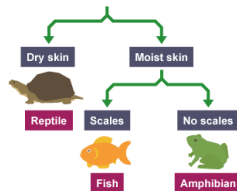
How to solve these problems?



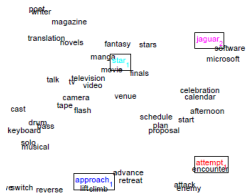
classify



regress (fit)



logical rules



model / embed

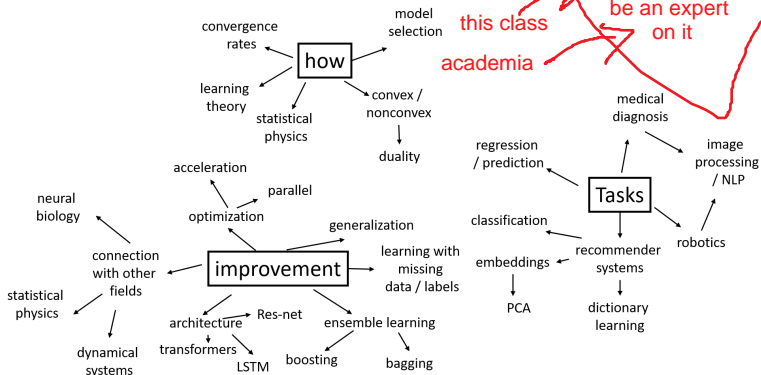


copy friends



imitate

Deluge of tools



... linear / ridge regression, feature selection / screening, CNNs/RNNs/GANs, Hidden Markov models, decision trees, random forests, boosting, XGBoost, bagging, bayesian inference, word embeddings, bidirectional RNNs, nearest neighbors, regularization / overfitting, support vector machines, ...

We cover depth. How to get breadth?

- Online resources: papers, arxiv, blogs, quora, twitter, other ML courses
- Many slides here are heavily based on last year's course offering
- You don't need to know everything to do machine learning, but you should know where to look
- This class: depth on a few fundamental topics.
- Classes in CS dept. for followup: CSE 519, CSE 544 (data science), CSE 527 Computer Vision, CSE 537 Artificial Intelligence, CSE 532 Speech Processing, CSE 545 Big data analytics

Tentative schedule

- Simple tools and Bayes
 - logistic / linear regression, point estimation, Bayes classifier, naive Bayes
- Margin based methods
 - convex optimization, max margin classifier, dual SVM, kernels
- Tree based methods
 - decision trees, bagging, boosting
- Generative models
 - Kmeans, GMM, Graphical models, HMM
- Dimensionality reduction
 - recommender systems, PCA, ICA, LDA
- Advanced topics
 - deep learning, online optimization, learning theory (more by popular demand)

Structure

Prerequisite:

- Comfortable in linear algebra, probability, statistics, logic
- Coding: Matlab or Python (or learn on your own).

Lectures

- Synchronous lectures on zoom, will be recorded and uploaded
- We will rely heavily on Piazza for Q/A

Assignments: 2 types of questions

- normal problems (graded)
- challenge problems (graded, bonus)

Please be flexible and proactive with feedback, as online synchronous teaching is uncharted territory for all of us.

Grading

Distribution

- 6 normal assignments, each 10 pts
- 5 challenge problems, 2 pts each
- 10 pts guided project
- 10 pts midterm, 20 pts take home final

sums to $110 > 100$ points (You can pick some things to drop!)

- 90 to 100 = A
- 80 to 89 = B
- 65 to 79 = C
- < 65 = fail
- Borderline grades: depends on class grade distribution

Note that you can fail both exams and still not fail the course

Misc policies

- You get 5 late days total (not counting weekends). After that, any late assignments will not be graded. You must declare that you are using a late day when submitting your assignment.
- If you have difficulty submitting on time because of internet issues, please contact me/TAs ASAP.
- Midterm: ~~Monday Oct 12 (given online)~~ ^{poll on piazza}. There will be **no** retakes. (Report scheduling conflicts within the next 2 weeks or forever hold your peace.)
- You may ask TAs / me for 1 regrade per assignment, but we will regrade the entire assignment and your score could decrease.

Academic misconduct

- Feel free to discuss assignments / projects with anyone, but do them problems yourself. Do not look up past year solutions or solutions online.
- Exams should be done individually (honor code).
- If we notice code being shared amongst groups of people, we will take the total score and divide them by the number of people sharing code. (5 people with exact code earning 10 pts = 2 pts each)
- If we detect plagiarism (in code or homework solutions) we will report it and you will get a 0. If we detect cheating in an exam you will automatically fail the course (and be reported).
- When in doubt, ask

We live in strange and unusual times

- Please report all special circumstances early, as there is a higher chance we can be flexible.
- Understand that we are all doing this for the first time, so be patient with technical difficulties, TA schedules, etc.
- Your health is the most important. Please do not prioritize this course over staying safe and healthy. Be responsible and attentive to your own needs.