

Overview of Machine Learning

Pst170030 / Parker Tate

CS4375.004

(a) Machine learning is a subset of artificial intelligence that focuses on creating systems that analyze and learn from data, as opposed to being hard coded like a traditional program. **(b)** In machine learning data is extremely important as it is the input for the models to learn. Pattern recognition is identifying patterns in said data to further identify relationships between the data points. The accuracy of a model is how well it can make predictions over a given data set and is the most crucial factor in evaluating a machine learning algorithm.

(c) Machine learning is a subcategory of the broader field of artificial intelligence, as AI wants to (eventually) create truly intelligent systems. ML on the other hand only learns from data and doesn't have the wider scope of understanding that a true artificial intelligence would have.

(d) An example of a modern machine learning application is image/facial recognition used on social media platforms. People will post pictures and automatically tag their friends present in the picture using a data set of pictures already posted of them. This isn't possible using traditional programming because although understanding and differentiating two people apart is very intuitive for humans, it is extremely difficult to express algorithmically. Another example that we discussed in class is ChatGPT, which is a conversational language model that generates human-like responses to text prompts. That description itself was generated by ChatGPT when I asked about it. For a program to read in any possible prompt and have a dedicated response is not feasible to hard code as it would consist of a near infinite number of if/else statements.

(e) An observation in machine learning is a single point of collected data, which may contain multiple features, which are properties of an observation. An example being students (observation) and their grades, parental income, hours spent studying per week, etc. (features). These are used in machine learning to draw relationships and make predictions. Quantitative data is data that can be measured numerically, while qualitative data is data that is non-numerical or categorical. These two types of data are used in supervised and unsupervised ML algorithms respectively.

(f) I honestly don't have much of a personal interest in machine learning as I've found that the data management side of things bores me. I do, however, understand the importance of it and the role it will play in the coming decades and want to know more about it to broaden my understanding of the general programming fields. I've had some personal project ideas in the past, but those ideas revolved around ways to make user data more obscure and harder for big companies that harvest it to use. Most notably a way to temporarily swap social media accounts with a large network of people so that the data collected over a period of time would consist of dozens of users' data, rendering it much less useful/harder to sell.