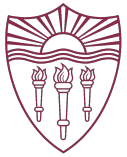


CSCI 552

Dr. Kristina Lerman

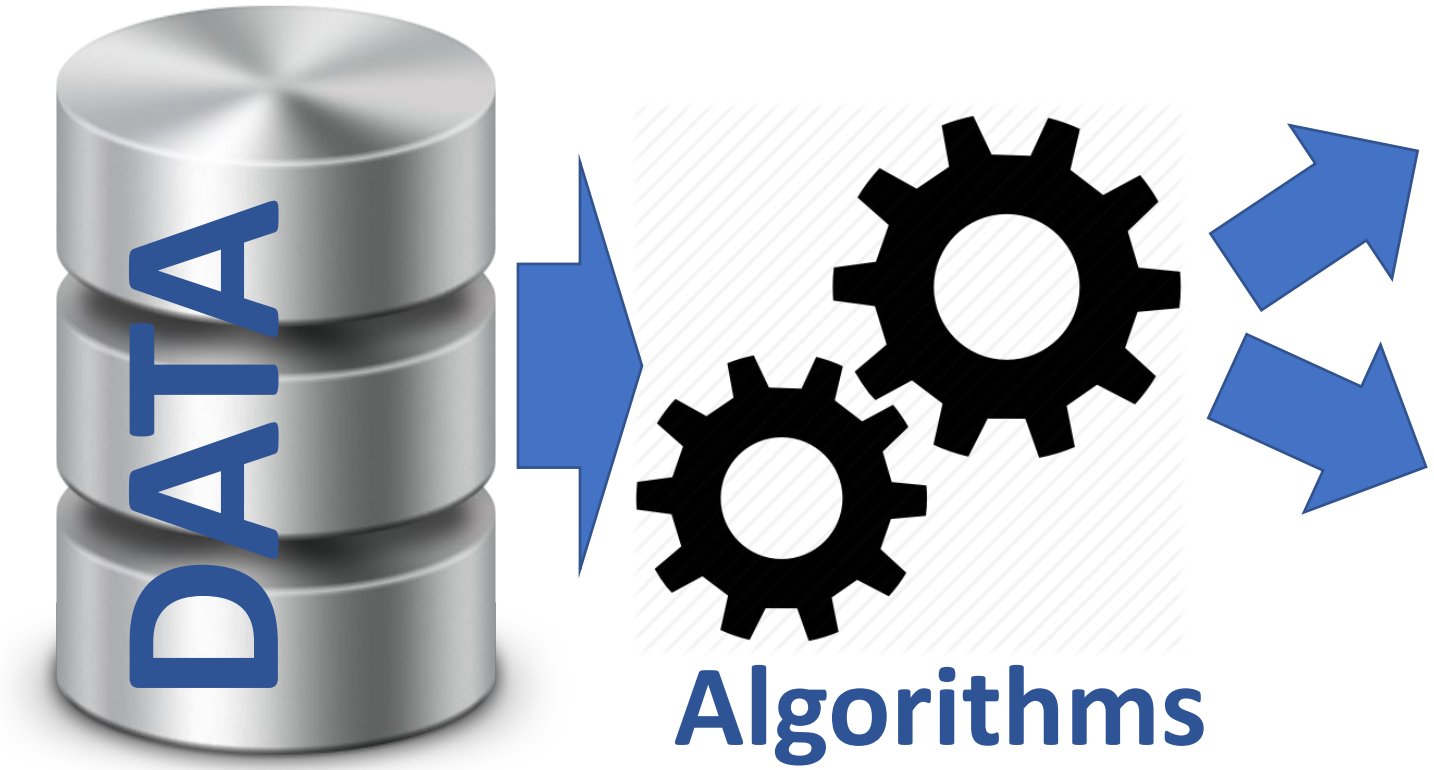
Dr. Keith Burghardt

USC Viterbi
School of Engineering

Outline

- Course description
  - Topics
  - Course goals
- Course details
  - Contact
  - Workload
  - Grading, etc.

What is Data Science*?



- Learns patterns from data
- *Predictive*: Uses learned patterns to make personalized predictions
- *Descriptive*: Uses learned patterns to better understand the data

* AI, Machine Learning, Data Mining

USC Viterbi School of Engineering

Big Data

- Widespread use of personal computers and wireless communication leads to “big data”
- We are both producers and consumers of data
- Data is not random, it has structure, e.g., customer behavior
- We need “big theory” to extract that structure from data for
  (a) Understanding the process
  (b) Making predictions for the future

What We Talk About When We Talk About "Learning"

- Learning general models from a data of particular examples
- Data is cheap and abundant (data warehouses, data marts); knowledge is expensive and scarce.
- Example in retail: Customer transactions to consumer behavior:

  People who bought "Blink" also bought "Outliers" (www.amazon.com)
- Build a model that is *a good and useful approximation* to the data.

AI is everywhere

Democracy

News ranking algorithms

- Does the algorithm create filter bubbles?
- Does the algorithm disproportionately censor content?

Algorithmic justice

- Does the algorithm discriminate against a racial group in granting parole?
- Does a predictive policing system increase the false conviction rate?

Kinetics

Autonomous vehicles

- How aggressively does the car overtake other vehicles?
- How does the car distribute risk between passengers and pedestrians?

Autonomous weapons

- Does the weapon respect necessity and proportionality in its use of force?
- Does the weapon distinguish between combatants and civilians?

Markets

Algorithmic trading

- Do algorithms manipulate markets?
- Does the behaviour of the algorithm increase systemic risk of market crash?

Algorithmic pricing

- Do algorithms of competitors collude to fix prices?
- Does the algorithm exhibit price discrimination?

Society

Online dating

- Does the matching algorithm use facial features?
- Does the matching algorithm amplify or reduce homophily?

Conversational robots

- Does the robot promote products to children?
- Does the algorithm affect collective behaviours?

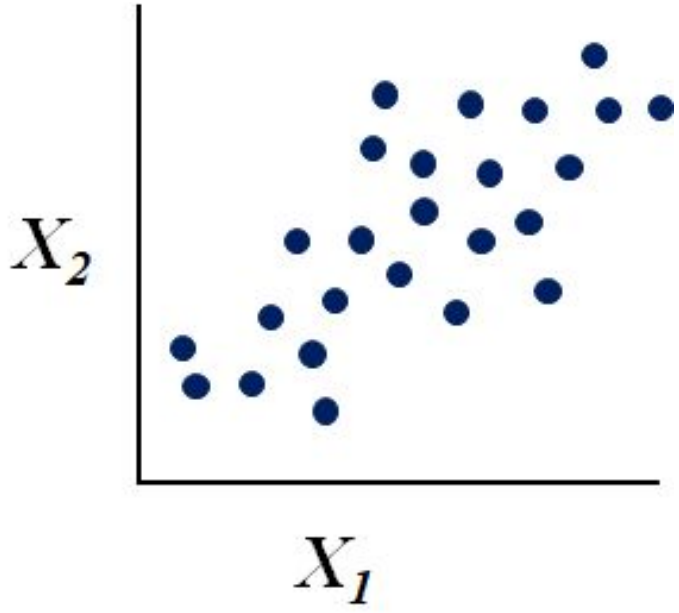
[Source: Rahwan et al. *Nature*, 2019]

USC Viterbi
School of Engineering

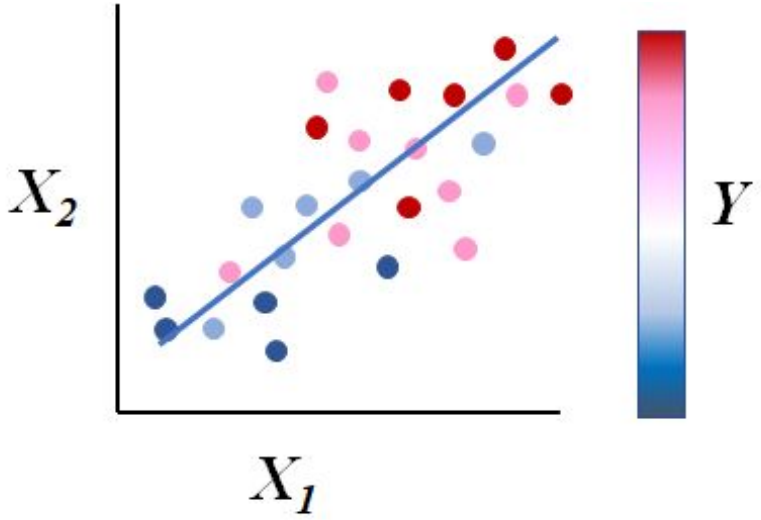
Types of machine learning

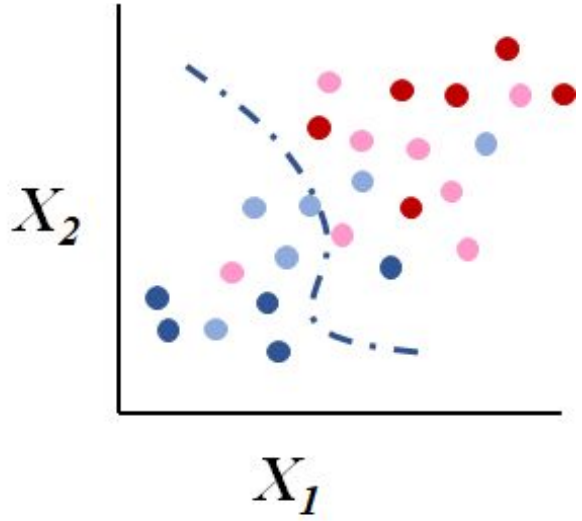
Unsupervised methods

- Input: a set of features
- Clustering
  - K-means, hierarchical, spectral, ...
- Dimensionality reduction
  - Principal component analysis (PCA),
  - Singular value decomposition (SVD), …
- Latent factors
  - Non-negative matrix factorization (NMF)
  - Topic modeling (LDA)
  - ...



Types of machine learning





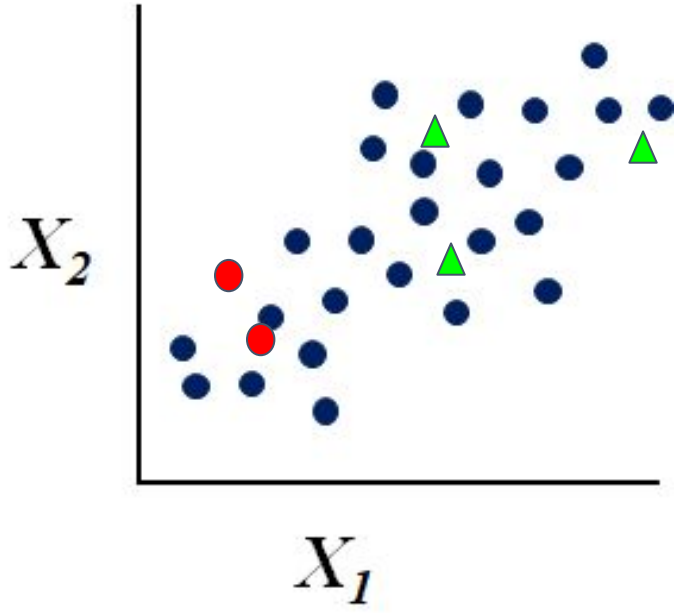
Supervised methods

- Input: features & labels/outcomes
- Regression
  - Linear regression, logistic regression, mixed effects…
- Classification
  - k-Nearest Neighbors (kNN), Random forest, Support Vector Machines (SVM), …
- Linear vs Non-linear

Types of machine learning

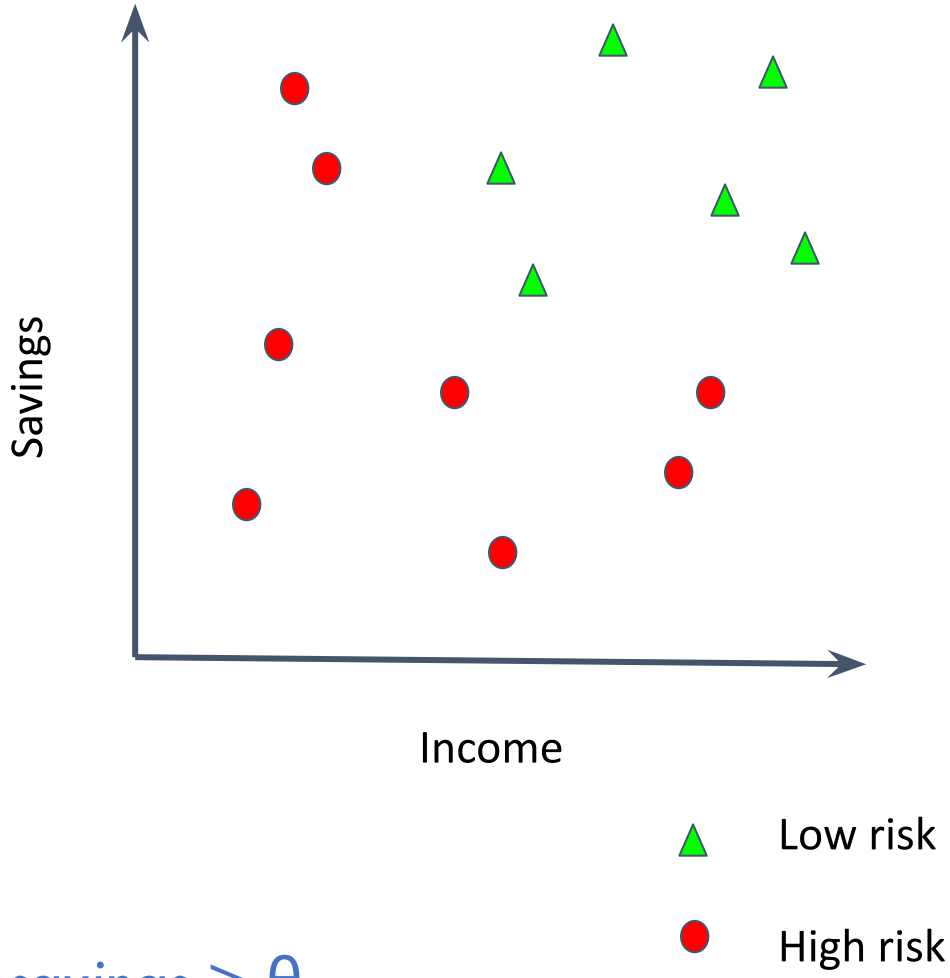
Semi-supervised methods

- Input: a set of features, and a few labeled data points
- Combination of supervised + unsupervised
- Active learning to find best data to label
- …



Classification

- Example: Credit scoring
- Differentiating between low-risk and high-risk customers from their *income* and *savings*



Discriminant: IF *income* $> \theta_1$ AND *savings* $> \theta_2$
THEN low-risk ELSE high-risk

Classification: Applications

- Aka Pattern recognition
- Face recognition: Pose, lighting, occlusion (glasses, beard), make-up, hair style
- Character recognition: Different handwriting styles.
- Speech recognition: Temporal dependency.
- Medical diagnosis: From symptoms to illnesses
- Biometrics: Recognition/authentication using physical and/or behavioral characteristics: Face, iris, signature, etc
- Outlier/novelty detection:

Face Recognition

Training examples of a person

Test images

ORL dataset,
AT&T Laboratories, Cambridge UK

Regression

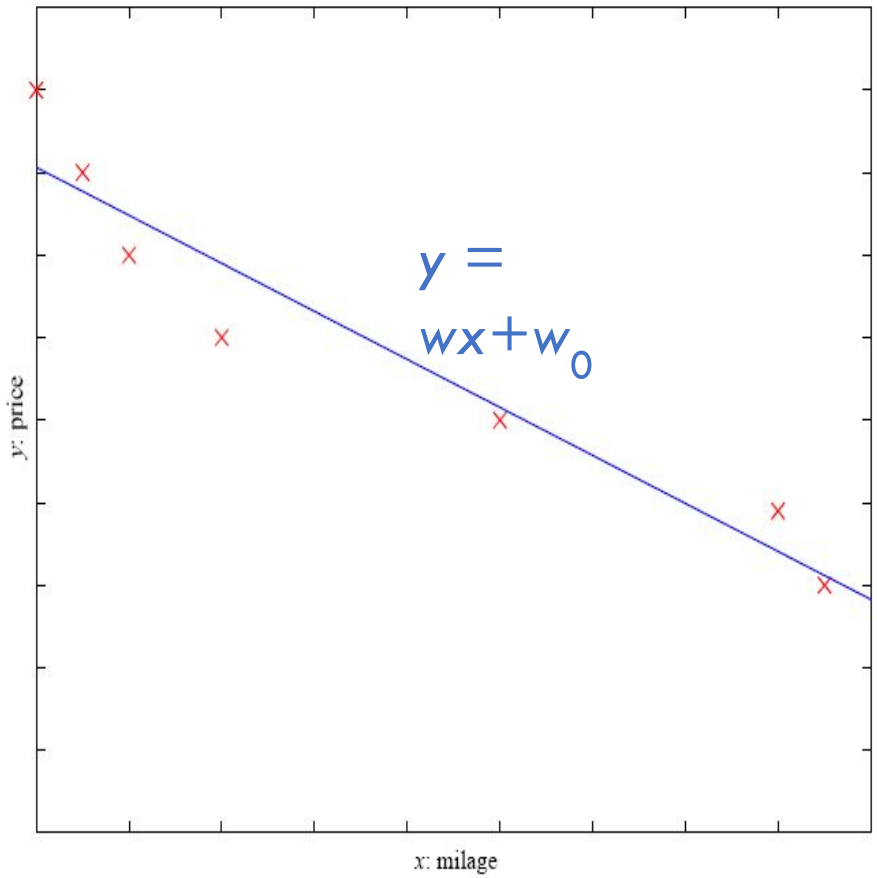
- Example: Price of a used car
- x : car attributes

  y : price

  $y = g(x \mid \theta)$

  $g(\ )$ model,

  θ parameters

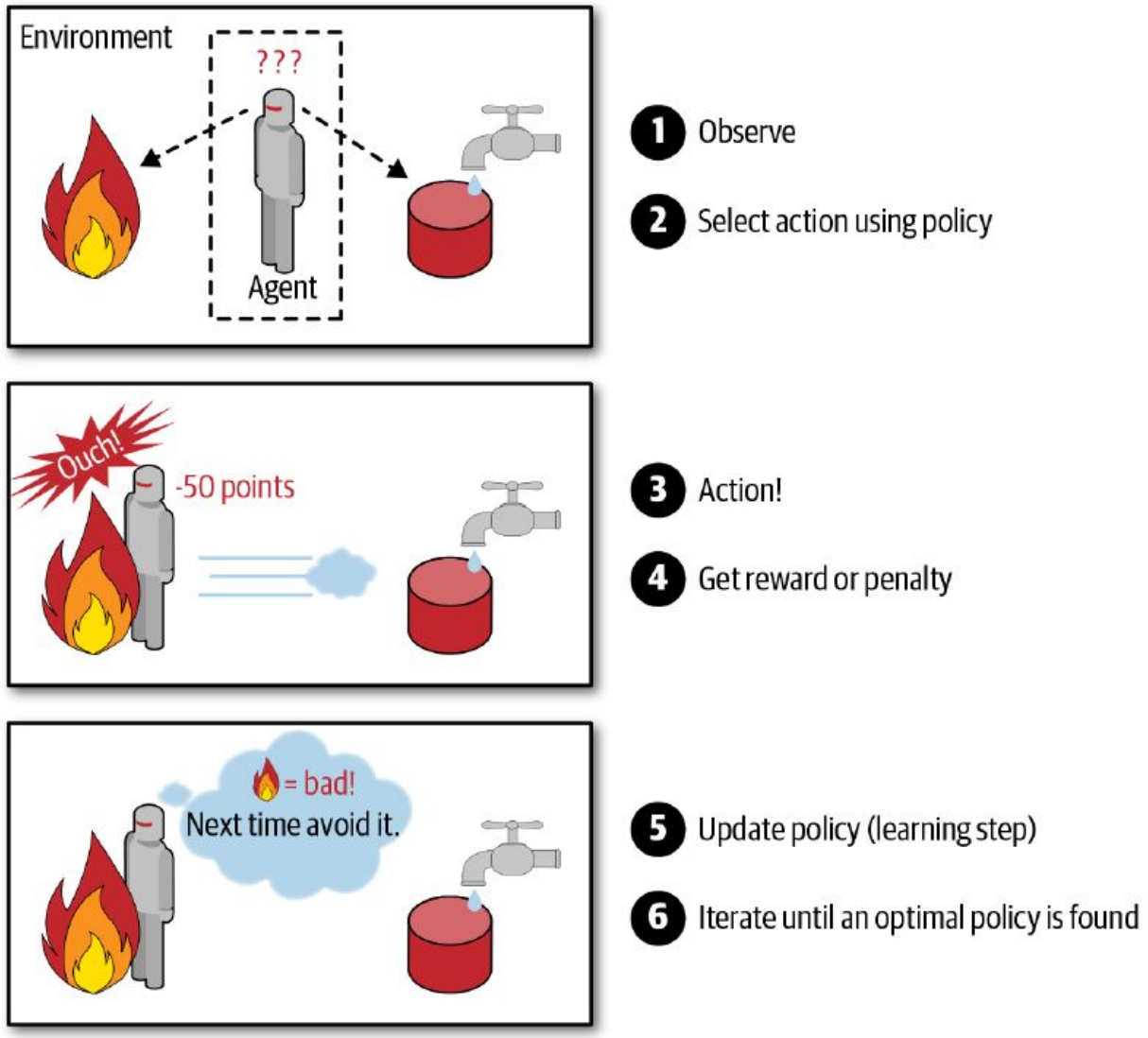


Supervised Learning Applications

- Prediction of future cases: Use the rule to predict the output for future inputs
- Knowledge extraction: The rule is easy to understand
- Compression: The rule is simpler than the data it explains
- Outlier detection: Exceptions that are not covered by the rule, e.g., fraud

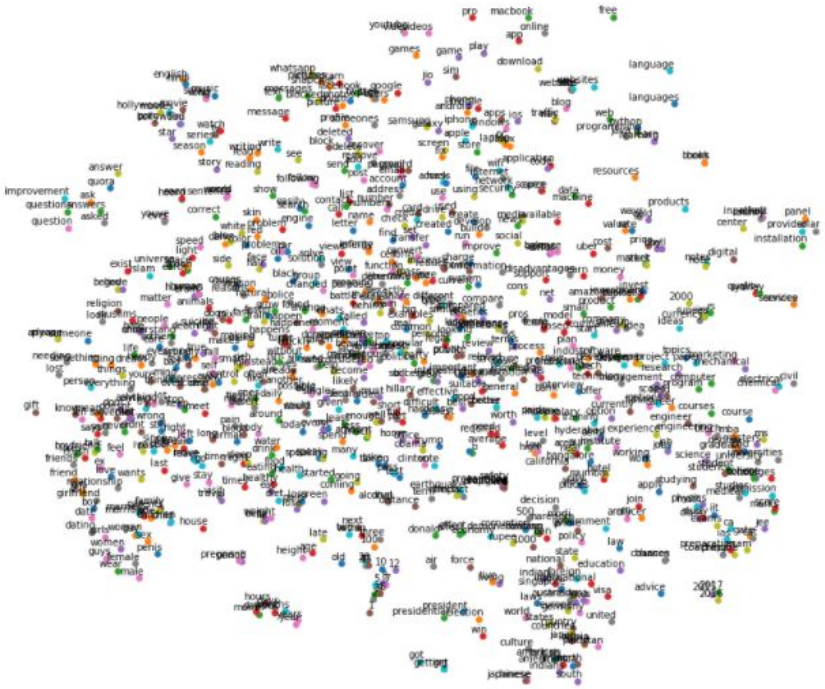
Reinforcement Learning

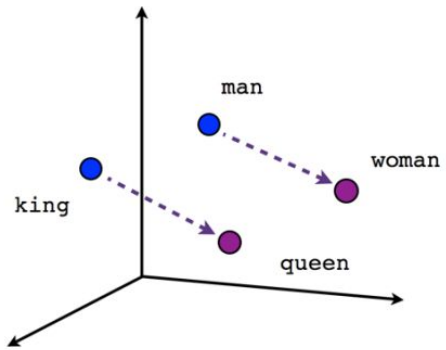
- Learning a policy: A sequence o outputs
- No supervised output but delayed reward
- Credit assignment problem
- Game playing
- Robot in a maze
- Multiple agents, partial observability, ...



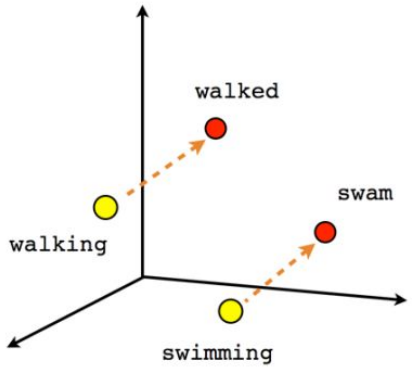
High dimensional data: NLP

Word Embeddings

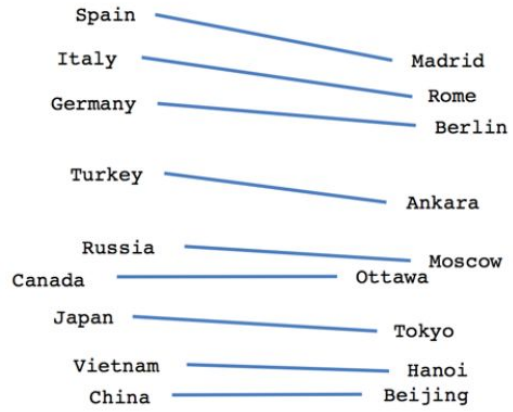




Male-Female



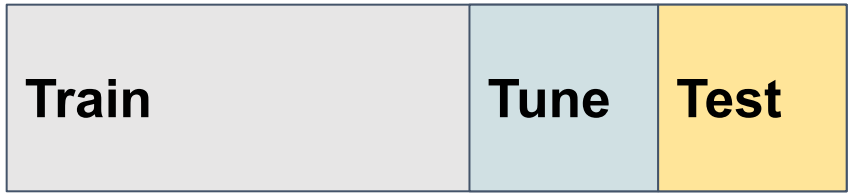
Verb tense



Country-Capital

How well does a model work?

- Testing simple models
  - Train on part of the data, test on held-out data
  - Testing data has labels (ground truth)
  - Repeat k times on random samples
    - k-fold cross validation
- Testing complex models
  - Some models require parameter tuning
    - Train on part of the data
    - Tune on separate data
    - test on held-out data
- Measure performance on test data
  - Accuracy, precision/recall, F1, ROC, AUC



Evaluation

- Precision

$$pr = \frac{TP}{TP + FP}$$

- Recall

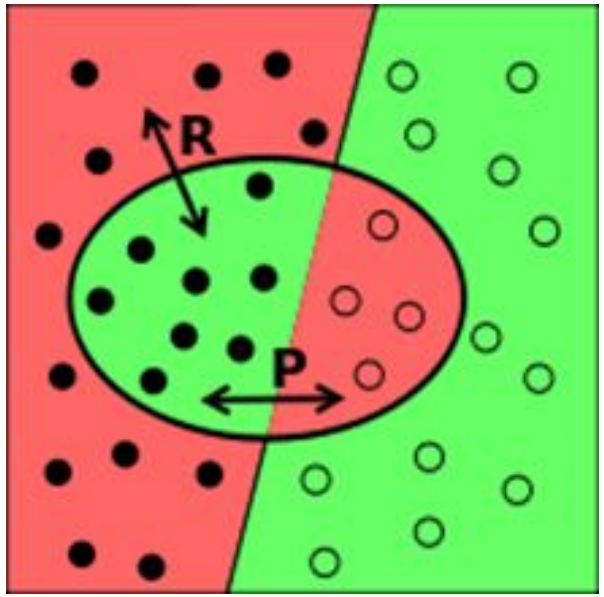
$$re = \frac{TP}{TP + FN}$$

- Accuracy

$$ac = \frac{TP + TN}{TP + TN + FP + FN}$$

R = positive examples of X

P = model predictions of X

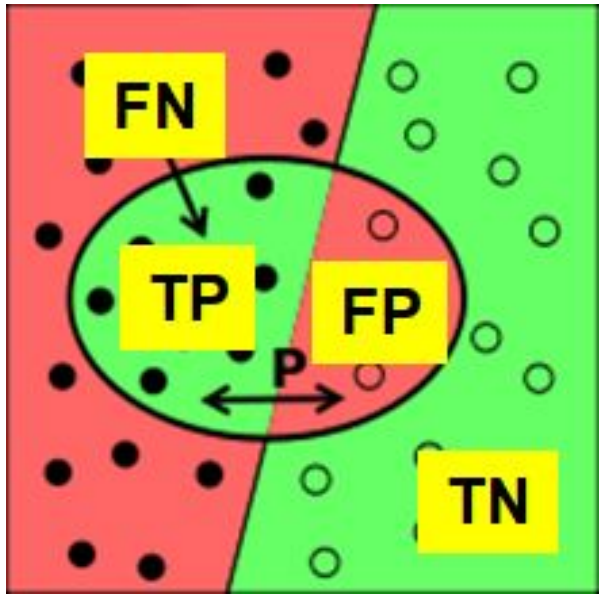


TP = true positives

FP = false pos.

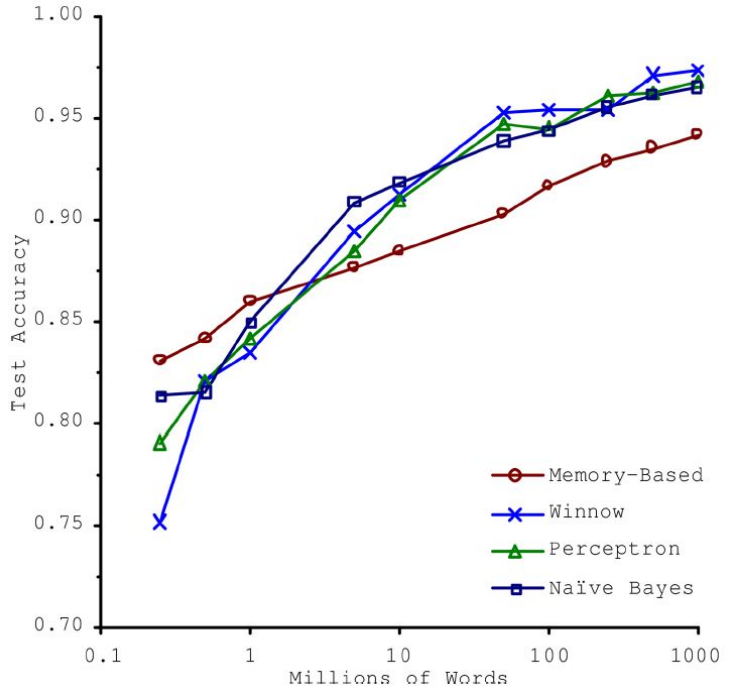
TN = true negatives

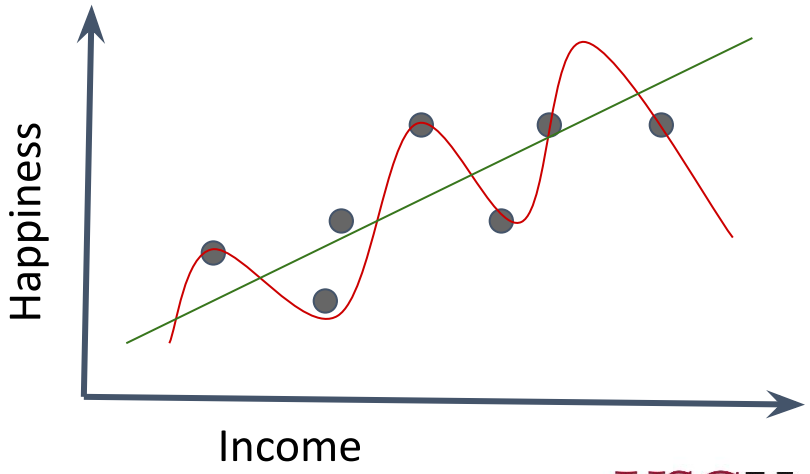
FN = false neg.



Challenges

- Insufficient training data
  - NLP, Neural nets are data-hungry
  - With enough data, all algorithms perform similarly well
- Non-representative training data
  - Biases
- Overfitting training data
  - Model too complex for the data - can detect subtle patterns, including patterns in the noise!
  - Will not generalize to new instances
  - Simplify the model, regularize
- Underfitting the data
  - Model is too simple for the data





USC Viterbi
School of Engineering

der was misidentified in **up to 7 percent of lighter-skinned fema**
photos.

sidentified in **up to 12 percent of darker-skinned male**

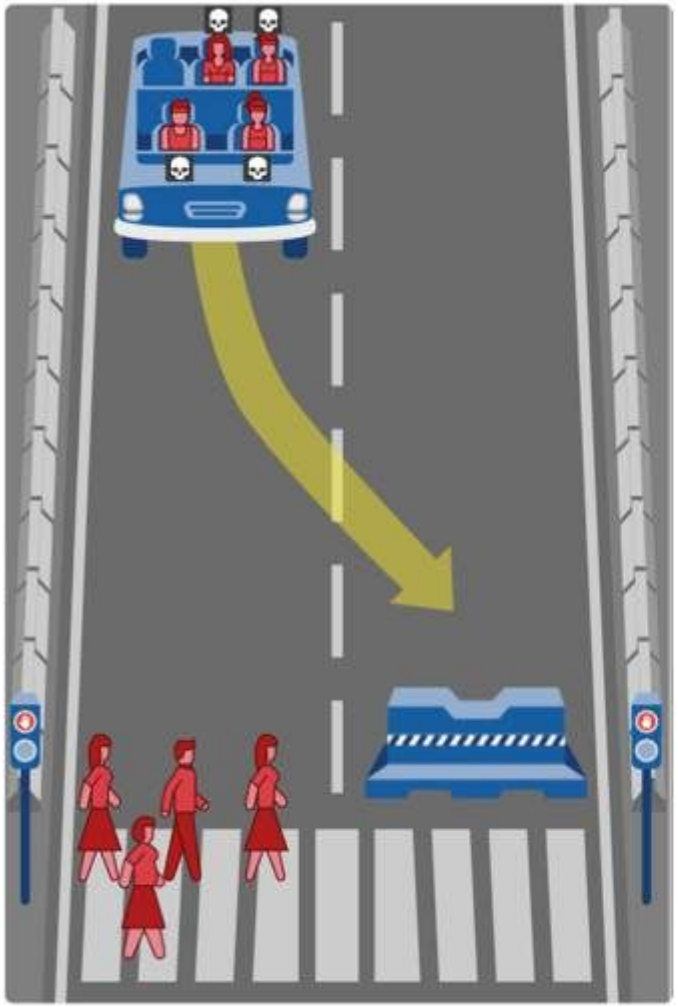
misidentified in **35 percent of darker-skinned females**

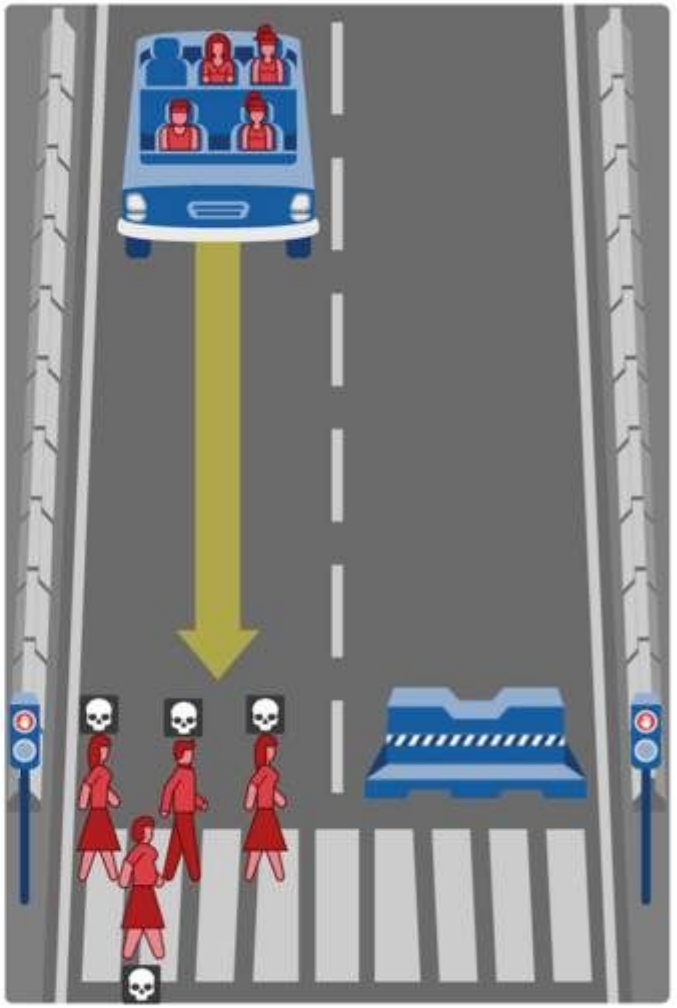
ler was misidentified in **up to 1 percent of lighter-skinned males**
photos.

Bias in AI: Errors in gender classification

https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html

AI & ethics: What should a self-driving car do?





USC Viterbi
School of Engineering

Other ethical dilemmas

- AI infers gender
  - ads for high-income jobs are presented more often to men than to women
- AI infers race
  - ads for arrest records are significantly more likely to show up on searches for African-American names
- AI infers a person's income
  - Alternative 1: provides better product recommendations
  - Alternative 2: uses the information to set product prices
- AI infers a person's health state
  - Alternative 1: recommends lifestyle changes to improve health
  - Alternative 2: uses it to set health insurance rates

Resources: Datasets

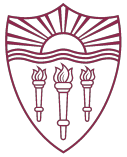
- Kaggle:
  - https://www.kaggle.com/competitions
- UCI Repository:
  - http://www.ics.uci.edu/~mlearn/MLRepository.html
- Statlib:
  - http://lib.stat.cmu.edu/

Resources: Journals

- Journal of Machine Learning Research www.jmlr.org
- Machine Learning
- Neural Computation
- Neural Networks
- IEEE Trans on Neural Networks and Learning Systems
- IEEE Trans on Pattern Analysis and Machine Intelligence
- Journals on Statistics/Data Mining/Signal Processing/Natural Language Processing/Bioinformatics/...

Resources: Conferences

- International Conference on Knowledge Discovery from Data (KDD)
- International Conference on Data Mining (ICDM)
- International Conference on Machine Learning (ICML)
- European Conference on Machine Learning (ECML)
- Neural Information Processing Systems (NeurIPS)
- Uncertainty in Artificial Intelligence (UAI)
- Computational Learning Theory (COLT)
- International Conference on Artificial Neural Networks (ICANN)
- International Conference on AI & Statistics (AISTATS)
- International Conference on Pattern Recognition (ICPR)
- …



Course details

Where to find Professor Lerman

- Research Professor
  USC
  Computer Science Department

- Principal Scientist
  Information Sciences Institute

- Email: lerman@isi.edu

- Office hours: 30 minutes after class

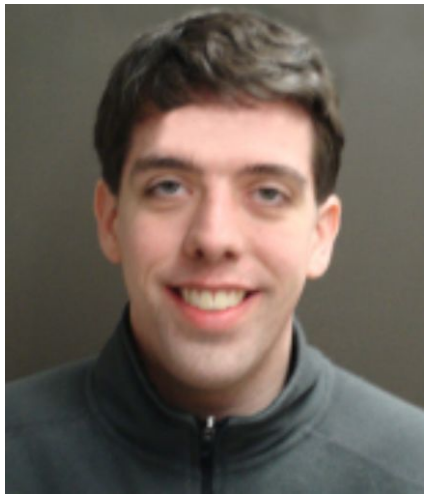
USC Viterbi
School of Engineering

Where to find Dr. Burghardt

- Computer Scientist
  USC Information sciences Institute

- Email: keithab@isi.edu

- Office hours: 30 minutes after class



USC Viterbi
School of Engineering

TA

Ashwin Rao

- Office Hours: **Wednesday 1 - 3**
- Email: mohanrao@usc.edu

Course Communication

- Blackboard – blackboard.usc.edu
  - Your USC login works on this account
  - If you are registered for 552, you will have access
- All assignments will be submitted through
- Slack - Course Discussion, ask questions (#general)
- All questions should be posted (not emailed!)
  - If you know the answer to a posted question, please try to provide helpful suggestions

Learning objectives

- Analyze quantitatively and qualitatively real-world datasets.
- Describe and compare standard machine learning algorithms.
- Choose or design learning algorithms suitable for a particular task.
- Train and evaluate machine learning models.
- Detect and assess biases in both datasets and trained machine learning models.
- Design a full machine learning pipeline.
- Create a technical report describing your work and presenting your results.
- Create a peer-review.
- Present your findings in the form of a short presentation.

Readings

- Main textbooks
  - Ethem Alpaydin, *Introduction to Machine Learning*, 3rd Edition, MIT Press (2014).
  - Aurélien Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 2nd Edition, O'Reilly (2019).
  - Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, *An Introduction to Statistical Learning*, Springer (2013, corrected at 8th printing 2017).
  - …
- Journal & conference papers for newer topics
- Please read all required readings before the class they are covered

Pre-requisites

- No formal pre- or co-requisites
- Basic knowledge of programming is required (Python 3 recommended but familiarity with any major programming language can be sufficient).

Workload: Weekly Quizzes

- Each Monday, we will publish a short quiz on Blackboard. The quiz will cover material for the previous week. You will have 7 days to complete each quiz. You can submit answers until closing time on the following Monday at 10 a.m. PT (Pacific Time).

★ The *first* quiz closes on Monday, January 25, 2021, at 10 a.m. PT.

★ The *second* quiz closes on Monday, February 1, 2021, at 10 a.m. PT.

★ …

★ The *twelfth* quiz closes on Monday, April 21, 2021, at 10 a.m. PT (see the full schedule below).

- Late Policy: We will discuss answers to the quiz in the Monday lecture. Therefore, late submissions will receive 0 points.

Workload – Homework (bi-weekly)

•Every second Wednesday we will publish a homework, which will require you to analyze a dataset, design a learning algorithm, or train and evaluate a machine learning model. The code must be published on GitHub and technical reports in the pdf format. You will all have 14 days to complete it. The deadline for uploading solutions is on Wednesday at 10 a.m. PT.

★ The deadline for the *first* homework is Wednesday, 2/4/2021, at 10 am PT.

★ The deadline for the *second* homework is on Wednesday, 2/18/2020, at 10 am PT.

★ The deadline for the *sixth* homework is on Wednesday, 4/15/2021, at 10 am PT

- Grading rubric: see syllabus
- Late policy: 2pts off each day

Workload - Projects

- An independent research project based on what you have learned in class
- An ideal project is one that you could publish a paper about
  - Empirical validation is important – show that your method beats state-of-the-art
  - Be creative!
  - Have fun!
- Teaming
  - 2 students
  - No individual projects

Project timeline

1. Prepare and post a work plan **by Wednesday, January 27, at 10 a.m**. Then Choose your topic.
2. Prepare a literature review **by Wednesday, February 10, at 10 a.m**.
3. Make a plan for doing your project. Submit your outline **by Wednesday, February 24 at 10 a.m**.
4. Write the first version of your article. You should have **an early draft by March 10**.
5. Proofread your article. Make sure that all key terms are defined. Make sure that the article has the right structure (abstract, introduction, the main content, discussion/summary, and bibliography). Remember, that the list of references at the end of your paper is not enough - your sources must be cited in the article (see the template that will be distributed).
6. Prepare a pdf of your article. Make sure that your name, affiliation, abstract and paper title are visible on the first page. Submit the pdf using a Blackboard **by Wednesday, March 24, not later than 10.00 a.m.**
7. Choose two articles prepared by your peers. Read those articles. Using the Blackboard forum, give each author suggestions on how they can improve the papers. To make sure that each person will receive an equal number of comments, only the first two comments under each project will count for credit (though you are still welcome to give comments to more than two papers if you wish; it will just not count as extra credit). You should complete this action **by Monday, April 5, at 10.00 a.m.**

Project timeline

8. Read the suggestions that you received from your peers. Address them (either incorporate the suggested changes or challenge them, describing why you think those changes would not improve the quality of your article).
9. Submit your final article.
10. Record a short summary of your work (2-3 minutes), either as a video-presentation or a narrated slideshow. Submit both your video and the final version of your article by **Wednesday, April 28, no later than 10.00 a.m.**

1.

Grading

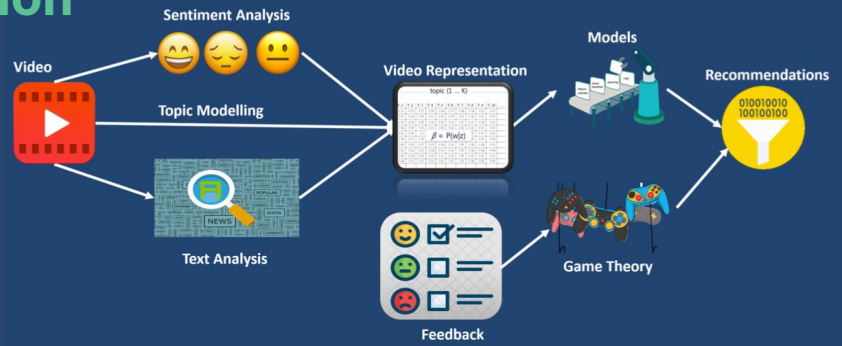
Course Element	Points
Weekly Quizzes(12)	120 (=12x10)
Bi-Weekly Problem Sets (6)	120 (=6x20)
Literature Review	20
Project Outline	10
Project Draft	20
Peer Reviews	20
Student Project	100
Final Presentation	20
Academic Reflection	10
TOTAL	**440**

Example Project

Diversification of YouTube Recommendations

Introduction

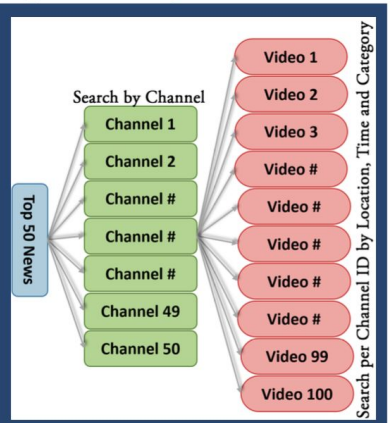
- Teenagers are heavily influenced by youtube videos.
- Youtube aims at maximizing time spent on platform and in turn castigated for guiding viewers towards extremist, inappropriate content or hoax/click bait videos e.g. Flat Earth Theory.

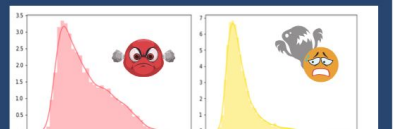


- Our project is mainly targeted towards solving this problems and giving the user a holistic view
analysis to g

Data

- Using Youtube data API we collected the initial top 50 news videos (seed)
- Using Youtube Search API we collect channel ID and for each video and search videos present for that channel
- Thus recursively calling search API we collected around 80K You tube videos with meta-data including likes, dislikes, comments, views, description, title and video closed caption.
- Carryout cleaning and explanatory data analysis on the initial corpus of data





Evaluation

- We made an interface for user's to provide their subjective opinion on the recommended videos
- The user feedback helps to fine tune machine learning and game theory framework
- Historical user data can be used to make better recommendation reducing the problem to a collaborative

USC Viterbi School of Engineering

- Detect or quantify bias in popular datasets;
- Compare or benchmark some techniques that promise to reduce bias;

Possible starting points: Joy Buolamwini, “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification” (2018), Moin Nadeem, Anna Bethke, and Siva Reddy, “StereoSet: Measuring stereotypical bias in pretrained language models” (2020), Jungseock Joo and Kimmo Kärkkäinen, “Gender Slopes: Counterfactual Fairness for Computer Vision Models by Attribute Manipulation” (2020) or Rachel Rudinger et al. “Gender Bias in Coreference Resolution” (2018).

Project Paper

- Length: ~3500 words
- Recommended Format: Latex – same as project proposal
  - Use Overleaf template : https://www.overleaf.com/read/crjbtrfftfhg
- Content:
  - Title block
  - Abstract
  - Body
  - References

Title Block

- Title of the paper
  - Choose a title that highlights the contribution of your paper
  - Keep it short (<16 words)
  - Title should summarize the main outcome of the study
- Full names of the authors
  - Address includes affiliation:
    - University of Southern California
    - Computer Science Department
    - Los Angeles, CA 90089
    - Your email address

Abstract

- An abstract is a 100-250 word summary of your paper
  - Identify the problem you are solving
  - Describe your solution
  - Summarize the results that support your solution
- An abstract is NOT the same as the introduction
  - NEVER repeat the abstract in the introduction
  - You might restate portions of it in the paper, but using different words

Introduction

- The Mini Paper
- The formula
  - The **problem** and why is it important
  - **State of the art** and its failing (~Related Work)
  - Your **contribution** (~Methods)
  - Its **benefits** (~Results)
  - End with the **big picture** (~Conclusion)

Body of Paper

- Introduction
- Sections might include:
  - Motivating application or example
  - Approach or Methods
  - Results or **Evaluation**
    - **Evaluation is one of the most important components of the paper!**
  - Related work
- Conclusion/Discussion
  - Summarizes the contribution, including any conclusions and directions for future research
- Use pictures, screen shots, and diagrams

References

- References to both related work and work that you build on
- Don’t just list references at the end of the paper, refer to them throughout the text.
- Use the “named” bibliography style [Ambite, 2004].
- Bibliography
  - Ambite, Jose Luis, 2004. Planning by Rewriting, Journal of Artificial Intelligence, Kluwer Academic Publishers, 4(2), pg 27—34.

- For more tips on writing papers
  - See slides

Rules of behavior

- Attendance
  - Logistically, we recognize that it may not be possible for you to attend lectures in real-time due to time zone or family responsibilties
    - Let us know if that is the case for you
  - Lectures will be recorded. Please review them before the next class.
- No distraction
  - Minimize alarms
  - No texting, web surfing, checking TikTok
  - Pay attention! You are paying for the privilege of attending the course – make the most of it

Camera Policy

- Seeing your faces can help us pace the lectures better.
- Please keep your cameras turned on if you can!
- Many reasons why you might wish to keep your privacy or be addressing bandwidth limitations.
- We encourage the use of virtual backgrounds and earphones/headsets whenever it is possible to mitigate privacy concerns.

Writing and academic integrity

- We will follow USC policies: https://sjacs.usc.edu/students/academic-integrity/
- Is there a meaning to authors order?
  - Alphabetical order in some fields (CS theory, economics)
  - Usually, first author is one who did the work/wrote the paper
- **What is plagiarism?**
  - More than 7 words copy and pasted from another work; unless put in quotes and referenced
- Is there such a thing as self-plagiarism?
  - Yes. Do not copy & paste from your own work – instead, rewrite.

Cheating

- Not tolerated!
- No second chances – all infractions will be reported
  - First offense is automatic failure in the class
  - Second offense is suspension from the University
- Examples:
  - Turning in someone else's work
  - Copying from someone else during a quiz
  - Doing a project that uses someone else's work without giving them credit

Once the course is over

- Directed research (1-2 MS or Phd Students)
- Research Assistantships (Phd Students)
  - We can also recommend you for positions in other groups
- Teaching Assistantships (for PhD students)
- Recommendation letters (anyone that gets at least an A-)
- Recommendations at companies affiliated with USC
  - Other companies are often looking for students