Week 4 _Review Articles

Bhadamkar, A., Bhattacharya, S. (2022). Tesla Inc. Stock Prediction using Sentiment Analysis. Australasian Business Accounting, and Finance (AABFJ). Vol 16(1). Pgs. 52-66.

The authors used Elon Musk's Tweets on Twitter to establish a relationship between his tweets and Tesla's stock performance. The authors used Exploratory Data Analysis to differential patterns in the dataset, which were preprocessed to remove any stop words and found that the increase of number of Musk's tweets corresponded top a increase in Tesla's closing price. Historically there have been a variety of approaches used to view the relationship between sentiment analysis and the stock market. For example. Mehtab, Sidra, and Jauydi in 2019 looked at social media in addition to stock prices. Mittal and Goel used sentiment analysis structured into 4 steps. Word lists, Tweet Filtering, Daily Scoring, and Score Mapping. There were also Hybrid models of Sentiment Analysis used ie., SVM, KNN and Random Forrests' were used (Seng and Yang, 2017).

In their research, the authors used Textblob, with noun phrase extraction, Tagging particles of speech and inflection and Lemmatization of Words. Then they used Facebooks Prophet to look at polarity vs subjectivity. They found in the short run, the number of tweets by Elon Musk and his interaction marginally corresponds to the Tesla Stock Price. In the long run over months and years, the Tesla Stock Price had a direct parallel association to Musk's engagement.

`Tan,K., L. Lee, C., P., and Lim K, M. (2023). A Survey of Sentiment Analysis: Approaches, Datasets and Future Research. Applied Sciences, 13, 4550.

The authors review the approaches used in sentiment analysis approaches, the datasets used and the future research areas. The Approaches used in Sentiment Analysis were categorized into Machine Learning, Deep Learning and Ensemble Learning. What was interesting were the Datasets used in sentiment analysis. These included (1) The Internet Movie Database. Which is a combination of storylines and personal opinions, the Twitter US Airline Sentiment dataset which challenges in brevity and informality which increases the misclassification of twitter tweets, and the Sentiment140 by Stanford University which like Twitter faces challenges in informality and Brevity. But it does according to the author reflects real world scenario which researchers are likely to face. Finally, there is the SemEval-217 Task 4 dataset which has challenges because it's in multiple languages.

Whatever dataset this group chooses, the limitations and the challenges need to be addressed.