26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

Sentiment analysis of historical Australian parliamentary speeches

William Pincombe [a]

[a] *School of Computer and Mathematical Sciences, University of Adelaide, South Australia, 5005*
Email: william.pincombe@adelaide.edu.au

Abstract: Sentiment analysis, within natural language processing, is the study of emotions in text. Models for sentiment analysis on contemporary English texts have been extensively developed, including sentiment dictionaries, which are lists of words with information about their associated sentiment. However, using these models on historical text carries the risk of inaccuracy due to shifts in the sentiment associated with words over time. For a sufficiently large text corpus, it is possible to infer a sentiment dictionary from the text using semi-supervised label propagation and distributional linguistics (Hamilton et al., 2016). We study this approach for historical Australian parliamentary speeches, as recorded in the Australian Hansard. The Hansard consists of transcripts of parliamentary speeches, has been entirely digitised and represents a large enough corpus to apply distributional methods.

There are two parts to constructing a domain-specific dictionary. First, it requires the construction of a distributional word embedding, which captures co-occurrence information from the corpus in numerical form. We investigated two embedding methods, the neural network-based Skip Gram with Negative Sampling (SGNS) and the traditional count-based Positive Pointwise Mutual Information (PPMI) with dimension reduction using Singular Value Decomposition (SVD). Second, we specify seed words, for which sentiment is assumed, and then estimate the sentiment for each word using label propagation from the seed words. We used the domain-specific sentiment dictionary to estimate the sentiment of parliamentary speeches, to study differences in average sentiment over time and by political party affiliation.

We found that the constructed sentiment dictionary successfully detected many words with domain-specific sentiment. Our results also emphasised the importance of numerous parameter choices in text preprocessing, embedding and label propagation. Further, we estimated that speeches by members of governing parties had, on average, more positive sentiment than those by members of opposition parties.

Keywords: *Sentiment analysis, natural language processing, Australian history*

References

Hamilton William L., Clark Kevin, Leskovec Jure and Jurafsky Dan (2016), Inducing Domain-Specific Sentiment Lexicons from Unlabeled Corpora, *in* Jian Su, Kevin Duh and Xavier Carreras, eds, 'Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing', Association for Computational Linguistics, Austin, Texas, pp. 595–605.
URL: *https://aclanthology.org/D16-1057*