IST 707 Applied Machine Learning

HW1: What is Machine Learning
Name: Jainish Savaliya

Task 1: What, Why and Where?

1. **What is machine learning?**
   Machine learning is the field where computer algorithms and data is used to get the intelligent actions. There are different methods and machine learning models which can be used for various use cases. Machine learning can be divided into two subcategories which is supervised and unsupervised machine learning. In the supervised machine learning it uses the labeled dataset which trains on the algorithms and gives the insights and forecasting. The unsupervised machine learning is used using untagged data to identify different patterns .

2. **Where can you use machine learning?**
   Machine learning has vast applications. It can be used to built AI based applications in the real world such as self-driving cars, , deep learning, forecasting , prediction , Financial services , sale & marketing , smart city initiatives, strategy , decision making Technology etc. Recent examples of the machine learning are chat GPT, DAL e, Tesla self-driving cars etc.

3. **Why should we use machine learning?**
   Machine learning increases the efficiency , Productivity in multiple areas and can provide high values to make people's life better. Machine learning can be used to transform business , enhance decision making & modernize systems. On the business side machine learning can increase productivity through automation, improve customer experience , innovate products and services & help developing new data based business models. It can also help detecting anomalies , reduce logistics cost by real time forecasts and behavioral coaching. In the healthcare sector it can be used to cure diseases, identify patterns and predict future pandemics.

4. **What benefits does machine learning have over other traditional methods of analysis, prediction, and decision-making?**
   Unlike traditional methods machine learning tools can help increase the productivity exponentially by using tools like automatic outlier detectors, Data cleaning, clustering, visualization by machine learning regression models. The accuracy of the prediction is much better than traditional methods. Opposite to traditional methods, Machine learning can work with much larger data and information by using various libraries which is helpful to recognize different patterns and get the meaningful insights.

5. **What business problems are machine-learning and artificial intelligence most suited to?**
   Machine Learning and artificial intelligence is most suited to business problems like , lack of technology improvement , issues with operations and maintenance , lack of understanding of customer experiences and feedbacks , coming up with the business strategy to increase the revenue , identifying perfect product market fit & development of a service , solving supply chain issues , increasing the sales & effective marketing , human resources & management , financial analysis , complexity of merging & acquisitions and areas like environment , social & governance which has problems.

Write one paragraph to summarize the criticism, and another paragraph for the defense. Write the third paragraph to offer your own thoughts, e.g., is the criticism valid? Does the defense make sense? What other problems or benefits do you see in Google Flu Trend or similar big data applications?

Task 2: Critical Thinking and Writing

1. **NYT: Google Flu Trends: The Limits of Big Data**
2. **Atlantic: In Defense of Google Flu Trends**

Criticism: The recent trend of CDC reports from doctors on influenza- like illness was made from the real data of the recent disease case and Google Flu Trends were predicted based on the previous data base collected by the systems so there might be a difference of the virus mutation or any other changes that caused the prediction to be far less accurate. So criticizing the google Flu Trends is only partially valid. Secondly, The Prediction algorithm might not work on every data set properly because of the prediction model learns with more testing over time and becomes more accurate based on the data set it was training on. So, Criticism from social scientists were valid from their point of view but from the point of view of the people who knew the technology google flu trends performed well.

Defense: A follow up analysis of four authors in the year 2013-14 that google updated its algorithm which could be improved by training the algorithm but still overshot by 30 percent which means that the algorithm is becoming more accurate by time and has a scope to increase the accuracy by time with more testing and training. Authors also mentioned that the 2 weeks recent illness CDC report would be more accurate which can be true because of the data would be new data which would provide the latest information regarding the flue. Lastly the authors' latest claim to just looking google flu trends and corelating is enough according to the recent situation would be justified because of the amount of data and the increase of the accuracy with time. As the authors said mashing up the data with CDC provides much better predictions so its can be a replacement

Thoughts: According to me the increase of the accuracy was related to testing of the data sets and improving the machine learning model. As the datasets becomes more wide and insightful and regression model improves the prediction will improve and regarding the news article disregard of the accuracy the prediction of the disease outbreak would be much better indicator to be alert and recently it can be improved at the level where as the authors said that google flu trends and correlations is enough can be very much possible to achieve maximum accuracy.