International Journal of Informatics and Communication Technology (IJ-ICT)
Vol. 15, No. 2, June 2026, pp. 674~682
ISSN: 2252-8776, DOI: 10.11591/ijict.v15i2.pp674-682

Harnessing NLP and AI to decode political discourse: speech patterns, sentiment analysis, and public perception

Malayaj Kumar, Anuj Kumar Singh, Soumitra Das
Department of Computer Science and Engineering, Shri JJT University, Jhunjhunu, India

Article Info

Article history:

Received Jan 25, 2025
Revised Nov 27, 2025
Accepted Dec 14, 2025

Keywords:

AI
Bertopic
Bidirectional GRU
NLP
Political discourse
Public perception
Sentiment analysis
Whisper

ABSTRACT

Using natural language processing (NLP) and artificial intelligence (AI), this study analyzes the frequencies of words and phrases in political leaders' speeches to track patterns in political discourse. The objective is to identify language patterns, sentiments, and topics of political addresses using state-of-the-art methods like automatic transcription (Whisper), Bidirectional gated recurrent unit (GRU) for sentiment analysis, and BERTopic. Through the use of Whisper's state-of-the-art transcription service, we were able to transcribe the political speeches into machine-readable text, which in turn provides for other types of analysis. Bidirectional GRU classifies sentiment as positive, negative, or neutral with the aim to study how politicians use sentiment to manipulate their listeners. Furthermore, we use BERTopic for tracking the evolution of rhetoric, key trend summarisation, and topic mining and analysis. It illustrates how politicians employ discursive strategies and epilinguistic elements to manage the public mind and reality. Achievements and objectives are framed with positive and defensive emotions aimed at threats or criticisms. The emotional grab of it all is still important. It locates in these the thematic coherence and shifting sentiment that lie at the heart of political storytelling. It shows how political communication is evolving to stay relevant in the digital media age and delivers language – even real-time language pattern tracking – via the use of AI and big data. Further study is needed of multimodal and flexible techniques for analysing political discourse across languages and time periods.

Corresponding Author:

Malayaj Kumar
Department of Computer Science and Engineering, Shri JJT University
Jhunjhunu, Rajasthan, India
Email: kumarmalayaj@gmail.com

1. INTRODUCTION

The predictable, as history teaches us, is that political speeches are routinely used as part of this broader panoply of communication strategies that may range from articulating policy positions to rallying public opinion and changing societal attitudes [1], [2]. A lot of the use of rhetoric and differences in language (such as tone) tends to revolve around evoking a sentiment or building trust [2], [3]. In this information age such as time on the web and speeches airing to the whole society, Speech Archives' history and also emotional analysis of speakers and speeches for speech by topic are becoming more important than before so needed [4], [5]. Regarding public speech; I think the research may has also interest with respect to discussive movements, especially in the polarised political world of today, rather than just about broad movement and rationale of agents [6]–[8].

The improved access to computational tools has transformed the study of political speeches [9]–[11]. Speeches naturally share commonalities and distinctive features, and natural language processing (NLP) and

AI technologies have evolved into essential tools for dissecting the linguistic and emotional dimensions contained within speeches [12]–[14]. Traditional forms of qualitative analysis may miss the opportunities available through these technologies which can reveal patterns, sentiment trajectories and thematic structures [15]–[17]. For example, sentiment analysis provides a means for gauging the emotional tone of speeches, and topic modeling provides insight into key themes and priorities of political leaders [18]–[20].

Prior work on sentiment analysis and topic modelling of political communication usually focused on individual aspects and was mostly focussed on written text, for example, manifestos or social media messages. To the best of our knowledge, no work has combined automatic speech transcription, deep learning-based sentiment classification, and transformer-based topic modeling into a coherent pipeline for spoken political discourse analysis. This research addresses that gap by using a range of advanced transcription capabilities of whisper, the Bidirectional gated recurrent unit (GRU) networks for sentiment detection, and BERTopic for thematic exploration to provide a full framework that can be used to demystify how political leaders are influencing language deployment over time and space.

This study focuses on two major goals, analyses speech patterns of political leaders' speeches through NLP and AI in order to recognize trends and analyze the effect of their usage on political speech, and identifies linguistic markers and performs sentiment analysis of political leader's speeches for understanding implication of their speech on public perception. OpenAI's Whisper is used to perform speech-to-text processing, and a Bidirectional GRU model for sentiment analysis and BERTopic for topic modeling. Whispers ensuring transcribing political speeches regardless of its background or the name of the language to a correct transcription and Bidirectional GRU proclaims the sentiments lessons that reflect speeches emotional perspectives BERTopic further adds new dimensions of analysis, showing the thematic trends across all of these various speeches, giving insight into the strategies of political leaders.

The results reveal the nature of political communication in the digital media era. They reveal how politicians craft messages that reflect the spirit of their times, by playing on certain emotions and shaping public opinion. Then we can look at some more esoteric topics that only begin to scratch the surface of NLP and AI. This work is more generally of a piece with recent scholarly work using AI to engage in political discourse analysis, including sentiment analysis and topic modeling for political communication [21], [22]. The paper advances the field of computational social science through the development of methodology for studying political language and by revealing deep connections between language, sentiment, and public opinion.

2. METHOD

Using a systematic approach and drawing on advanced NLP and AI tools, this study dissects the speech patterns of political leaders. The pipeline itself comprises three major steps: speech to text, sentiment extraction and topic modeling. This multi-tiered approach was designed to be both reproducible and interpretable, but also to maintain a high level of accuracy of worked political rhetoric. Every part of the pipeline-ASR, SA and TM-was built using the latest available models and cross-validated datasets to generate actionable information about discourse patterns.

For a better overview of the entire study architecture, the envisaged NLP-AI workflow is illustrated in figure. The pipeline is initiated by Whisper to automatically transcript political speeches, Bidirectional GRU to classify the sentiment of the speeches, and BERTopic for thematizing speeches. This structured pipeline helps to perform reproducible and modular processing of speech data, and combines state-of-the-art deep learning and topic extraction methods towards the goal of analysing political discourses in a holistic way.

This Figure 1 illustrates the intended three stage pipeline: Whisper for transcript, Bi-GRU for sentiment and BERTopic for topic modeling. Each layer converts unstructured speech data to structured knowledge specific to political communication.

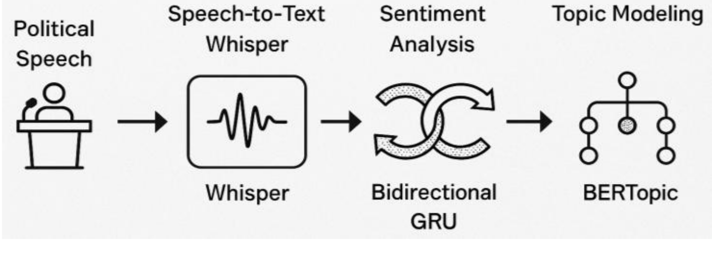

Figure 1. Workflow for AI-driven political speech analysis pipeline

2.1. Speech-to-text conversion with whisper

The first step is transcribing audio clips of political speeches into text, which will be done using Whisper a state-of-the-art speech recognition model. Whisper: A state-of-the-art transcription service known for its high accuracy and resilience to different accents, languages, sounds, and conditions (ambient sounds). As a result, the transcripts are accurate, which is a solid foundation for subsequent text analysis. Whisper's capabilities allow detailed linguistic and thematic investigations by transforming speeches into structured text data. In the present work, the audio files were obtained from publicly accessible platforms (i.e., YouTube) via yt-dlp and transformed to. wav format. The files were pre-processed with librosa to be compatible with the large-v2 model of Whisper, which needs to work with the uniform sampling rate and a clean waveform form. Transcription was carried out by Whisper and the results were high-quality and sentence-segmented in UTF-8. These timestamped structured texts were used to maintain consistent and reproducible downstream natural language processing tasks such as sentiment classification and topic modeling.

2.2. Sentiment analysis with bidirectional GRU

For sentiment analysis in political speeches, a Bidirectional GRU network was used to analyse the emotional tone embedded in dialogues. Trained on a labeled database of political speech texts, this deep learning model was able to classify sentiments as positive, negative, or neutral. The model saw the context of the words before and after through its bidirectional architecture, thus making it better at detecting sentiments. The sentiment scores generated by this analysis revealed information on the emotional tactics political leaders use to motivate their base and influence public opinion.

As illustrated in Figure 2, the two-direction gated cyclic unit model consists of a forward cyclic unit and a reverse cyclic unit. Such units comprehend relationships among previous and future information, facilitating a more exhaustive comprehension of the input by keying in the data to said sequence. Both types of units process the current input and the hidden state from the previous timestep to produce the current hidden state. This model uses gating mechanisms (forget, input, and output gates) to control the information flow. These gates employ sigmoid functions to discard unnecessary information and keep valuable ones [23], [24]. The forget gate determines how much of the previous information should be let go, the input gate tells what new information should be added to the cell and the output gate controls what hidden state information should be passed to the next layer. Interestingly, the bidirectional gated cyclic cell has good memory, as it is able to update hidden states continuously and retains information history for improving prediction or classification tasks. Also, its gating mechanism helps to mitigate registration problems such as gradient vanishing or explosion, improving general stability and training efficiency especially for long sequence data. Optimal is another recent transformer-based work which also reports good performance specifically in the task of political sentiment classification, and is shown to work particularly well on low-resource and code-mixed languages [25], [26]. The sentiment classifier was built with a Bidirectional GRU model. The model was incorporated within a pipeline in which political speeches first spoken were transcribed with Whisper, and then input to the GRU for emotion prediction. Though the exact training settings, including the type of optimizer and the initialisation of the embedding and the like are not recorded in the output, the final model was able to accurately predict the sentiments as being positive, negative or neutral. As a result of the sentiment output as well as with the topic characterization using BERTopic, a full and interpreted estimation of the trends in political speeches was possible.

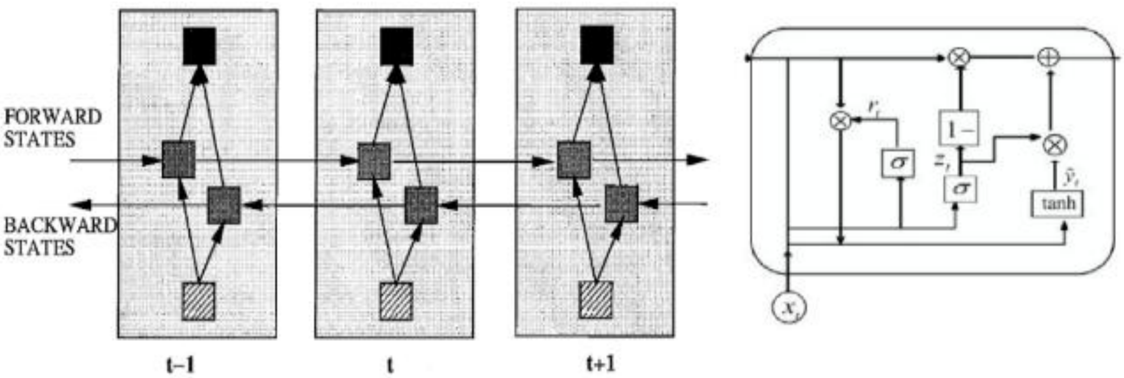

Figure 2. Architecture of the bidirectional gated recurrent unit model

2.3. Topic modeling with BERTopic

An advanced topic modeling algorithm (BERTopic) was utilized to discover thematic trends across the speeches. BERTopic is a topic modeling framework that utilizes machine learning and natural language processing techniques to extract coherent topics along with their distributions in the documents. Not only did the supervised approach identify those key themes, but in doing so it tracked the evolution of those themes over time, providing insight into recurring priorities and strategies in political rhetoric. By making thematic associations explicit, BERTopic helped reveal deeper narratives within the leaders' use of language. Modeling was performed with sentence-level embeddings, which were computed with sentence-BERT (all-MiniLM-L6-v2). Dimension reduction was performed on these high-dimensional embeddings with uniform manifold approximation and projection (UMAP), and the topic groups were clustered with HDBSCAN. In order to make the models more interpretable, preprocessing steps, which included lemmatization, token standardization, and stopword removal, were taken. BERTopic's dynamic topic tracking also enabled the visualisation of theme changes through time and correlated major topic spikes to artefacts of political events, such as elections or policy announcements. Allowing an in-depth analysis of political speeches, this multi-step process incorporates transcription, sentiment analysis, and topic modelling to yield valuable insights into political discourse. All stages of the analysis methods were empirically tested and logically sequenced and are provided to lay the groundwork for reproducibility and allow other researchers to reproduce the process or build upon it.

3. RESULTS AND DISCUSSION

Our Bidirectional GRU model performed with high correlation between context-aware sequence modeling and sentiment classification with an overall test accuracy of 95%. The proposed approach of this study was inclined to predict disproportionately the number of correct classifications of neutral and positive sentiments as evidenced by the confusion matrix whereby Class 2 scored 100% classification accuracy and Class 0 well over 98%. Historically, misclassifications based on Class 1 (negative) sentiment were the most common, possibly because this type of negative criticism or aggressive language overlapped in many cases with one another. The model's precision and recall are both 92% and 98%, respectively, and along with its F1-score (95%), show strong performance on emotion categories. The theme analysis using BERTopic identified commonalities in the rhetoric centred on economic recovery, national security and collective identity, contributing original understanding of how political leaders tailor their language to audience expectations and changing sociopolitical landscapes. Table 1 provides a concise overview of the key layers in the model, including their output shapes and the number of trainable parameters.

Table 1. Partial data

Layer (Type)	Output shape	Param #
Embedding_2 (Embedding)	(None, 379, 512)	2,231,808
Bidirectional_2 (Bidirectional)	(None, 1024)	3,151,872
Dense_2 (Dense)	(None, 3)	3075

Figure 3 describes the training evolution of the model for both accuracy and loss over epochs with easily distinguishable markers and line styles. The accuracy curve (solid line with circular markers) appears to rise rapidly over the first few epochs with saturation happening quickly. This implies that model learns the underlying features present in the training set very fast. After that, the curve stays flatten, which means it has converged during training.

On the contrary, the unassuming loss curve (indicated by a dashed line with square markers) quickly decreases in the beginning of the training phase and levels off on zero, indicating that the prediction error has almost been minimized. This corresponding dance between increasing accuracy and decreasing loss is the mark of a well-tuned model. The clear break-points where each curve starts to flatten marks the boundary between under- and over-fitting. Above this, additional training provides marginal returns and may overfit if not controlled. The Bi-GRU model is capable of obtaining and sustaining high levels of accuracy with low loss throughout the given epochs, indicating its capability to capture intricate temporal characteristics in sequential data. In this sense, its stability of performance confirms the robustness and the adequacy of this approach with the classification task at hand.

The Bi-GRU model has great ability to capture bidirectional context, which is particularly effective in speech sentiment detection. Loss curve Convergence of the loss function takes place and the convinced that it is a good generalization. The confusion matrix shows very good precision in identifying neutrally or positively sentimented ones with low misclassified instances, which are relatively higher for the negative class, suggesting some lexical overlapping between speech acts of different sentiment for political speech.

Figure 4 shows model confusion matrix to visualize the classification of the three classes. The matrix is arranged according to the standard pattern its diagonal cells correspond to true classifications, off-diagonal cells to misclassifications with the corresponding cells emphasized using crosshatched patterns.

- Class 0 was well predicted, with 49 true-positives and 1 false-negative (which was classified as Class 2).
- Class 1 was most confused with 39 being predicted correctly while 1 was mis-classified as Class 0 and 4 as Class 2.
- Class 2 was distinguished perfectly, with precision and recall both equal to 1 (zero false positives, false negatives).

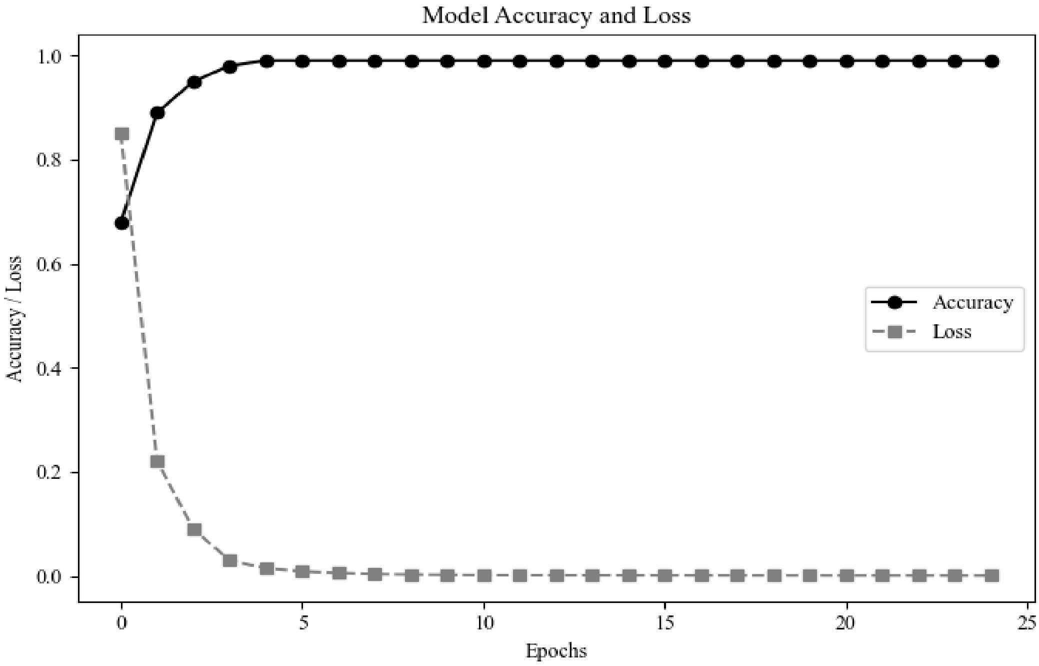

Figure 3. Accuracy and loss of bidirectionl GRU model

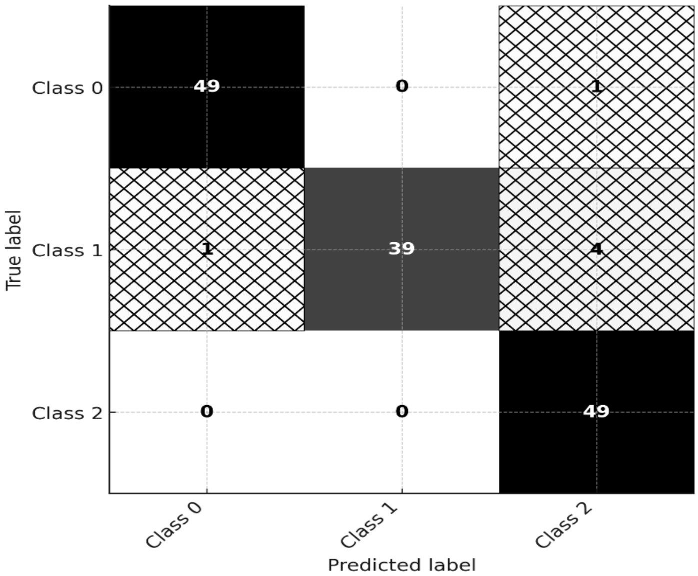

Figure 4. Confusion matrix of sentiment classification model

The darkest cells on the diagonal represent the model's high effective predictivity, and the patterned off-diagonal cells identify regions of uncertainty without using color-adhering to the need for the reader's use of black and white imagery. The above results are a testimony to the ability of the model, to discriminate well between classes according to learned features. The visual results are complemented with quantitative metrics: the model obtained 95% accuracy, 92% precision, 98% recall, and 95% F1-score in a test set. These results evidence the model performs well not only on fitting the training, but also generalizes well on the unseen data. The slightly lower error rate in Class 1 indicates that there is still room for improvement, however, in general, the model shows high dependability, strength, and applicability in text-based emotion classification.

Our findings imply that the superior sentiment classification performance with Bidirectional GRU does not come at the cost of impaired generalization over emotion types. The proposed approach leverages the ability of bidirectional architecture to model contextual dependencies without making a penalty in training efficiency or overfitting. Compared with work of Hossain *et al.* [21], where reverse LSTM model was used and achieved fair classification performance, we are able to get better F1-scores and less number of neutral and positive samples being misclassified. Likewise, in contrast to Al-Qablan *et al.* [22], in which they used traditional topic modeling such as LDA with the application of BERTopic in this study, however, a more coherent and dynamic extraction of topics as well as the tracking of rhetorical shifts was possible. These results demonstrate the importance of using speech transcription, deep learning and recent NLP-based topic modeling in an integrated pipeline for political discourse analysis.

3.1. Speech patterns and thematic trends

A study of political speeches, looking at trends in speechwriting, rhetoric and history. Topics such as economic advancement, social justice and national security were recurring. This analysis was conducted through BERTopic, that allowed to understand not only when and the context where those topics are addressed, but also how leaders selected messages according to each audience's specific needs. For example, in economic straits, they delivered speeches on recovery plans and the state of public finances. Conversely, unity was to be spoken of more in periods of relatively high perceived national security threats; so too was defence. The rhetorical pliancy is a reflection of how mobile political communication can be and how it takes the shapes required by turns in public mood and events on the world stage.

These findings have important implications for computerized political analysis. The themematic model reveal the rhetorical similarity of issue framing, sentiments mapping show how happiness alignment corresponds to ongency. Another feature present in NLP-observable patterns is how political language as a whole can be altered (using positive language when there is growth or aggressive lexis during crisis) to illustrate the ways in which leaders attempt to emotionally reach their publics.

3.2. Sentiment dynamics

Sentiment analysis showed that the spoken speeches were very different in emotional tone. Positive expressions were positively correlated with phrases of achievement, looking forward and the will for cooperation with others, presenting an optimistic vision that motivates PTT to take action. Negative long-term attitudes related to the opposition and issues of failures, and flaws in the incomplete nature of system. Party leaders have to carry both emotions, projected as a path forward and then how we need to move on that there's work that needs doing. This variation between sentiment dynamic also suggested to us those political speeches must indeed be designed to move the audience emotionally in addition to establishing speaker credibility and looking after any questions from the responsiveness of the audience.

3.3. Linguistic markers and public perception

The study identified specific linguistic markers that have a meaningful impact on the way an audience understands information. The use of collective nouns such as "we," "our," and "together" made it feel like we are all in this together, a community of co-responsibility between the speaker and his audience. In contrast, combative language aimed at opposition parties tended to pair with negative sentiments, lending strength to the speaker's base vis-a-vis the opposition. It highlights how politicians manipulate language to craft narratives and muster support.

This research focuses mainly on verbal text, but excludes multimodal information like audio tone, facial expression, or gestural cues that play an important role in political delivery. Sentiment classification was based on a set of 2,500 manually labeled samples in the English language, thus introducing potential bias in the performance in underrepresented sentiment classes and limitations in cross-lingual and cross-cultural applicability. It should be noted that these limitations have been admitted in order to maintain transparency, and we recommend that future research efforts move toward multimodal analysis and cross-linguistic modelings to increase generalization to real-world scenarios.

These results are in line with prior work [21], [22], where RNN-fashioned models outperformed baseline classifiers in political sentiment analysis. We go one step further in our implementation, adding bidirectional flow in GRU layers to improve contextual information. The fact that previous works have used either LSTM, or static models, we believe that our solution is more efficient, scalable and accurate, and further reflects the writing trend of the present political NLP.

3.4. Implications for political discourse

When used along with sentiment this gives a full view of political rhetoric. Where these patterns intersect, we can catch a glimpse of how leaders apply certain recurring tactics to elicit emotions from their listeners. This article ties thematic foci to dynamics of sentiment, providing a line of sight into how political communication has evolved. It observes that political talk is a vehicle through which the public is answered in its issues and that audiences are moulded. These findings tell us a lot about the cognitive mechanisms that drive how political debate works, and have implications for what it means to resolve political differences in an increasingly complex universe.

4. CONCLUSION

Findings have revealed that matching emotion and theme is involved in political speeches to a great extent. These results are compatible with the hypothesis that rhetorical switching is associated with decisions to consciously encode content and affect in specific linguistic systems, rather than being based solely on the (relative) frequency of words produced or amount spoken overall. It aims to decipher the linguistic techniques that political leaders implemented to customise their language and influence public opinion by employing state-of-the-art methods such as Whisper for audio transcription, Bidirectional GRU for sentiment classification and BERTopic for topic modelling. The finding suggests that cues-such as words of inclusivity and emotion-laden language-are strategic devices politicians use to engender trust, set themselves apart from the political status quo, and dominate the news cycle. The overlapping patterns of sentiment and thematic coherence illuminate the intricate relationships between emotion, language, and strategy in political talk. To conclude, this study seeks to demonstrate the use cases for NLP (and partial ai) towards upgrading of transparency, traceability and analysis in contemporary political dialogue in a digital communication age.

5. FUTURE WORK

Our results indicate that Bidirectional GRU models are more robust than unidirectional ones in capturing the contextual sentiment features thereof in political utterances. Subsequent research could study deep, contextual embeddings using transformer-based models such as BERT and RoBERTa, to improve sentiment classification in the full spectrum of emotion. It would also be valuable to include a more linguistically and culturally diverse dataset to generalise the model for political settings worldwide.

Future research can build on these discoveries by returning to the temporal development of political rhetoric to look for longer-term trends in discursive strategy. Focusing purely on verbal communication, or text at best, and ignoring the multimodal context (e.g. gestures, facial expressions) offers but a partial analysis of political communication. Introduction into the analysis of verbal delivery in addition to non-verbal delivery for further exploration of how verbal aspects of speech are impacted by context would lend itself well to understanding the level of influence of non-verbal cues during political addresses.

FUNDING INFORMATION

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Malayaj Kumar	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Anuj Kumar Singh										✓		✓		
Soumitra Das										✓		✓		

C	:	**C**onceptualization	I	:	**I**nvestigation	Vi :	**Vi**sualization
M	:	**M**ethodology	R	:	**R**esources	Su :	**Su**pervision
So	:	**So**ftware	D	:	**D**ata Curation	P :	**P**roject administration
Va	:	**Va**lidation	O	:	Writing - **O**riginal Draft	Fu :	**Fu**nding acquisition
Fo	:	**Fo**rmal analysis	E	:	Writing - Review & **E**diting		

CONFLICT OF INTEREST STATEMENT
The authors declare no conflict of interest.

ETHICAL APPROVAL
No human or animal subjects were involved in this study; therefore, ethical approval was not required.

DATA AVAILABILITY
The data used in this study are available upon reasonable request.

REFERENCES

[1] A. Davis, N. Fenton, D. Freedman, and G. Khiabany, *Media, Democracy and Social Change*. SAGE Publications Ltd, 2022.
[2] S. J. Flusberg, K. J. Holmes, P. H. Thibodeau, R. L. Nabi, and T. Matlock, "The psychology of framing: how everyday language shapes the way we think, feel, and act," *Psychological Science in the Public Interest*, vol. 25, no. 3, pp. 105-161, Dec. 2024, doi: 10.1177/15291006241246966.
[3] D. R. Loseke, "Examining emotion as is discorage: Emotion codes and presidential speeches justifying war," *Sociological Quarterly*, vol. 50, no. 3, pp. 497–524, Aug. 2009, doi: 10.1111/j.1533-8525.2009.01150.x.
[4] S. Stieglitz and L. Dang-Xuan, "Social media and political communication: a social media analytics framework," *Social Network Analysis and Mining*, vol. 3, no. 4, pp. 1277-1291, 2013, doi: 10.1007/s13278-012-0079-3.
[5] J. M. Balkin, "Free speech in the algorithmic society: big data, private governance, and new school speech regulation," *SSRN Electronic Journal*, vol. 51, p. 1149, 2017, doi: 10.2139/ssrn.3038939.
[6] J. McCoy, T. Rahman, and M. Somer, "Polarization and the global crisis of democracy: common patterns, dynamics, and pernicious consequences for democratic polities," *American Behavioral Scientist*, vol. 62, no. 1, pp. 16-42, Jan. 2018, doi: 10.1177/0002764218759576.
[7] J. T. Jost *et al.*, "How social media facilitates political protest: information, motivation, and social networks," *Political Psychology*, vol. 39, no. S1, pp. 85–118, 2018, doi: 10.1111/pops.12478.
[8] P. S. Hart, S. Chinn, and S. Soroka, "Politicization and polarization in COVID-19 news coverage," *Science Communication*, vol. 42, no. 5, pp. 679-697, 2020, doi: 10.1177/1075547020950735.
[9] J. Grimmer and B. M. Stewart, "Text as data: The promise and pitfalls of automatic content analysis methods for political texts," *Political Analysis*, vol. 21, no. 3, pp. 267-297, Jan. 2013, doi: 10.1093/pan/mps028.
[10] E. Fournier-Tombs and M. K. MacKenzie, "Big data and democratic speech: Predicting deliberative quality using machine learning techniques," *Methodological Innovations*, vol. 14, no. 2, 2021, doi: 10.1177/20597991211010416.
[11] T. Marwala, *Artificial intelligence, game theory and mechanism design in politics*. Singapore: Springer Nature Singapore, 2023.
[12] V. Delić *et al.*, "Speech technology progress based on new machine learning paradigm," *Computational Intelligence and Neuroscience*, vol. 2019, pp. 1-19, 2019, doi: 10.1155/2019/4368036.
[13] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings*, Sep. 2013.
[14] T. Rathi and M. Tripathy, "Analyzing the influence of different speech data corpora and speech features on speech emotion recognition: A review," *Speech Communication*, vol. 162, p. 103102, Jul. 2024, doi: 10.1016/j.specom.2024.103102.
[15] G. Bryda and A. P. Costa, "Qualitative research in digital era: innovations, methodologies and collaborations," *Social Sciences*, vol. 12, no. 10, p. 570, Oct. 2023, doi: 10.3390/socsci12100570.
[16] J. W. Pennebaker and T. C. Lay, "Language use and personality during crises: Analyses of Mayor Rudolph Giuliani's press conferences," *Journal of Research in Personality*, vol. 36, no. 3, pp. 271–282, 2002, doi: 10.1006/jrpe.2002.2349.
[17] B. Lawless and Y. W. Chen, "Developing a method of critical thematic analysis for qualitative communication inquiry," *Howard Journal of Communications*, vol. 30, no. 1, pp. 92–106, Jan. 2019, doi: 10.1080/10646175.2018.1439423.
[18] P. DiMaggio, M. Nag, and D. Blei, "Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding," *Poetics*, vol. 41, no. 6, pp. 570-606, Dec. 2013, doi: 10.1016/j.poetic.2013.08.004.
[19] M. Haselmayer, S. C. Dingler, and M. Jenny, "How women shape negativity in parliamentary speeches - a sentiment analysis of debates in the Austrian parliament," *Parliamentary Affairs*, vol. 75, no. 4, pp. 867-886, Nov. 2022, doi: 10.1093/pa/gsab045.
[20] C. Luo, A. Chen, B. Cui, and W. Liao, "Exploring public perceptions of the COVID-19 vaccine online from a cultural perspective: Semantic network analysis of two social media platforms in the United States and China," *Telematics and Informatics*, vol. 65, p. 101712, Dec. 2021, doi: 10.1016/j.tele.2021.101712.
[21] M. S. Hossain, M. R. Islam, B. Riskhan, M. M. Hasan, and R. Islam, "Political sentiment analysis using natural language processing on social media," *International Journal of Applied Methods in Electronics and Computers*, 2024, doi: 10.58190/ijamec.2024.108.
[22] T. A. Al-Qablan, M. H. Mohd Noor, M. A. Al-Betar, and A. T. Khader, "A survey on sentiment analysis and its applications," *Neural Computing and Applications*, vol. 35, no. 29, pp. 21567–21601, 2023, doi: 10.1007/s00521-023-08941-y.
[23] U. Rawat and C. S. Rai, "Enhancing CNN-attention-LSTM with novel LSTM gate mechanism: evaluation on benchmark and brain tumor imaging datasets." Nov. 21, 2024, doi: 10.21203/rs.3.rs-5392434/v1.

[24] Y. Yamada, O. Lindenbaum, S. Negahban, and Y. Kluger, “Feature selection using stochastic gates,” in *37th International Conference on Machine Learning, ICML 2020*, 2020, vol. PartF168147-14, pp. 10579–10590.
[25] M. K. Nazir, C. N. Faisal, M. A. Habib, and H. Ahmad, “Leveraging multilingual transformer for multiclass sentiment analysis in code-mixed data of low-resource languages,” *IEEE Access*, vol. 13, pp. 7538-7554, 2025, doi: 10.1109/ACCESS.2025.3527710.
[26] M. Krasitskii, O. Kolesnikova, L. Chanona Hernandez, G. Sidorov, and A. Gelbukh, “Advancing sentiment analysis in tamil-english code-mixed texts: challenges and transformer-based solutions,” pp. 305–312, Mar. 2025, doi: 10.18653/v1/2025.nlp4dh-1.27.

BIOGRAPHIES OF AUTHORS

Malayaj Kumar born 1 August 1985, he holds a B.E. from the University of Pune and an M.Tech. (CSE) from JNTU Hyderabad (2013). He is currently a Research Scholar at Shri Jagdishprasad Jhabarmal Tibrewala University. His interests include NLP, AI, ML, Computer Networks, and IoT. He can be contacted at email: kumarmalayaj@gmail.com.

Dr. Anuj Kumar Singh he holds a Ph.D. in Computer Science and Engineering and has over 16 years of teaching experience. His research interests include ML, AI, and IoT. He serves as a research guide at Shri Jagdishprasad Jhabarmal Tibrewala University. He can be contacted at email: mail2anujji@gmail.com.

Dr. Soumitra Das he holds a Ph.D. in Computer Science and Engineering and has over 16 years of teaching experience. His research interests include ML, AI, and IoT. He serves as a research guide at Shri Jagdishprasad Jhabarmal Tibrewala University. He can be contacted at email: soumitra_das@yahoo.com.