Alejo Vinluan

Abv210001

CS 4375.004

Overview of ML

Machine Learning is the act of a computer making predictions based on available data. When more data becomes available, a computer could process predictions more accurately.

Machine Learning relies on data to make its predictions. Data enforces patterns that the models find within the data. Furthermore, data can also help reject the predictions that models can believe to be correct. Models rely on Pattern Recognition within the data to base its predictions off. Trends within the data can be found using the Models which use Pattern Recognition. Finally, the Models create Predictions. By utilizing data and finding patterns in the data, predictions can be made. This data can find something similar within the data or predict a target value that can be quantitative or qualitative [1].

An example of a modern Machine Learning approach is utilizing Computer Vision to predict whether a cat is in a picture. By feeding a Machine Learning model millions of pictures of cats from all angles, the model can learn the general "rules" of whether or not a cat is in the photo. Then, the model can make a prediction in a brand new, unseen photo of if there's a cat in the picture or not. A traditional application could not be built that does this since pictures can be extremely dynamic. A photo could be taken from any angle. Furthermore, the subject of a photo can be different, like of a dog. The photo could even include multiple subjects, like a cat, dog, and chicken. Since photos can have multiple possibilities and angles, a traditional, rule-based application would not be able to create a prediction model on a picture.

Another example of a Machine Learning approach is a self-driving vehicle. A self-driving vehicle is a vehicle that can drive on the road without human intervention or input. A self-driving vehicle would not be possible without Machine Learning. The road is extremely dynamic and difficult to predict. Additionally, the vehicle has multiple possible actions it could take while handling the road. For example, a pedestrian exiting their parked vehicle on the side of the road would require the self-driving vehicle to swerve left. What would occur in

the instance that there was also a pedestrian walking on the left side of the road? Since the road is extremely dynamic and cannot be defined within a set amount of rules and boundaries, the vehicle must be dynamic as well. A traditional application cannot capture every scenario and situation that would occur on the roadways.

Machine Learning relies on data being organized in tables. A sample data point on the table would be called an observation. Observations can also be called instances or examples. Each column within a table is a feature. Data can be quantitative or qualitative. Quantitative data is data that is based on numbers [1]. For example, GPA, hours worked, and SAT score are all quantitative data points. Qualitative data is data that can only take on a finite set of data. A student's year, Freshman, Sophomore, Junior, or Senior are all examples of qualitative data. We could utilize all of the data available to make predictions for feature students. For example, if a student scored 1450 on their SAT score, what is their predicted GPA? All of the observations, features, and data types are important for the model to make more accurate predictions.

I am interested in Machine Learning because making predictions based on statistical models is essential to making improvements within multiple fields of science. There was an article I read of using Machine Learning to make accurate predictions on certain cancers. Creating the solution to self-driving trucks would alleviate the truck driver shortage. All of these solutions have Machine Learning at its core and would help improve multiple industries. I'd like to learn how to train models and learn how to use the proper model for a given dataset. Professionally, I'd like to be able to utilize Machine Learning to create features wherever possible.

Sources

[1] Mazidi, K. (2020). Machine Learning Handbook Using R and Python (K. Mazidi, Ed.; 2nd ed., pp. 21–24) [Review of Machine Learning Handbook Using R and Python].