Applied Machine Learning!!!

W207 Section 9
Rasika Bhalerao
rasikabh@berkeley.edu

Aug 23: Welcome!
Nov 8 and 22: No classes

Schedule

Supervised learning methods

	Sync	Topic
2	Aug 30	Linear Regression / Gradient Descent
3	Sep 6	Feature Engineering
4	Sep 13	Logistic Regression
5	Sep 20	Multiclass classification / Eval Metrics
6	Sep 27	Neural Networks
7	Oct 4	KNN, Decision Trees, Ensembles

Unsupervised learning methods

	Sync	Topic
8	Oct 11	KMeans and PCA
9	Oct 18	Text Embeddings
10	Oct 25	CNNs
11	Nov 1	EDA, Real data, Baselines
12	Nov 15	Fairness / Ethics
13	Nov 29	Fancy Neural Networks
14	Dec 6	Final Presentations

Assignment Schedule

Due Date	Assignment
Aug 28	HW1
Sep 4	HW2
Sep 11	HW3
Sep 18	HW4
Sep 25	HW5
Oct 2	HW6
Oct 16	Group project baseline
Oct 23	HW8
Nov 6	HW9
Nov 20	HW10
Dec 4	Final project notebook + presentation

Hopefully everyone has signed up for a group by the end of this week!

https://docs.google.com/document/d/1R3J_X1Rz6WP8eMQ2cyMC0wAr5iQdhMK_httdoNO6L0w/edit?usp=sharing

Behavior expectations

- Healthy disagreement is expected
- Be mindful of one another’s schedules
- Be a good listener
- Have fun in a professional manner
- Share related real-world experience
- Ask questions when something is confusing
- Keep it 100 but be respectful
- Be open-minded to new ideas in the real world and when coding
- On time for group meetings

View options (at the top) → Annotate

Async Practice Quiz Questions (vote!)

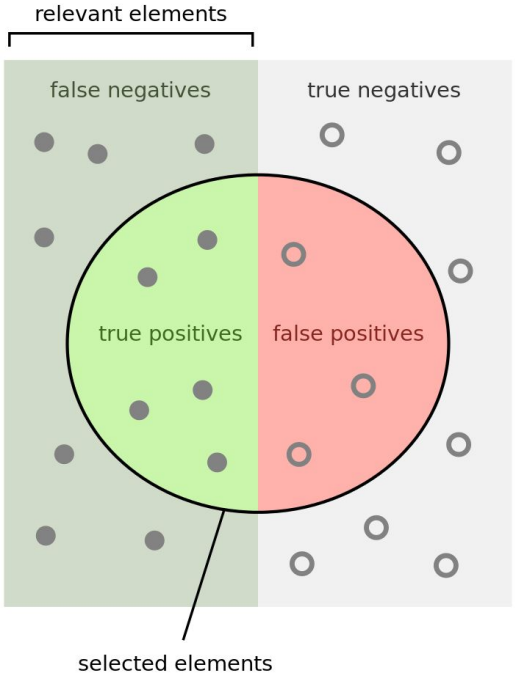
Linear regression cannot be applied to binary-valued labels.	True	False
The logistic function maps any real-valued number to the probability scale.	True	False
For a given model, different values of the classification threshold will always produce the same accuracy.	True	False
Log loss is undefined when $\|y-y'\|=1$.	True	False

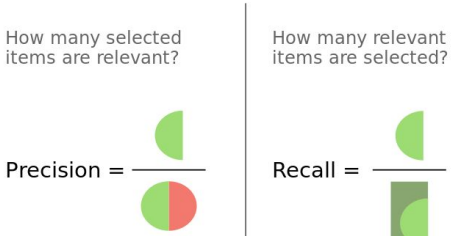
Precision, Recall, and F1

	Ground truth positive	**Ground truth negative**
Tested positive	True positive (TP)	False positive (FP)
Tested negative	False negative (FN)	True negative (TN)

- Precision
  - Out of those tested positive, how many are truly positive?
  - TP / (TP + FP)
- Recall
  - Out of those truly positive, how many tested positive?
  - TP / (TP + FN)
- F1

$$\frac{2}{\mathrm{recall}^{-1} + \mathrm{precision}^{-1}}$$





https://en.wikipedia.org/wiki/F1_score

Linear Regression versus Logistic Regression?

What is the output of Logistic Regression?

Notebook!

To access later:
https://github.com/MIDS-W207/rasikabh/blob/main/live_sessions/Week4.ipynb