Name: Pureum Kim

Title: Sentiment Analysis Using Multinomial Naïve Bayes Model

GitHub Link: https://github.com/pkim93/INFO523

Presentation Link: https://drive.google.com/file/d/14UMcquRuB44yesfSbeA2AEbvv7toGO1-/view?usp=sharing

Summary:

Sentiment analysis is used to classify the emotion expressed in a given text, often classified as either negative or positive. There are multiple approaches to conducting sentiment analysis, including lexicon, machine learning, and hybrid approach. Of the three approaches, machine learning approach, more specifically multinomial naïve bayes model, is discussed in this presentation. I chose to discuss multinomial naïve bayes model because it is relevant to the topics of the course. For instance, this model is a part of machine learning technique to take a set of data to calculate the probability and make predictions. Also, this model is a supervised technique for text classification, a topic that we discussed in the last part of our course.

Multinomial naïve bayes model predicts the tag of a text, such as a review, using the Bayes theorem. The algorithm calculates each tag's likelihood for a given training sample and outputs the tag (i.e., positive or negative) with the greatest likelihood. This model is suitable for text classification problems where the output is nominal but is not suitable for regressions that involve numerical values. Sentiment analysis (in extension multinomial naïve bayes model) is applied widely in marketing. Businesses use sentiment analysis to analyze social media discussion around the company's products or services, evaluate survey responses (e.g., free response questions), and determine whether reviews are positive or negative.

Altogether, multinomial naïve bayes model is a supervised machine learning technique that is used to classify a text as either positive or negative. The model has a low processing power, and with a smaller sample set, it can outperform other types of models. However, once the sample set becomes bigger, the accuracy decreases, and other types of models, such as XLNet or BERT, outperform bayes model.