Project Proposal CPSC 481

Group Members:

Adam Harb

Shiyun Zhou

Description of Project:

Our project is on developing a Sentiment Analysis model using Python. Sentiment Analysis covers on building a system to gather and establish the emotional effect behind words. Its also a sub-field of Natural Language Processing (NLP). This is significant as it enables us to acquire an overview of the extensive public opinion behind precise topics. The project will touch on the following course of actions such as Data Collection (We will collect data from a publicly accessible datasets; in this case, an article summary or a review of an Amazon product), Model Training (We will use the preprocessed data to teach a machine learning model. The model will grasp on identifying certain words and phrases with positive or negative sentiment), and Testing and Optimization (After mastering the model, we will test it on new data to see how accurate it can predict sentiment). The ultimate objective of this project is to make a model that can accurately predict the sentiment of a piece of text whether its a summary of an article or a review of a certain product. This includes real-world applications in areas like marketing, politics, etc.

Proof of Real-World Applications:

Sentiment Analysis has a broad list of real world applications such as: Business Intelligence(Companies use sentiment analysis to grasp on customers opinion on the product and when businesses examine customer reviews, they can recognize what customers like or dislike and use those information to better their offerings), Politics(Sentiment Analysis is used to articulate the summary of an article related to politics and test if the article is positive or negative), and Social Media Monitoring(Brands use sentiment analysis to observe their social media channels. This aids them to respond to negative comments or reviews in a prompt approach, and also decipher what aspects of their product or service are admired).

Plan:

The plan for our Sentiment Analysis project is to gather articles from reliable sources such as news or blogs to form a summary. The articles will include a broad range of topics and will cover both positive and negative sentiments. We will use TextBlob which is a Python library for processing textual data, to perform sentiment analysis on the preprocessed data. TextBlob will pin on a polarity value to the text, which ranges from -1(most negative) to +1(most positive). This type of value will enable us to know the sentiment of each article. We will also test the sentiment analysis model on a list of articles. This includes articles about positive event such as a significant stock market increases or a charitable activities and articles that discusses negative events such as a violent incident or an economic crisis. The performance of the model will be judged based on how accurately it can predict the sentiment of these articles.