CS 4375.003

Jonathan Ho

Overview of ML

1. I believe machine learning is a computer “learning” to detect patterns within data using given, preset models. Over time, the computer would be able to ideally gain an increase accuracy in predicting patterns within data.

2. Data is important to machine learning because it is the essential basis of it. No machine can learn anything without data. Pattern recognition in machine learning is the aspect that allows a computer to understand what they are interpreting in the data. For example, it would help a computer distinguish if there was an actual trend in data, or if it is just coincidental. These pattern recognitions can also lead to developing predictions in the data with models. Lastly, accuracy is important in machine learning as this helps to support a model a computer determines within data. Otherwise, if there is no accuracy, then the model a computer “sees” is no better than a guess in the dark.

3. The relationship between ML and AI is that ML is a subset of AI. For example, AI is viewed as a machine being able to “mimic” human intelligence, whereas ML can gain knowledge from data like a human. Both try to encompass human capabilities, but ML focuses more on learning from the data.

4. One application of machine learning is Duolingo, a mobile application used to learn another language. Duolingo uses machine learning to collect answers from a user to create a model of how long a user would retain learned words, and if a refresher course is needed. This cannot be done with traditional programming because some users can remember a word better than others, and this applies to every single word learned. It would be impossible for a programmer to code exactly when is the optimal time for each individual user to receive a refresher lesson. Another application of machine learning is the suggestions given in a Google search bar. The search bar suggestions are most likely based on global data of the billions of searches done every day. This combined with cookies from google help to create personalized search suggestions to a user. Traditional programming would not be able to statically set personalized suggestions to each individual user that used the Google search bar.

5. Observations are a sample data point within data, such as a row within a table of data. A feature is an attribute of data, which is something that is common throughout all the data like a column in a table of data. Quantitative data is numerical features of data which is only given through numbers. Qualitative data is descriptive features of data which is described through language. Each aspect of the data is important in making up a data set to use for machine learning. Observations in machine learning help the machine determine what attributes make up one data which may help in distinguishing one data point from another. A feature would help to determine what set of numerical or descriptive data belongs to a certain attribute. Finally, qualitative and quantitative data can be used to potentially improve accuracy and predictions for a machine so it may produce better models of the data.

6. Throughout my life, I occasionally would hear of the term “machine learning”. It was something I knew was a growing field, but I did not understand what it was. When I heard this class was being offered, I thought it would be a great opportunity to learn more about machine learning and how it is applied to everyday applications. I hope by the end of the semester, I would have not only gained a deeper understanding of machine learning, but I would be able to apply what I have learned. I believe it

would be nice to learn about the professional applications of machine learning as it would help to answer my own personal question on understanding the field. However, the way the course sounds like it is being taught with project-based learning, it sounds like a great hands-on way of understanding machine learning. Outside of the class, if I do not use it in industry, I hope to be able to do a fun personal project which incorporates machine learning.