Abigail Smith

ARS190011

CS 4375.004

Dr. Mazidi

Github link: https://msabigailscs4375.github.io/Portfolio-Setup/

Overview of ML

Machine learning is the process of training a computer to correctly identify patterns in data for data analysis, prediction, and/or action selection. The field itself is composed of four main spheres: computer science, artificial intelligence, statistics, and probability. While probability and statistics form the mathematical foundation for machine learning, computer science and artificial intelligence have catalyzed advancements in machine learning. For machine learning to be possible, there are a few unique requirements that programmers should understand in order to create successful, informed models.

In order for machine learning to accurately identify patterns, programmers have to understand the importance of data, pattern recognition, and accuracy. Without data, machine learning could not be possible. For learning scenarios to be created, data has to be clean and, usually, in the form of a table where each row represents one data sample and each column represents an attribute/feature of that sample. From this data, programmers must identify patterns in order to create algorithms that correctly utilize provided data to perform data analysis, predictions, and/or action selection (pattern recognition is the translation of raw data into an tool for informed predictions). Lastly, accuracy is paramount to creating a successful machine learning module. Accuracy draws the line between predictions and guesses by fortifying a program's ability to correctly identify and label data. Data, pattern recognition, and accuracy act as the foundational elements of machine learning, and without them, machine learning would not be possible.

There exists both similarities and differences between machine learning and artificial intelligence. While artificial intelligence is a sphere composing machine learning, artificial intelligence "refers to the general ability of computers to emulate human thought and perform tasks in real-world environments" (Artificial Intelligence (AI) vs. Machine Learning). In contrast, machine learning focuses on training an application to accurately identify patterns in data for data analysis, prediction, and/or action selection. This core difference in definition means artificial intelligence has a focus on creating systems that perform numerous, complex actions while machine learning focuses on performing tasks that they are trained to as accurately as possible. However, like machine learning, artificial intelligence does not rely on traditional programming practices in which the program and its algorithms have to be hard-coded with an limited set of results. Both artificial intelligence and machine learning lend themselves to a unique set of programming practices that rely on data and non-traditional programming practices to create dynamic, educated applications.

Machine learning models for both image recognition and commute predictions are two examples of modern machine learning applications that traditional programming cannot replicate. In both image recognition and commute predictions, previous data, pattern recognition, and algorithms are required to train a computer to accurately work with new input data. Unlike traditional programming, machine learning requires input data, an algorithm to be trained, and an output model to be produced. In contrast, traditional programming requires input data, a process using an algorithm, and results to be output. Since, in traditional programming, all information is explicitly written in an algorithm, information cannot be developed nor predicted from new data. Since image recognition and commute predictions rely on user-provided data (either an image or a desired transportation route), accurate results rely on machine learning to process new data and compare the new data to what the application has been trained on in order to produce meaningful results.

In terms of machine learning, there are a few terms paramount to understanding how machine learning operates functionally: observation, feature, quantitative data, and qualitative data. An observation, also called an example or instance, is simply a row in data. Each observation has a feature, or attribute/predictor, that is simply a column. Each feature can be labeled as either quantitative, a numeric value, or qualitative, a descriptor from a finite set of values. How data is labeled and accessed impacts how well a machine learning model can accurately make predictions, and as such, these elements of data must be fully understood and explored in order for developers to properly create informed, accurate machine learning modules. Without a thorough understanding of these terms, (accurate and meaningful) machine learning is not possible.

Personally, I am interested in machine learning for a variety of reasons. First, one promise I made to myself when I received the opportunity to pursue my degree was to actively explore different fields of computer science, and as such, I've eager to learn about machine learning. My second reason is that I've been fortunate to work on projects in a professional setting where my employers and mentor showed an interest in machine learning and provided the opportunity for me to attend some introduction to machine learning classes during my internship. From hearing how machine learning was becoming a desired skill and seeing how machine learning can be implemented (or is desired to be implemented) in a professional setting, I understand that machine learning would be a valuable skills for me to develop. My third and last reason is, from conversations I've had with others that have worked in machine learning, the process of creating machine learning models is something that really interests me and I'm very thankful for the opportunity to gain hands-on experience in a guided environment for machine learning. I'm thankful for this chance to learn more and explore this field of computer science.

Works Cited

"Artificial Intelligence (AI) vs. Machine Learning." *AI Engineering Columbia*, https://ai.engineering.columbia.edu/ai-vs-machine-learning/.

Mazidi, Karen. *Machine Learning Handbook Using R and Python*. 2nd ed., 2020.