OVERVIEW OF ML

By: Bridgette Bryant

ML – MACHINE LEARNING

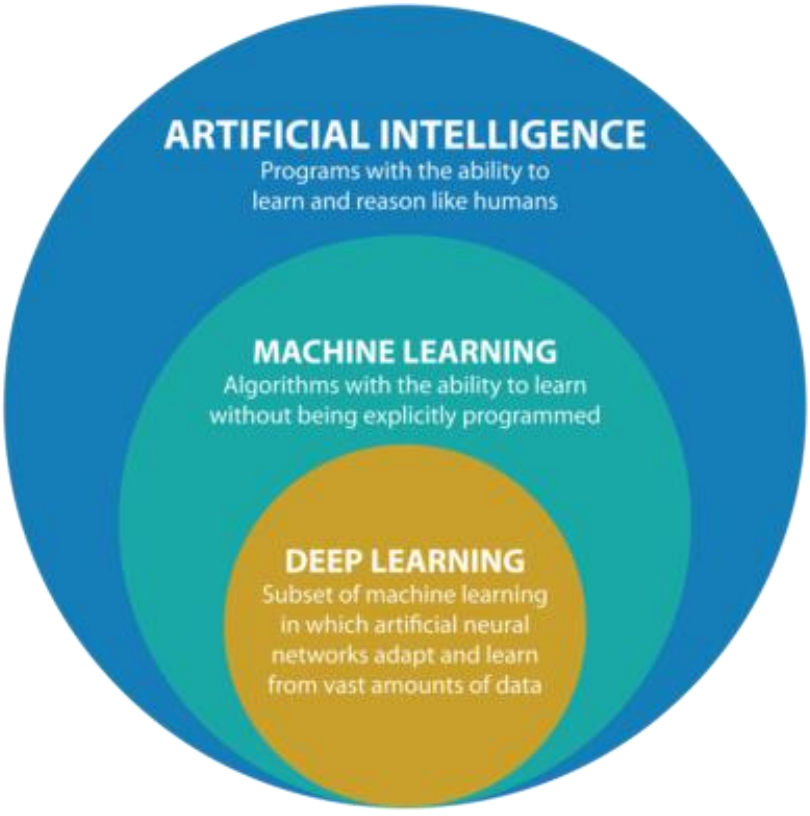
Machine learning is the process of training computers to accurately detect, analyze, and utilize patterns in data to make predictions, give feedback on the data, and/or perform actions based on the analysis autonomously.

DATA, PATTERN RECOGNITION, AND ACCURACY IN ML

If our program was bread data would be the flour used to make it. Data is what we analyze and use to create, train, test, and implement our machine learning. Without the data there is nothing to learn with or for, it is essentially the source of knowledge and purpose for our program. Pattern recognition is the program learning and analyzing the patterns in the data searching for correlations/causations. The program then uses what it has learned to give useful feedback, make predictions given incomplete data, and perform actions based on the given data. However, without accuracy the feedback, predictions, and actions would either be completely useless or possibly harmful. Therefore, accuracy is just as important as the pattern recognition and even data itself, at times having a wrong answer is much worse than having no answer. Especially when it comes to actions through machine learning, it may mass spread misinformation or even be dangerous/fatal to the public. Thus, destroying the purpose and goals of machine learning.

AI & ML

Artificial intelligence (AI) is a large umbrella for programs which mimic human behavior. This can range from making logical predictions, apps, chat bots, self-driving cars, robot factories, and even enemies in a video game. It covers a huge range of different types of programs and technology with diverse purposes. Machine Learning fits right under the umbrella as it uses data and pattern recognition to try to mimic human logic to give feedback, predict, and perform actions based on the given data/training sets. In other words, Artificial Intelligence is a vast tree with many branches, while machine learning is large branch of that tree.



Examples of Modern Machine Learning

There are thousands of machine learning applications today. One of the most popular that everyone uses is Google Maps. Google Maps is developed with machine learning by using statistics from all drivers using Google Maps to predict and update your fastest route to your destination in an efficient and accurate way (most of the time). Manually creating such a vast and extensive program would simply be impossible without machine learning. Because no matter how extensive you make your rules system, you based statistics built into your program, and even the implemented map itself. You wouldn't be able to predict and update your software fast enough to keep up with the constant changes of driving, roads, statistics, etc. to give your users an efficient and accurate app.

On the other hand, something that we all must see and may never use are the personalized adds google gives us. With that same data that assists us with driving and everything else google is constantly collecting on our devices. Any time you look up nearly any product I am sure you become swamped with related adds. To program such an app with traditional programming, you would need to write similar items for nearly every product on the internet that could be googled. Including the millions of new products uploaded every day. I am sure as you can see that would be utterly impossible, even for google. Thankfully we have machine learning so we can always see the personalized adds we love, or don't.

Observations, Features, Quantitative Data, and Qualitative Data

In a paragraph, define the terms observation, feature, quantitative data, and qualitative data and discuss their importance in machine learning

An Observation is an example or instance of a sample data point, they are represented as rows in the datasheet. These are the individual data inputs we have that make up the data, for example if we had data about cars, an observation would be a specific car from the data. A Feature is an attribute or predictor of data, there are usually several, they are represented as columns in the datasheet. These are the different categories of information we have on each sample data point. If we continue with our cars example, the different categories would be year, make, model, milage, etc.

With this data we have a target, basically is it what we are trying to achieve/predict with our machine learning. The two main types of target values in Machine Learning we have quantitative and qualitative. A quantitative value is a real-number value, it is used in Regression. For example, if you want to predict the cost of buying a car based on current cars on the market and car loans the prediction would be a quantitative value. However, a qualitative value is an unmeasurable value used in classification. For example, if you wanted to group different types of cars using the data by color, body styles, etc., the different categories created would be qualitative values. These values makeup the different targets we attempt to predict/create using machine learning, they are the program's goal and measurement tool for accuracy.

My Interest in ML

ML is very fascinating to me, and I have always been very curious about how it works under the hood and how to implement it to create useful applications. I also have a few projects in mind such as creating my own secure version of Google Home/Alexa that doesn't constantly send data, creating an app that can recognize different snakes found through your phone camera to designate if they are venomous, and possibly even creating an app that mimics google maps but for hiking for national parks. I also understand machine learning is booming in the tech industry as well.