[Name removed for privacy]

CS 4375.003

https://gh-ml-class.github.io/cs-4375

a. Machine learning involves training computers to recognize patterns in data and use those patterns to automate tasks that would be difficult or impossible for a human to manually program a computer to do.
b. An abundance of data is essential for machine learning, as it is the source of all information which is "learned" by the machine. Pattern recognition methods enable the machine to actually learn from the provided data and determine consistent criteria for making predictions. Of course, inaccurate predictions are of no use, so the accuracy of the predictions must be measured to determine whether the machine learning model is worth using.
c. Machine learning is a type of artificial intelligence.
d. Demucs (https://github.com/facebookresearch/demucs) is a machine learning model that is capable of separating individual instruments/vocals from songs. DALL-E (https://github.com/borisdayma/dalle-mini) is a model which can generate artistic and realistic images from text. Humans have not yet manually invented a decent algorithm for either of these problems, as they are so complex to understand, but machine learning enables these problems to be modeled based on huge amounts of data that have been collected.
e. Observations are instances of sample data; for example, a row in a table. Features are the components of an observation; for example, a column of that row. Quantitative data are measured numerically and qualitative data are non-numeric. The data collection and preparation stage of the machine learning pipeline involves gathering observations and determining the features which should be observed. These features may either be qualitative or quantitative.
f. For a long time, machine learning has interested me personally but I am lacking in mathematical/statistical skills so I've always struggled to learn these concepts on my own. I'm particularly fascinated by creative multimedia uses of ML, such as image generation and audio source separation. I've built small utilities and demos using MediaPipe and Demucs, and over time have wanted to understand how these models work so I can build additional models for my own use case (if possible).