1. Machine learning is a way to analyze and classify data in ways that normal programs cannot through pattern recognition.
2. When writing a machine learning algorithm, the first step is which data the algorithm will look over and the gathering of said data. It is important that the data be representative of whatever the algorithm will be trained to find as well as the sourcing of the data be as ethical as possible. Once the data is collected the ML algorithm will look over the data in a variety of ways predetermined by the person and look for patters in the data that can be used as predictors for some specific trait. Once the algorithm finds these patterns and makes predictions the person checks to see if the predictions match the expected predictions better than random guessing or better than just guessing with the obvious trends in the data.
3. Artificial Intelligence in a term used to describe computer systems that will get “smarter” by some measurable degree over time without the intervention of humans. Machine Learning is a specific field of AI in which machines learn to recognize patterns in data.
4. Two modern applications of machine learning that come to mind for me are Image Recognition and Advertising targeting. These two applications, like most machine learning applications, could not be achieved with typical algorithms because the large number of factors that are present in the data, as well as “invisible” factors that may not present themselves to the person writing the code, but the ML algorithm might learn to distinguish over time.
5. When it comes to training an algorithm in ML, breaking the data into different categories can be very helpful. The most common categories are qualitative and quantitative data, where quantitative data is any data that can be easily represented with a number or as a feature and qualitative data is data that can be represented with a finite set of values. An example would be average temperature over a week vs the average favorite color in a class. The first data point can be represented by a decimal value and has an infinite number of possibilities while the second requires words and is limited to the predefined spectrum of visible color for humans. In the data collected they can also be broken into features and observations where the observation is meant to be a baseline truth that is looked for by the trained algorithm and features are the points of data used to train the algorithm to find the observation.
6. My personal interest in Machine Learning is surface level, I find the topic interesting and of obvious importance in the coming years for Computer Science. I would like to learn more about the field in general for the sake of knowing it as well as seeing if I would be interested in learning more or even pursuing a career involving AI.