SENTIMENT ANALYSIS OF NEWS ARTICLES USING BIDIRECTIONAL RECURRENT NEURAL NETWORKS

ALEXANDER TAN KA JIN

A project report submitted in partial fulfilment of the
requirements for the award of the degree of
Your Degree

Faculty Of Computing
Universiti Teknologi Malaysia

21/12/2024

Conclusions

Summary and Future Work

The study introduced it's goals for uncovering media bias and propaganda with the usage of Bidirectional Recurrent Neural Networks (RNN). The study was done under the assumption that rules-based techniques is insufficient in detection of finer patterns within news and media and inspired by the recent success of utilizing Bidirectional RNNs models for Natural Language Processing. Using rules-based sentiment analysis, the study was able to get an general overview of the corpus and was able to draw correlations between media bias and sentiment scores.

From the initial findings in the exploratory data analysis, there is a linear correlation between media bias scores from AllSides and sentiment scores in both the polarity and subjectivity areas. Using TexBlob sentiment labelling, the study was able to account for surrounding word contexts within a particular sentence and identify sensational and opinionated language within news media. The initial findings has demonstrated a moderate linear relationship between the use of sensational/opinionated language and sentences and AllSides media bias scores, which emphasizes how emotional and opinionated coverage plays an important part in media bias. In addition to the initial findings, the data was embedded for future analysis using BiLSTM and BiGRU models.

This project only targets the possibility that deep learning driven sentiment analysis has on the identification of media bias. Further research can be done using a more finer grained labeller that specifically targets the form of propaganda technique and media bias that is present within the sentences of news articles like in the study done by (Da San Martino et al., 2019).

Limitations

The thesis was limited in time and computation. The authors did not have sufficient time in doing a more comprehensive literature review and as mentioned in the Exploratory Data Analysis section, needed to split the data into smaller chunks in order to handle computation of the data.

References

Da San Martino, G., Seunghak, Y., Barrón-Cedeno, A., Petrov, R., Nakov, P., et al. (2019). Fine-grained analysis of propaganda in news article. In *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (emnlp-ijcnlp)* (pp. 5636–5646).