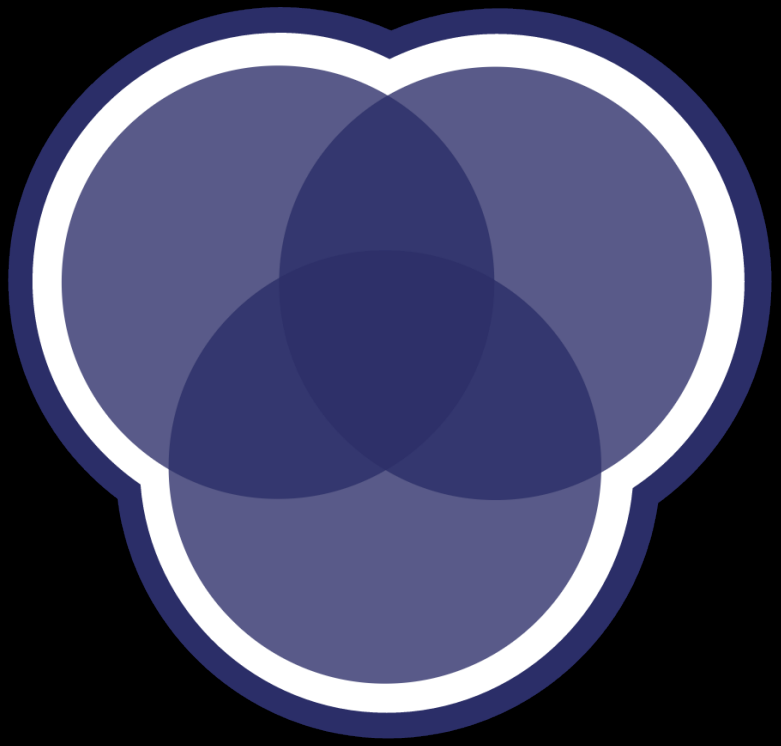
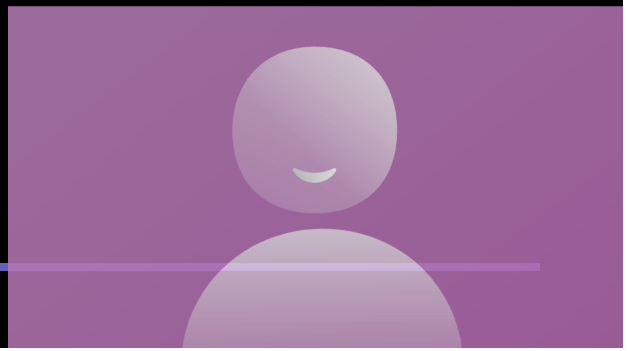
CDS Education

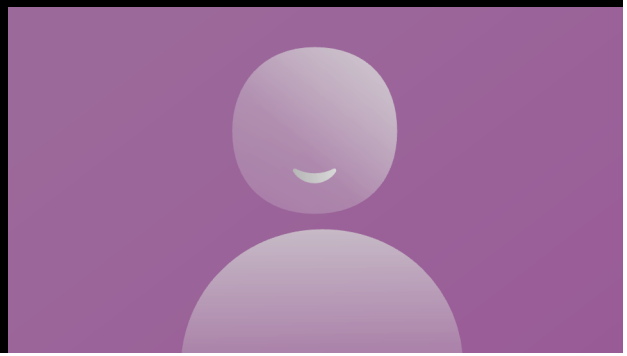
INFO 1998: Intro to ML

Lecture 1

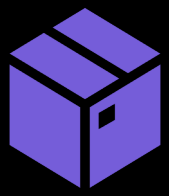




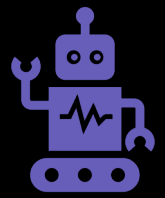
Why should machines learn?



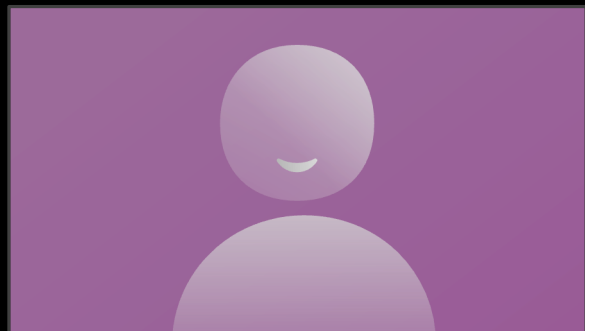
Manual methods fall short at scale.



White-box models are not complex enough.



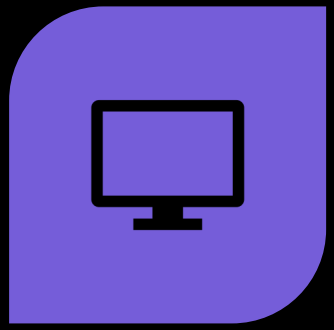
Automation.



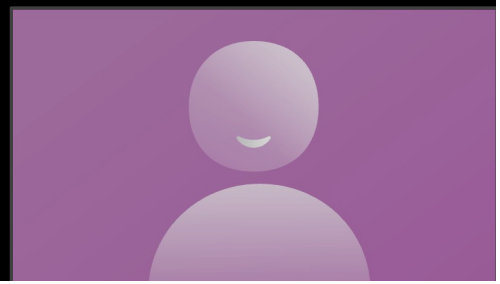
Frontiers

MEDICAL
DIAGNOSTICS

MOVIE
RECOMMENDATIONS



GENERATING MEDIA

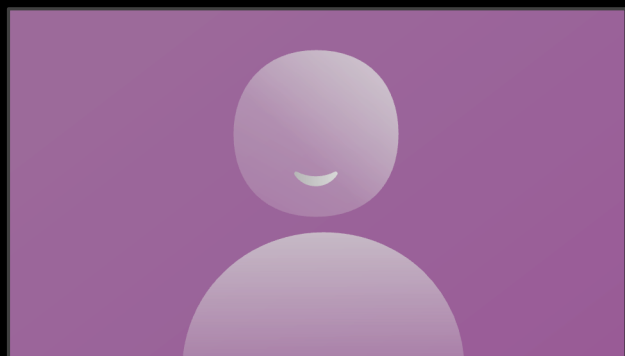


Bottom Line

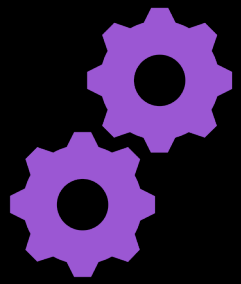
We **need** efficient + robust machines that can **make**

1. Decisions
2. Predictions
3. Content

…and we need people to make them.

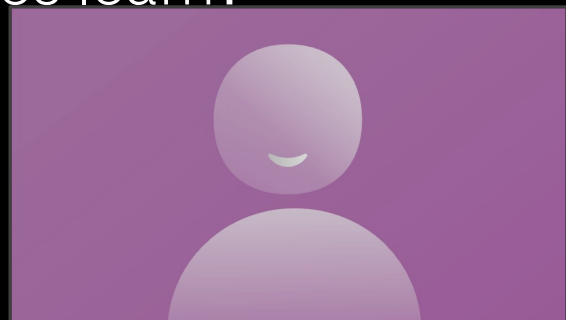


Natural Questions

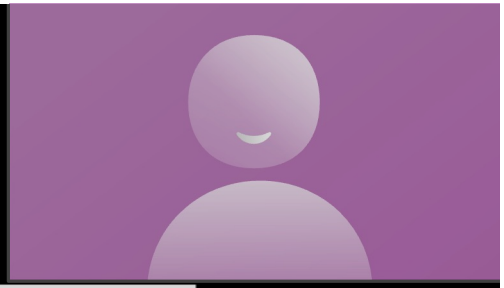


What do machines learn?

How do machines learn?



Course Specifics

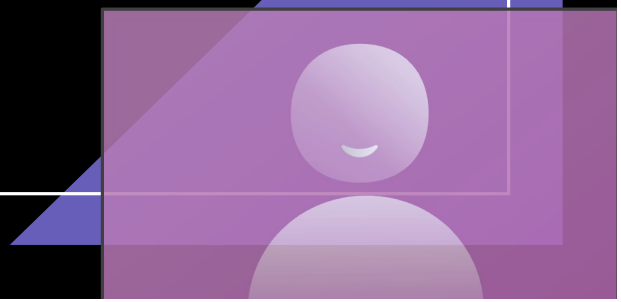


OBJECTIVES	SYLLABUS
Manipulating Data	Data Manipulation / Visualization *Lectures 1-3*
Communicating Data	
Understanding of ML as a concept	Fundamentals of Machine Learning *Lectures 4-5*
Intuitive understanding of ML models	
Implementation of ML models	Supervised Learning *Lectures 6-8*
Comfort Using Python	
Applications in Industry	Unsupervised Learning *Lecture 9*
Project Experience	

Syllabus is posted online.

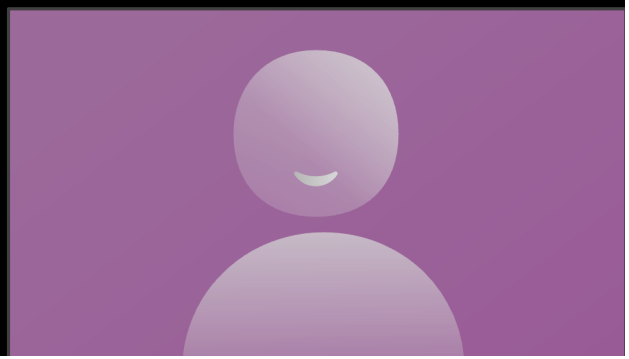
Why this course?

1. An introduction to the ML modeling pipeline.
2. An *intuition* for the theory behind ML models.
3. A framework for reasoning about model selection + performance.
4. Hands-on skills in data cleaning & training models.



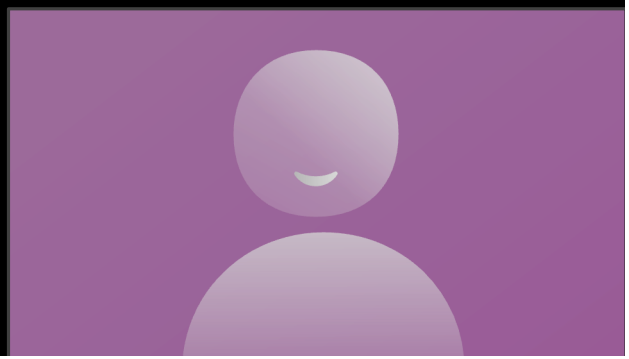
Logistics

- 1 credit S/U ONLY
- Advisor Approved Elective
- 10 lectures (required attendance)
- Several planned workshops
- 9 assignments (~1/week)
- Discussions: EdStem
- Office Hours (Calendar TBD)



Grading

9 assignments	(lowest dropped)	55%
Final Project	(2-3 people/group)	40%
Attendance	(mandatory)	5%



Final Project

Pre-Processing & Data Manipulation

obtain + clean + manipulate data

Visualizations

understand your dataset

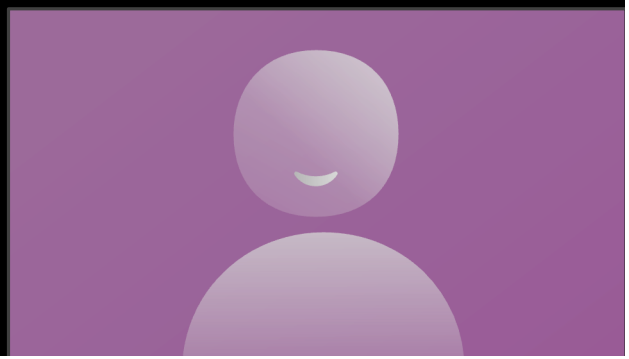
Models

effective + optimized solutions to your problem

FAQs

1. Will this course make me a Data Scientist?

No, it will not. The course covers a breadth of concepts, helps build understanding of some models, but does not dive too deep into the mathematical complexities. Intuition is our goal.



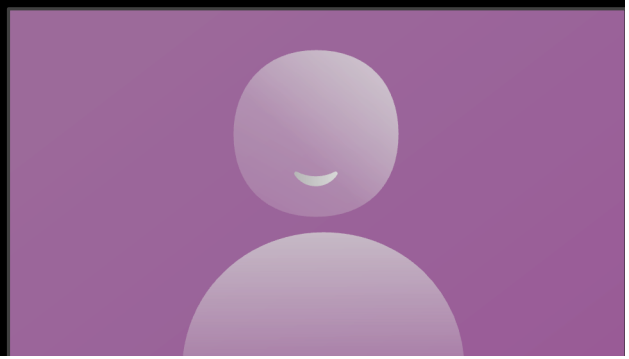
FAQs

2. Time Commitment?

Your call.

If all you want is basic street-fighting machine learning skills – < 1 hr/week.

If you want to engage passionately and build a creative data science project – that will take more time. But incredibly rewarding.



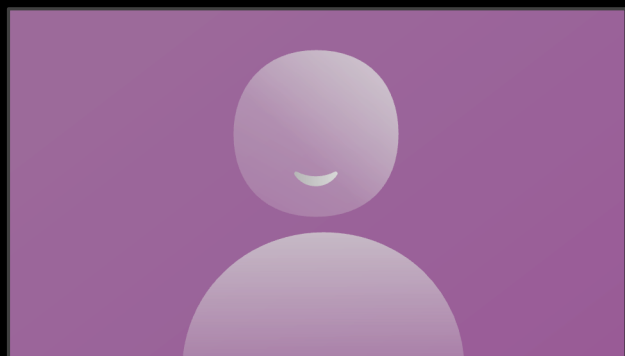
FAQs

3. I have no background in CS/Stats. Am I prepared?

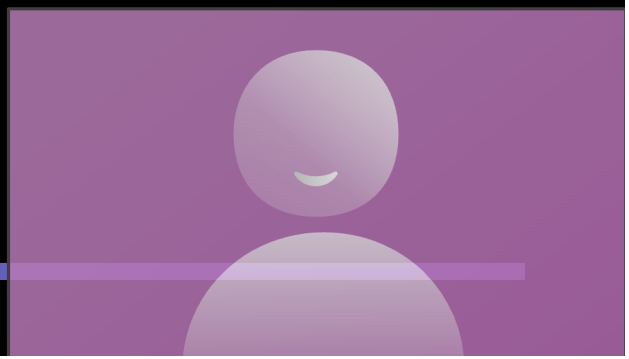
Not a problem.

We make no assumptions and encourage non-majors to take this course.

While you may need to put in some more work, we're here to help.



Who are we?



Course Director

Sri Kundurthy
cs + math
srk247@cornell.edu

Course Staff

Mericel Tao

Deniz BT

Skai Nzeuton

Jay Talwar

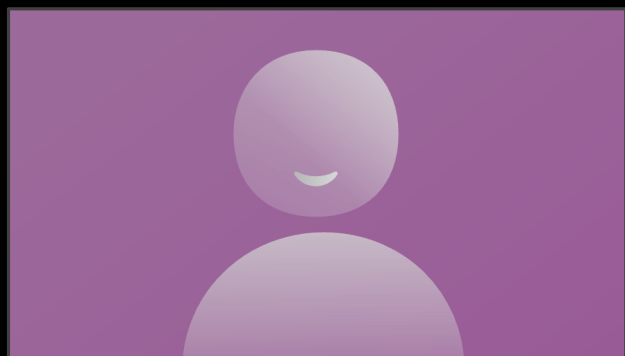
Leon Jiao

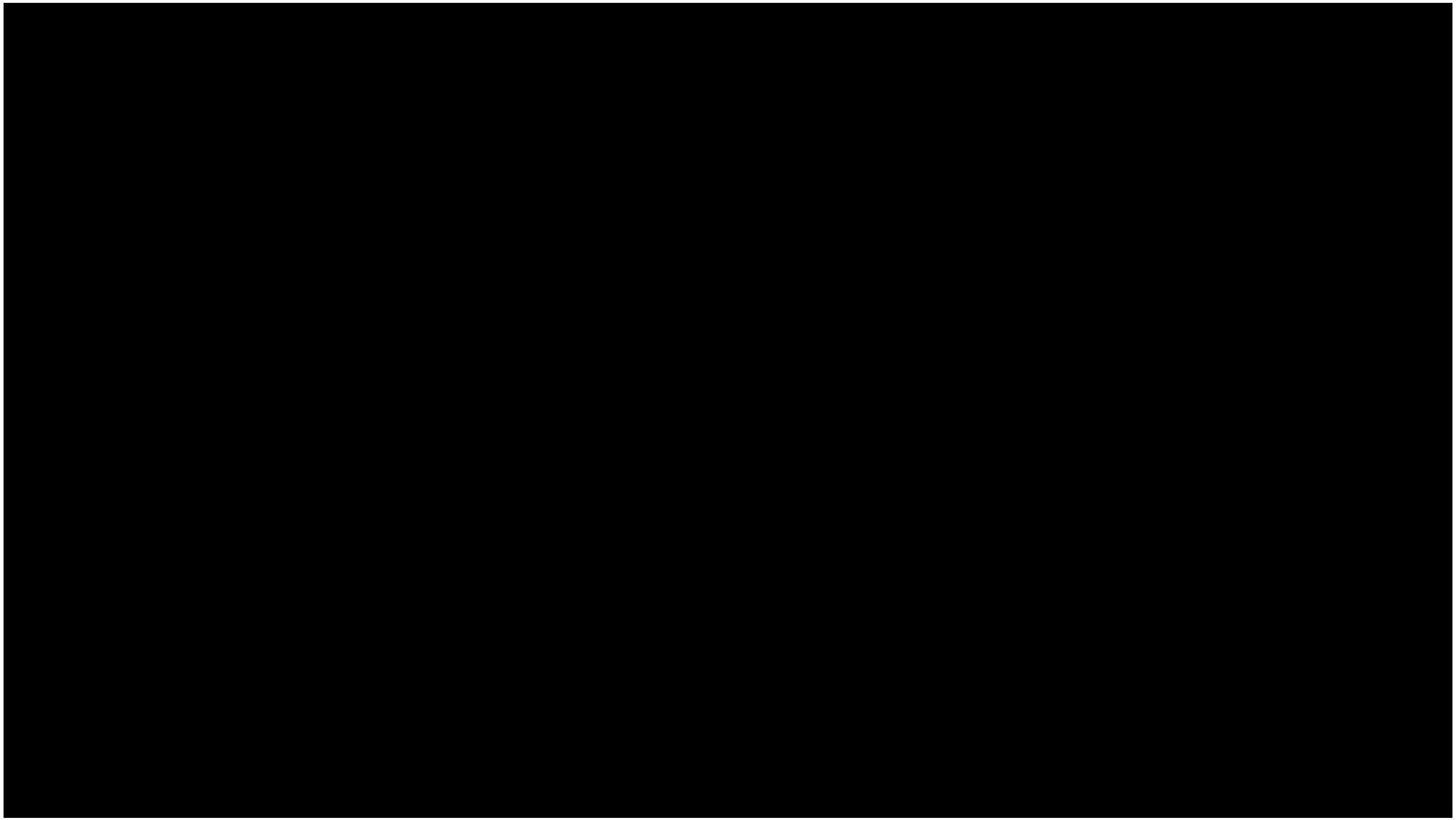
Minhaj Fahad

Tanvi Bhave

Eric Do

...



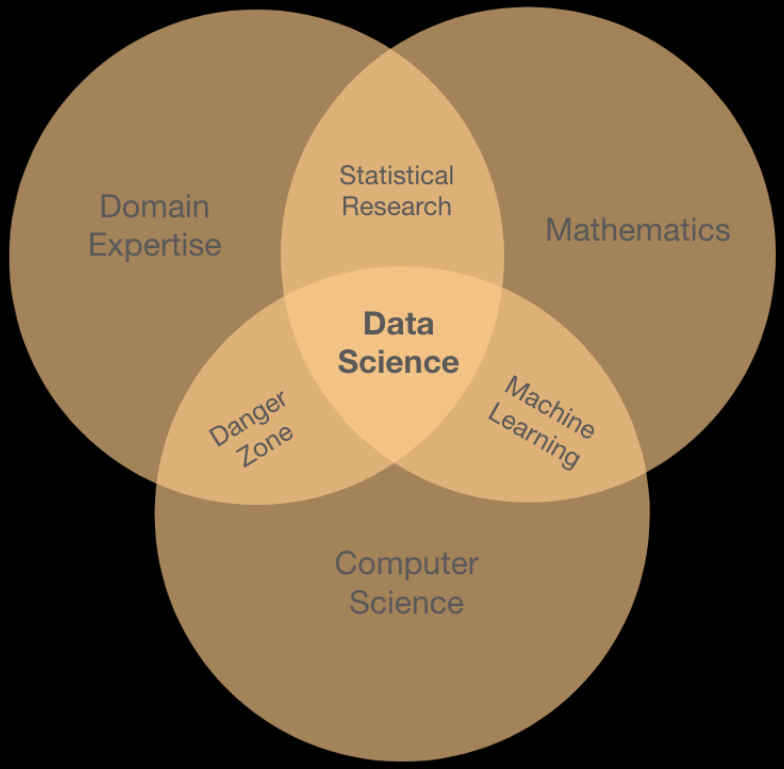


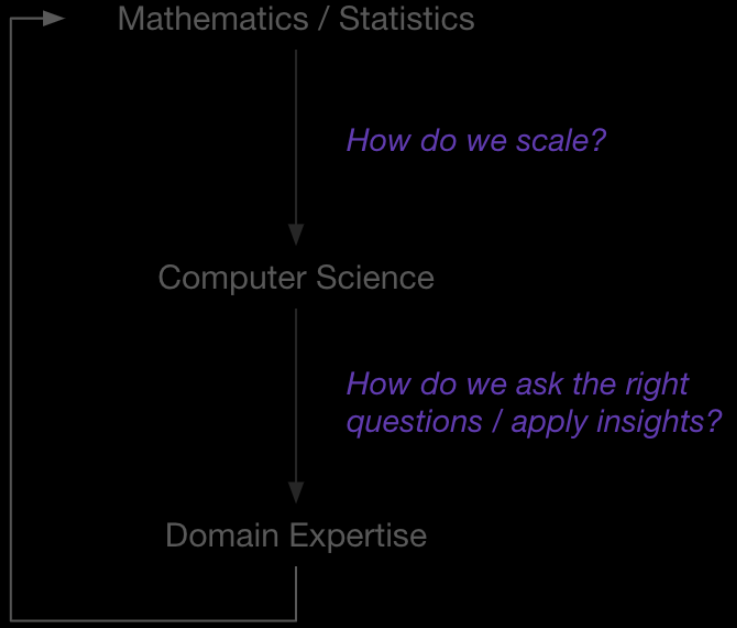
Data Science vs Machine Learning

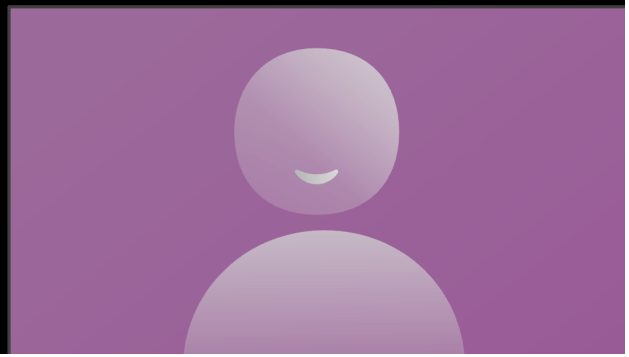
“By "Data Science", we mean almost everything that has something to do with data: Collecting, analyzing, modeling..... yet the most important part is its applications --- all sorts of applications.”

- Journal Of Data Science

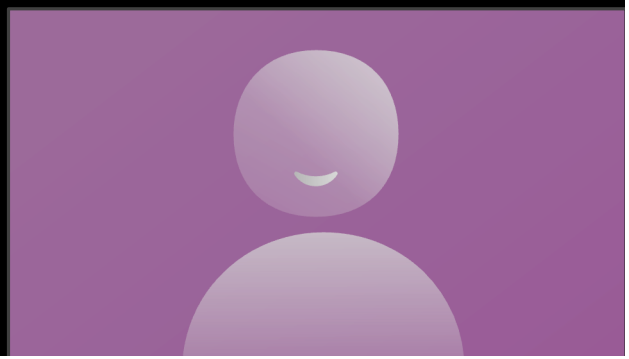
Data Science





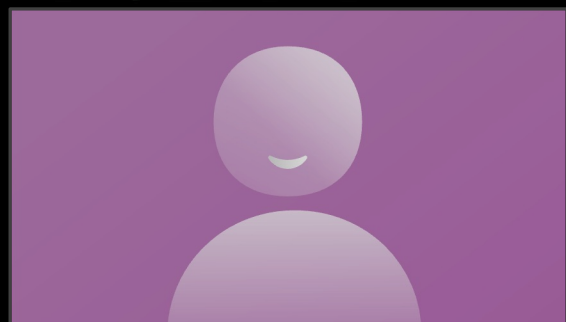


Machine Learning $\subseteq$ Data Science



Installation: Anaconda

- Go to https://www.anaconda.com/distribution/
- Open **Terminal** (MacOS) / **Command Prompt** (Windows),
- **Type and enter:** jupyter notebook



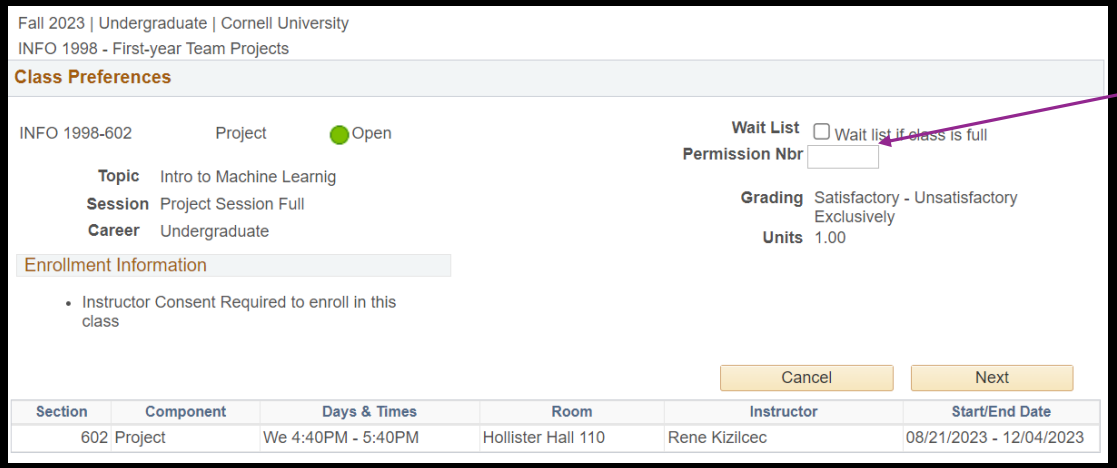
Enrollment

Fill out by Friday to get a pin.
Also counts for attendance.

https://forms.gle/sUV3wvPeMJ3k2Ba37

Enrollment in Student Center

- You'll receive pin via email soon.
- Add INFO 1998 Section 602 (Kizilcec)
- Please enroll as soon as you receive a pin.



Pin

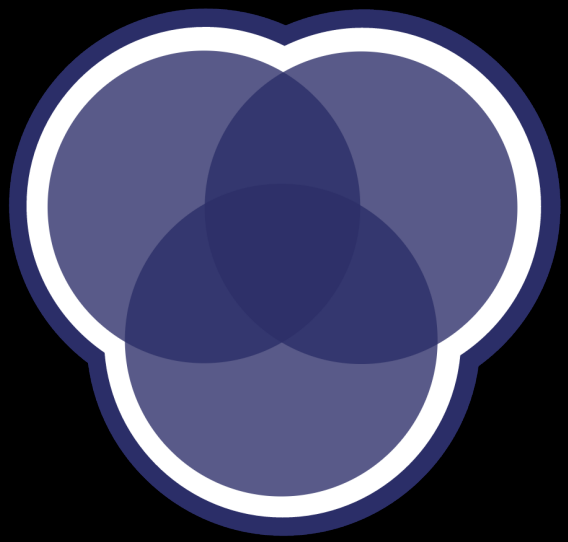
Course Links

- Ed Discussion
  - All queries + discourse (avoid email if possible)
- CMS
  - Assignment Files + Submission
- Course Website
  - cornelldatascience.github.io/info1998/

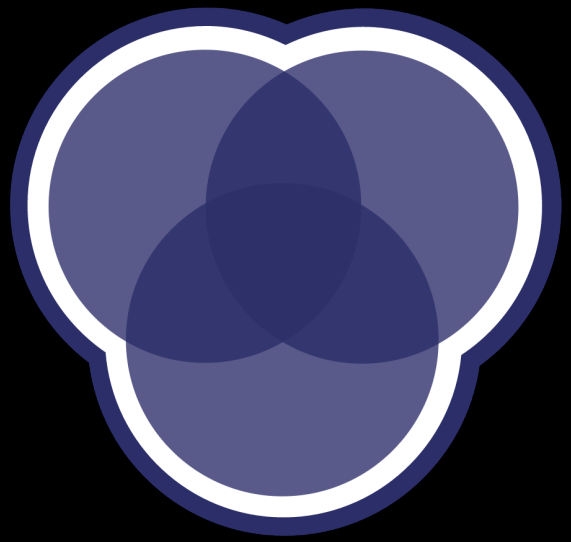
Jupyter Notebook Demo

Next Steps

- Assignment 1 Released Soon!
  - Due 11:59 PM Friday, 02/14/25.
- Enroll on Student Center
- Next Lecture: Data Manipulation



CDS Education



CDS Education