

1) I hope to learn the whole data science pipeline; how to frame the business problem, clean the data, visualize the data, selecting the models, testing, and deploying the solution. And also, how to take this process and make it scalable across a company. Due to the rapid growth of data, I also hope to learn how to process diverse types of data and big data sets that are potentially too big for memory on a regular computer. Some of the ideas in machine learning are very technical and difficult to explain in layman terms. I hope to learn how to effectively communicate machine learning concepts to non-technical people and also handle machine learning interview questions. By the end of the semester, I hope to be very inspired by machine learning.

2) I think machine learning is such a popular topic because of the current availability of data and the rapid increase in computing power. Some of these machine learning techniques have been known long ago, but there wasn't the resources to make them mainstream and effective. Right now technology is such an integral part of our lives that there's more opportunities for us to interact with computers, leading to companies exploring machine learning. Also, there's no denying that machine learning is heavily marketed as data scientists are called the sexiest job of the 21st century, so that leads more and more people continuing to jump in. The type of machine learning problem this is

3) I am still not exactly sure the direction of my project yet. One possibility is to do something related to YouTube videos. I can get the data from Kaggle, data.world, or potentially other places if I find any. There is the 'Trending YouTube Video Statistics' on Kaggle. This is an interesting area to explore because many people are able to make money on the side or even full time on YouTube. YouTube is a platform that's dominated by big players and new content creators have trouble breaking in. If new YouTubers can better understand how to grow their channel and views, it would be very valuable in saving time. The type of problem may depend on how it's framed. Trying to predict the number of views and subscribers is a regression problem. This problem can be applied to other situations where someone needs to rapidly grow their content like on Facebook, LinkedIn, etc.

4) I think AI is very much real. Many of the top companies are investing greatly into AI, so these companies (with lots of smart people) probably have done their work to make sure there's a good chance AI is the future. More concretely, we can look at how AI is used in China. The streets of major cities have cameras with advanced facial recognition that can track down anyone. Shenzhen has many self driving buses where the driver only acts as a safety in case of system failure. The 'super app' WeChat can do pretty much everything in one place. While there is no denying AI has brought convenience to people in China, this has led to the Chinese government being able to monitor and punish anyone who engages in behavior 'not approved by the government' (religion, protests, criticism, etc.). It's clear AI is capable of achieving unimaginable things, but it's important for people realize that there may be definite downsides in regards to privacy and freedoms.