Valari Graham

CS 4375

Dr. Mazidi

8/29/2022

# Overview of ML

**a.** **Machine learning** is used in data analysis for the purpose of training computational models for algorithmic purposes.

**b.** **Data** is essential to learning and making decisions. Real-world data is messy and disorganized. Therefore, more time will be spent gathering and cleaning data to make it useful for machine learning. From organized data, the experts decide what they expect to learn. **Patterns** can be recognized from the data which allows the algorithm to learn and enables experts to predict outcomes on future test data. These patterns are used to make informative decisions. Predictions made should hold **accuracy** or they are merely guesses. To achieve accuracy, there must be a baseline to prove the prediction is reliable. This baseline should hold for training and test data to make predictions.

**c.** Machine learning is used to create more automation in our lives every day. Machine learning algorithms become smarter as they meet new data and make further predictions. Artificial intelligence uses machine learning algorithms to make decisions on new data.

**d.** Self-driven cars use machine learning algorithms to gather real-time information from the environment around them to make immediate decisions and predictive analysis. Automated assistants, like Siri (Apple) or Cortana (Windows), are used to make technology use more convenient by asking questions or making commands. Self-driven cars and automated assistants could not be built with traditional programming due to the need for real time analytics and adaptation to new data.

**e.** In a dataset, each row is made up of data points also known as an **observation** and the columns are **features**. The target in regression analysis is a numeric quantitative value, and the target in classification is qualitative. **Quantitative** data is specifically for real-numbered data and their values result from measuring or counting. **Qualitative** data is for variables that are not measured, like a political party or class. Observations are important as they make up the substance of the data set and the features tell us what each observation is regarding.

**f.** In high school, I fell interested in statistics immediately. I loved learning about probability and mathematics. Not just the process of performing analysis, but what the analysis can help predict. Predictions are very important in making decisions and many businesses rely heavily on machine learning for guidance. I believe machine learning is the future as it influences artificial intelligence and automated processes. I want to be a part of that future. Recently, I have been interested in sports analytics, specifically player stats. I am currently working on a project to analyze the team and player stats for the NBA 2021-2022 season. This project will help me build my prediction analysis skills and create a machine learning algorithm to be used on future seasons. After graduation, I am looking to join a team of data scientists in the tech or financial industry to analyze big data and make business decisions beneficial to the growth of the company.