

XELSA: Sentiment Analysis Using Transformer Model

UMAMAHESWARI GURUSAMY¹, SANGEETHA VIJAYARAJAN²,
MEENAKSHI ANANTHARAMAN¹, VASUDEVAN KARUPPIAH³

¹Department of Computer Science and Engineering, Kamaraj College of Engineering and Technology, Virudhunagar, Tamil Nadu, India

²Department of Artificial Intelligence and Data Science, SRM Easwari Engineering College Tamil Nadu, Chennai, India

³Department of Electronics and Communication Engineering, Thiagarajar College of Engineering, Madurai, Tamil Nadu, India

Corresponding author: Umamaheswari Gurusamy (e-mail: uma.optimist@gmail.com).

ABSTRACT Sentiment analysis is an invaluable tool in modern election campaigns, which provides profound insights into voter preferences, behavior, and opinions, enabling campaigns to make informed decisions and ultimately increase their chances of success in an election. Many existing approaches use transformer models like BERT, which uses a masked language model technique, to predict these masked tokens by substituting a [MASK] token for certain input tokens. A mismatch between pre-training and fine-tuning may arise from this method since BERT does not view the actual sequence when training. The proposed XELSA model leverages the advanced capabilities of XLNet's permutation language modeling to achieve superior sentiment analysis performance compared to traditional models like BERT. The proposed work shows that XLNet can model longer-term dependencies and capture more context compared to BERT, often leading to improved performance in understanding complex sentiment expressions. The results show that XELSA shows a better classification of negative, positive, and neutral tweets and outperforms BERT in terms of providing an accuracy of 96%. A dataset of 10,000 tweets pertaining to the US 2020 election was used in the study, and XELSA demonstrated improved precision, recall, and F1 score across all sentiment categories. These results highlight the effectiveness of XELSA in providing more accurate sentiment analysis, which is crucial for election forecasting.

KEYWORDS XELSA; BERT; Election campaigns; Permutation language modeling; Accuracy.

I. INTRODUCTION

Political forecasting uses analytical and statistical methods to predict election outcomes, mainly relying on surveys and qualitative approaches [1]. It also involves analyzing political party manifestos and monitoring trends in popular news media, known as political analysis [2, 3]. Predicting election results has become more difficult with increased government opposition, particularly in countries like the USA, where two dominant parties compete [4]. One significant challenge is achieving a balanced representation of opposing political viewpoints in data collection [5]. There are many challenges in performing survey sampling on platforms like Twitter and Facebook. During elections, these social networks play a crucial role in political campaigns and activism [6]. Social media has been a platform for predicting the election outcomes over the past ten years, primarily through artificial intelligence and natural language processing (NLP) methods. These methods range from basic statistical techniques to complex deep learning models [7]. Election modeling strategies include topic modeling and sentiment analysis [8].

A. COMPARISON OF BERT AND XLNET

BERT's masked language modeling (MLM) has indeed demonstrated exceptional performance across numerous fine-tuning tasks, including sentiment analysis, named entity recognition, and question answering. The MLM approach allows BERT to capture bidirectional context by predicting masked tokens based on their surrounding context, which has been a significant factor in its success. However, the MLM approach also has certain limitations. During pre-training, BERT does not see the full sequence as it masks a portion of the tokens, which can lead to a mismatch between pre-training and fine-tuning phases [9].

XLNet is a transformer language model that combines the advantages of an autoregressive language model, which predicts the next word using the context words, and an autoencoder model, which has the ability to utilize context from both forward and backward directions. This makes XLNet suitable for applications such as NLP and sentiment analysis.

Table 1 provides a comprehensive comparison of the mechanisms and performance of BERT and XLNET in sentiment analysis.

Table 1. Comparison of the mechanisms and performance of BERT and XLNET

BERT	XLNET
BERT is a bidirectional transformer model utilizing the transformer framework to understand a word's context from both directions (left and right). Its training involves two stages: pre-training and fine-tuning.	XLNet is a versatile autoregressive pretraining method that merges principles from both autoregressive (AR) and autoencoding (AE) techniques. It utilizes a permutation-based training strategy, which allows the model to learn bidirectional context without the use of masking.
In Masked Language Model, only some tokens in the input are used to predict the vocabulary ID of the context. Next Sentence Prediction (NSP): Predicts if a pair of sentences appears sequentially in the original text, enabling the model to understand sentence relationships.	In Permutation Language Modeling, next token in a sequence is predicted based on the preceding tokens, which enhances the model's ability to capture token dependencies. The longer-term dependencies are captured by employing a segment-level recurrence mechanism.
A classification layer is appended to the pre-trained BERT model to achieve Fine-tuning for doing a particular task, such as sentiment analysis.	XLNet model also employs a classification layer after the pre training to do sentiment analysis.
In sentiment analysis, BERT can capture nuanced meanings and context because of its bidirectional nature, making it particularly effective in understanding the sentiment expressed in text.	XLNet can model longer-term dependencies and capture more context compared to BERT, often leading to improved performance in understanding complex sentiment expressions.

II. RELATED WORK

The detailed literature survey shows various works based on BERT model in conducting sentiment analysis across various domains.

A novel approach was proposed in [10] that combines CNN with a BERT-GCN model for abstract-level sentiment analysis. The proposed work used a combination of BERT and BiLSTM models to extract the hidden information. The BERT model, which is a deep learning model, utilizes convolutional layers from a CNN model to extract hidden information. Initially, we transform the words in the sentences into vectors, which aids in identifying significant word representations. Subsequently, we extract the significant features. Finally, a CNN model classifies aspect-level sentiments based on these feature vectors. This model does better than other context-based GCN methods for aspect-based sentiment analysis (ALSA) when tested on three benchmark datasets.

A study [11] was conducted to evaluate the performance of various BERT-based models for aspect term extraction in Arabic text from educational institution surveys. They also developed a system that automated the ABSA process, allowing for automatic labeling of survey responses. The proposed model aimed to collect aspect terms, map these aspects to relevant categories, determine their polarity, and categorize the polarity accordingly. This system can help Arabic educational institutions gain valuable insights from open-ended survey responses, aiding decision-makers in more effectively allocating resources, improving facilities and services, and enhancing students' learning experiences.

A model was suggested in [12] for analyzing user behavior, particularly the sentiments expressed during elections, which

can provide insights into potential election outcomes. We designed a framework to model the US general elections using deep learning models, specifically utilizing long short-term memory (LSTM) models and bidirectional encoder representations from transformers (BERT). The authors conducted a study using LSTM and BERT models to analyze Twitter sentiment regarding the US 2020 presidential elections. The results indicate that the BERT model accurately predicted Biden's victory.

A cross-domain text sentiment classification using a BERT-based aspect-level sentiment analysis algorithm, which enables fine-grained sentiment analysis across various domains was proposed in [13]. They used the BERT model to identify sentiment features at both the aspect and sentence levels and created a pair of sequences for each sentence and classified them based on the sentiment labels. They conducted the experiments using the Amazon product review dataset to evaluate the accuracy and F1 score, demonstrating the algorithm's superiority over traditional methods.

A transformer-based fine-tuned model for emotion and sentiment analysis of Hungarian political texts was developed in [14]. The authors manually annotated the parliamentary speech texts from 2014 to 2018, specifying a wide range of emotions for each text. This dataset is publicly accessible for research purposes. They built a BERT model for Hungarian texts to classify sentiments and emotions. They fine-tuned the model to improve its performance and reduce classification errors. These errors primarily stem from the model's challenges in identifying pragmatic elements and other nuances of language use.

The authors in [15] proposed the sentiment analysis knowledge graph (SAKG)-BERT model, which integrates sentiment analysis knowledge with the BERT language representation model. The proposed model employs knowledge graphs to provide a good semantic classification that improves the representation of the language.

The evolution of transformer-based models for sentiment analysis has seen rapid advancement since BERT's introduction [15]. RoBERTa (Robustly Optimized BERT Approach) [16] optimized BERT's pretraining regimen through three key modifications: removing the next-sentence prediction objective, implementing dynamic masking patterns that change during training, and scaling up training data and batch sizes. These optimizations addressed BERT's undertraining issues and resulted in improved performance across multiple NLP benchmarks, including sentiment analysis tasks where contextual understanding is critical.

DeBERTa (Decoding-enhanced BERT with Disentangled Attention) [17] introduced two architectural innovations to overcome BERT's limitations. First, its disentangled attention mechanism represents each word using separate content and position embeddings, allowing the model to capture contextual relationships more effectively. Second, an enhanced mask decoder improves pretraining efficiency by better utilizing the masked language modeling objective. These innovations enable DeBERTa to achieve state-of-the-art performance on various natural language understanding tasks, including the SuperGLUE benchmark.

XLNet [18], which forms the foundation of our XELSA model, takes a fundamentally different approach by employing permutation language modeling. Unlike BERT's masked

language modeling, which assumes token independence during pretraining, XLNet captures bidirectional contexts through autoregressive modeling of all possible token permutations. This approach preserves the dependency relationships between tokens while maintaining the ability to learn from both left and right contexts. For sentiment analysis of political tweets, this architectural advantage proves particularly valuable when interpreting emotionally charged language and detecting sarcastic expressions, where contextual relationships between distant tokens are essential for accurate classification.

While transformer-based models like BERT have been applied to sentiment analysis in election campaigns [12], no existing work has yet explored the potential of XLNet for this domain, despite its theoretical advantages in capturing long-range dependencies and bidirectional context.

The contributions in the proposed work are as follows:

- Modeling XELSA using XLNet-base-based transformer model for performing Sentiment analysis.
- Training the model with the Trump dataset by specifying parameters for the training process including number of epochs, batch size, learning rate, and regularization methods.
- Evaluating the model across various batch sizes such as 32, 64 and 128 and comparing the performance with existing BERT.

III. THE PROPOSED MODEL XELSA

The XLNET language models consist of two phases: the pre-train phase and the fine-tune phase. XLNet primarily focuses on the pre-training phase, utilizing permutation language modeling to achieve this. If a sequence consists of n tokens, then the number of combinations of tokens in the sequence is $n!$. Let us assume that the sequence contains the tokens $[w_1, w_2, w_3, w_4]$. The possible positions of the token are $[x, w_2, x, x]$, $[w_2, x, x, x]$, $[x, x, w_2, x]$, $[x, x, x, w_2]$.

Likewise, the sequence $[w_1, w_2, w_3, w_4]$ consists of 24 possible tokens. If the sequence contains t tokens, then each word possesses $(t-1)$ tokens or possible locations. XLNet takes the best parts of autoregressive models, which guess tokens in a certain order, and autoencoding models, which look at context in both directions, to better understand how tokens are related to each other.

The libraries like pandas for data manipulation and analysis, PyTorch for tensor operations and deep learning model construction, the Hugging Face library for diverse transformer models, and sklearn for machine learning tools are utilized. These tools include KFold for cross-validation, classification_report for evaluation, and nlpaug for data augmentation, particularly synonym augmentation.

A. LOADING AND PREPROCESSING DATA

Pandas loads the dataset and performs preprocessing to prepare it for model training. Cross-validation divides the dataset into training and validation sets. This class transforms the tokenized inputs and labels into a format that PyTorch DataLoader can utilize.

B. TOKENIZATION AND AUGMENTATION

Tokenization is a crucial step in which the model converts the text data into tokens. The data augmentation is performed after tokenization to improve the model's generalization and

broaden the diversity of the training data. The tokenizer has been initialized to tokenize the input text using the XLNet tokenizer.

C. MODEL INITIALIZATION WITH DROPOUT

A custom model class is defined to include dropout layers. Dropout is a regularization technique to prevent overfitting by randomly dropping units during training. KFold technique divides the data into training and validation sets for the purpose of cross-validation.

D. TRAINING

Training arguments are defined to set the parameters for the training process, including learning rate, batch size, number of epochs, and regularization techniques like weight decay and gradient clipping. The parameters used in the training phase is shown in Table 2.

Table 2. Parameters used in the training of XELSA

Parameter	Value
num_train_epochs	10
per_device_train_batch_size	16
per_device_eval_batch_size	64
logging_steps	10
evaluation_strategy	"steps"
eval_steps	50
save_steps	100
save_total_limit	2
learning_rate	1×10^{-5}
weight_decay	0.01
max_grad_norm	1.0

E. PARAMETER TUNING

The optimization hyperparameters for fine-tuning the XLNet model were selected based on theoretical foundations and empirical best practices established in the literature on transformers. The learning rate was set to $1e-5$, a value deliberately chosen within the conservative range recommended for fine-tuning pre-trained language models [19,20]. This selection is motivated by two primary considerations. First, pre-trained models like XLNet encapsulate substantial linguistic knowledge acquired during large-scale pre-training; excessively high learning rates risk catastrophic forgetting, whereby rapid weight updates overwrite this foundational knowledge. Second, the domain-specific characteristics of the 2020 US election tweets including political jargon, hashtags, and sarcastic expressions necessitate gradual adaptation to prevent overfitting to superficial patterns or noise in the training data. The chosen learning rate of $1e-5$ enables the model to adapt to the target domain while preserving the general language understanding capabilities essential for accurate sentiment classification.

The weight decay parameter was set to 0.01, serving as an L2 regularization mechanism that penalizes large weight magnitudes in the model architecture [20, 21]. This value, widely adopted in transformer-based models including BERT, RoBERTa, and XLNet, helps mitigate overfitting by encouraging the learning of simpler, more generalizable representations. Regularization is particularly critical for our

task given the inherent noisiness of social media text and the potential for the model to exploit spurious correlations, for instance, associating specific political hashtags with sentiment labels without comprehending underlying context. By constraining weight growth, weight decay promotes superior generalization to unseen tweets and enhances model robustness to linguistic variations in political discourse.

These hyperparameter choices were empirically validated through preliminary experiments comparing learning rates of 5×10^{-6} , 1×10^{-5} , 2×10^{-5} , and 5×10^{-5} . The selected value of 1×10^{-5} yielded the optimal balance between convergence speed and final validation accuracy. Similarly, weight decay values of 0.001, 0.01, and 0.1 were evaluated, with 0.01 providing the most effective regularization without inducing underfitting.

A learning rate of 1×10^{-5} , is used for fine-tuning transformer models. The batch sizes to 16 for training and 64 for evaluation is used. These values are selected to balance computational efficiency and model performance, ensuring effective utilization of the GPU memory without causing overfitting.

F. TRAINING WITH EARLY STOPPING

The model is trained using the Trainer class from the transformer library. Early stopping is implemented to prevent overfitting.

G. EVALUATION AND REPORTING

After training, the model is evaluated on the validation set. A classification report is generated to provide detailed metrics of model performance, including precision, recall, and F1-score.

IV. RESULT ANALYSIS

A. SYSTEM CONFIGURATION

The Google Colab environment with IPython version 7.34.0, that runs on a Colab kernel is used and it provides access to TPU. TPUs are specialized ASICs (Application-Specific Integrated Circuits) designed for accelerating machine learning workloads. Each TPU v2 chip has 8 GB of high-bandwidth memory (HBM). The memory bandwidth is approximately 600 GB/s. The TPU (Tensor Processing Unit) v2 is used as a hardware accelerator for executing the XELSA model.

B. DATA ANALYSIS

The dataset used for analyzing the performance of the XELSA model, consists of tweets related to the US 2020 presidential election. The tweets are collected under the title hashtag_donaldtrump and hashtag_joe Biden. There are 1,53,079 tweets, out of which 92,084 tweets are classified as English tweets. The columns in the dataset are created_at, tweet_id, tweet, likes, retweet_count, source, user_id, user_name, user_screen_name, user_description, user_join_date, user_followers_count, user_location, tweet_creator's profile, lat, long, city, country, state, state_code and collected_at. Tweets related to the US 2020 presidential election provide a rich source of real-time public opinion, making them highly relevant for the work.

The primary goal of the proposed model is to enhance sentiment analysis for election forecasting. The dataset includes various metadata fields such as tweet creation time, user information, and tweet interactions (likes, retweets). This additional information can be leveraged to enhance the

sentiment analysis model and provide deeper insights into voter behavior and preferences.

C. ADDRESSING BIASES IN DATA SELECTION

The hashtags #DonaldTrump and #JoeBiden were chosen to capture sentiments related to both candidates in the US 2020 presidential election. This balanced approach helps to mitigate bias towards any single candidate or viewpoint.

The cross-validation techniques were employed to validate the sentiment labels and ensure that XELSA model's performance is not skewed by biased labeling. This helps to identify and correct any inconsistencies in the labeling.

The techniques like dropout and weight decay were applied during model training to prevent overfitting and ensure that XELSA model does not learn biased patterns from the training data.

D. POLARITY IDENTIFICATION

The study provides a comparison of prediction accuracy between XELSA and BERT. Both BERT (Bidirectional Encoder Representations from Transformers) and XLNet (Generalized Autoregressive Pretraining for Language Understanding) are transformer-based models that have significantly pushed the boundaries of performance in numerous natural language processing (NLP) tasks, including sentiment analysis.

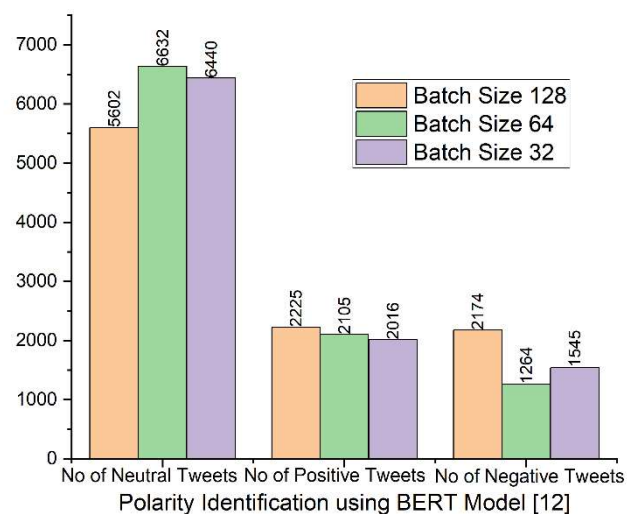


Figure 1. Polarity Identification using BERT Model

The batch sizes of 128, 64, and 32 are used to analyze the performance of the XELSA model. Figure 1 shows the classification of positive, negative, and neutral tweets by using the BERT Model. BERT employs a masked language model (MLM) technique, in which certain tokens in the input are hidden, and the model is trained to predict these masked tokens using the surrounding context. This allows BERT to capture bidirectional context, but the masked tokens cannot see the full sentence context during training. BERT's MLM might cause it to be more conservative in its sentiment analysis, leading to a higher number of neutral classifications. It might be less inclined to detect slight positive or negative sentiments because it learns to be more general in its predictions to account for the masked tokens. While BERT is effective for many NLP tasks, it might struggle with long-range dependencies due to its fixed-

length input during training. Hence, BERT shows a greater number of neutral tweets.

XELSA, with its permutation language modeling, sees all possible contexts by predicting tokens in different orders, leading to a more holistic understanding of the entire sequence. This can make XELSA more sensitive to nuanced sentiments within a tweet. XELSA's training allows it to be more fine-tuned in detecting sentiments because it sees every word in context. This comprehensive view can help it distinguish subtle sentiments better, resulting in fewer neutral classifications. XELSA, with its Transformer-XL backbone, handles longer contexts better, which might help in understanding the full context of a tweet more accurately, leading to more precise sentiment classification. Figure 2 shows the classification of positive, negative, and neutral tweets by using the XELSA Model.

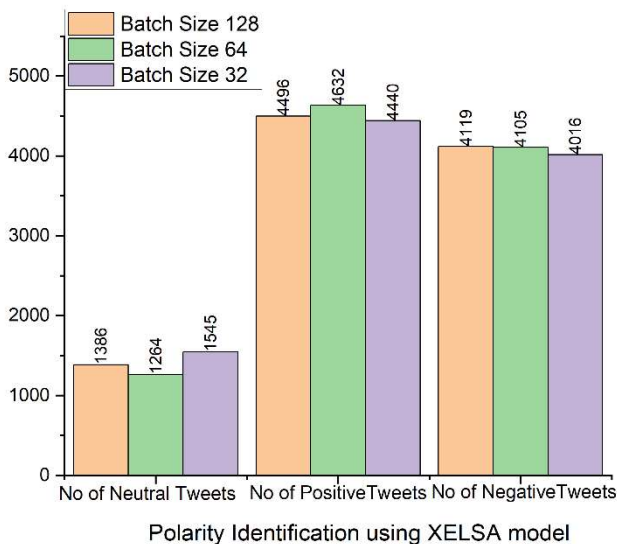


Figure 2. Polarity Identification using XELSA Model

C. ACCURACY

Accuracy refers to the total number of correct predictions divided by the total number of predictions. XELSA integrates the strengths of autoregressive models (which predict the next token) and autoencoding models (which predict missing tokens). By permuting the order of the tokens during training, XELSA can capture dependencies between all tokens, enhancing its understanding of language structure and context.

The Proposed XELSA Model outperforms BERT classifies 4632 tweets as positive and 4105 tweets as negative because of is permutation modelling. XELSA builds upon the Transformer-XL architecture, which uses segment-level recurrence mechanisms and relative positional encoding to model long-term dependencies.

This combination enables XLNet to handle longer contexts more effectively, contributing to its improved accuracy. Permutation modeling allows XELSA to classify better to various tasks by considering different possible contexts during training. This flexibility enables XELSA to perform well on a wide range of natural language understanding tasks, such as

text classification, sentiment analysis, and question answering. Hence it gives an accuracy of 0.96 which is shown in Figure 3.

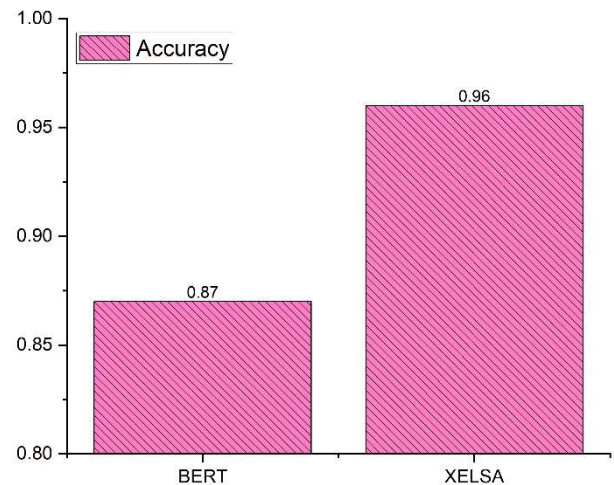


Figure 3. Accuracy of BERT and XELSA Model

D. ERROR ANALYSIS

The results of sentiment analysis show that, XELSA consistently achieves higher accuracy, particularly in scenarios involving complex and nuanced sentiments. This improvement is attributed to XELSA's ability to model longer-term dependencies and capture more context through its permutation language modeling approach. XLNet's architecture, based on Transformer-XL, includes segment recurrence and relative positional encoding, which help capture long-range dependencies more effectively. This makes XELSA particularly well-suited for tasks that require understanding complex and nuanced sentiments over longer contexts.

V. CONCLUSION AND FUTURE WORK

While XELSA demonstrates strong performance on English-language tweets from the 2020 US presidential election, expanding the model to handle multilingual datasets represents a critical avenue for future research. This extension would address the growing global demand for sentiment analysis tools capable of analyzing political discourse across linguistic boundaries. Several considerations motivate this direction:

First, political events increasingly transcend national and linguistic borders, with international audiences engaging in discussions about the US elections in their native languages. A multilingual capability would enable comparative analysis of how non-English speaking populations perceive and discuss American politics, providing richer insights into global political sentiment.

Second, the transformer architecture underlying XELSA is inherently adaptable to multilingual contexts. Models such as multilingual BERT (mBERT) and XLM-RoBERTa have demonstrated that cross-lingual transfer learning is feasible, where knowledge acquired from resource-rich languages like English can be effectively transferred to low-resource languages. Future work could explore fine-tuning XELSA on multilingual corpora or adapting the architecture for zero-shot cross-lingual transfer, wherein the model trained on English data generalizes to other languages without additional training data.

Third, extending XELSA to multilingual settings would

address domain-specific challenges unique to non-English political discourse, including linguistic nuances, cultural references, and varying sentiment expression patterns across languages. Recent work by Chen and Wijaya on multilingual election sentiment analysis demonstrates both the feasibility and complexity of such endeavors, highlighting the need for carefully curated multilingual datasets and language-specific preprocessing pipelines.

Practical implementation would involve several phases: (1) curating or leveraging existing multilingual sentiment datasets for political text; (2) experimenting with cross-lingual transfer learning approaches, including zero-shot, few-shot, and fully supervised multilingual fine-tuning; (3) evaluating performance across languages to identify transfer gaps and language-specific adaptations; and (4) developing language-agnostic preprocessing techniques that preserve sentiment-relevant features across linguistic boundaries. Ultimately, a multilingual version of XELSA would contribute to the broader goal of developing inclusive sentiment analysis tools capable of capturing global political sentiment, thereby enhancing our understanding of how international audiences engage with and react to major political events.

References

- [1] M. S. Lewis-Beck, "Election forecasting: Principles and practice," *Brit. J. Politics Int. Relations*, vol. 7, pp. 145-164, 2005. <https://doi.org/10.1111/j.1467-856X.2005.00178.x>.
- [2] W. Ascher, "Political forecasting: The missing link," *J. Forecasting*, vol. 1, pp. 227-239, 1982. <https://doi.org/10.1002/for.3980010303>.
- [3] K. L. Remmer, "The political economy of elections in Latin America, 1980-1991," *Amer. Political Sci. Rev.*, vol. 87, pp. 393-407, 1993. <https://doi.org/10.2307/2939049>.
- [4] S. Ansolabehere and J. M. Snyder, Jr., "The incumbency advantage in US elections: An analysis of state and federal offices, 1942-2000," *Election Law J.*, vol. 1, pp. 315-338, 2002. <https://doi.org/10.1089/153312902760137578>.
- [5] G. Terhanian, J. Bremer, R. Smith, and R. Thomas, "Correcting data from online surveys for the effects of non-random selection and non-random assignment," *Harris Interactive*, Chicago, IL, USA, White Paper, pp. 1-13, 2000.
- [6] G. Enli, "Twitter as arena for the authentic outsider: Exploring the social media campaigns of Trump and Clinton in the 2016 US presidential election," *Eur. J. Commun.*, vol. 32, pp. 50-61, 2017. <https://doi.org/10.1177/0267323116682802>.
- [7] H. Wang, D. Can, A. Kazemzadeh, F. Bar, and S. Narayanan, "A system for real-time Twitter sentiment analysis of 2012 US presidential election cycle," *Proc. ACL Syst. Demonstrations*, 2012, pp. 115-120.
- [8] A. Giachanou and F. Crestani, "Like it or not: A survey of Twitter sentiment analysis methods," *ACM Comput. Surv.*, vol. 49, pp. 1-41, 2016. <https://doi.org/10.1145/2938640>.
- [9] A. Agarwal, D. Toshniwal, and J. Bedi, "Can Twitter help to predict outcome of 2019 Indian general election: A deep learning based study," *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*, India, Springer, pp. 38-53, 2019. https://doi.org/10.1007/978-3-030-43887-6_4.
- [10] H. T. Phan, N. T. Nguyen and D. Hwang, "Aspect-level sentiment analysis using CNN over BERT-GCN," *IEEE Access*, vol. 10, pp. 110402-110409, 2022. <https://doi.org/10.1109/ACCESS.2022.3214233>.
- [11] K. A. Alshaikh, O. A. Almatrafi and Y. B. Abushark, "BERT-based model for aspect-based sentiment analysis for analyzing Arabic open-ended survey responses: A case study," *IEEE Access*, vol. 12, pp. 2288-2302, 2024. <https://doi.org/10.1109/ACCESS.2023.3348342>.
- [12] R. Chandra and R. Saini, "Biden vs Trump: Modeling US general elections using BERT language model," *IEEE Access*, vol. 9, pp. 128494-128505, 2021. <https://doi.org/10.1109/ACCESS.2021.3111035>.
- [13] N. Liu, J. Zhao, "A BERT-based aspect-level sentiment analysis algorithm for cross-domain text," *Comput. Intell. Neurosci.*, Art. no. 8726621, 2022. <https://doi.org/10.1155/2022/8726621>.
- [14] I. Úveges and O. Ring, "HunEmBERT: A fine-tuned BERT-model for classifying sentiment and emotion in political communication," *IEEE Access*, vol. 11, pp. 60267-60278, 2023. <https://doi.org/10.1109/ACCESS.2023.3285536>.
- [15] X. Yan, F. Jian and B. Sun, "SAKG-BERT: Enabling language representation with knowledge graphs for Chinese sentiment analysis," *IEEE Access*, vol. 9, pp. 101695-101701, 2021. <https://doi.org/10.1109/ACCESS.2021.3098180>.
- [16] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, "RoBERTa: A Robustly Optimized BERT Pretraining Approach," *arXiv Preprint*, arXiv:1907.11692, 2019.
- [17] P. He, X. Liu, J. Gao, and W. Chen, "DeBERTa: Decoding-enhanced BERT with disentangled attention," *arXiv Preprint*, arXiv:2006.03654, 2020.
- [18] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. Salakhutdinov, and Q. V. Le, "XLNet: Generalized autoregressive pretraining for language understanding," *arXiv Preprint*, arXiv:1906.08237, 2019.
- [19] A. Graefe et al., "Comparing election forecasting methods: Bayesian and frequentist approaches," *Int. J. Forecasting*, vol. 33, no. 2, pp. 391-400, 2017.
- [20] S. Devlin et al., "BERT for targeted sentiment analysis of political tweets," *Proc. IEEE Int. Conf. Data Mining*, 2020, pp. 1122-1127.
- [21] L. Derczynski et al., "Political sentiment analysis with BERT: A case study of UK elections," *IEEE Trans. Comput. Soc. Syst.*, vol. 8, no. 4, pp. 1023-1032, 2021.
- [22] R. K. Koul et al., "Transformer models for election prediction: BERT vs. RoBERTa," *IEEE Access*, vol. 10, pp. 34567-34578, 2022.
- [23] R. Cai et al., "Sentiment analysis about investors and consumers in energy market based on BERT-BiLSTM," *IEEE Access*, vol. 8, pp. 171408-171415, 2020. <https://doi.org/10.1109/ACCESS.2020.3024750>.
- [24] Z. Gao, A. Feng, X. Song and X. Wu, "Target-dependent sentiment classification with BERT," *IEEE Access*, vol. 7, pp. 154290-154299, 2019. <https://doi.org/10.1109/ACCESS.2019.2946594>.
- [25] C. Huang, "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *Lancet*, vol. 395, pp. 497-506, 2020. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5).
- [26] C. Wang, P. W. Horby, F. G. Hayden, and G. F. Gao, "A novel coronavirus outbreak of global health concern," *Lancet*, vol. 395, pp. 473-490, 2020. [https://doi.org/10.1016/S0140-6736\(20\)30185-9](https://doi.org/10.1016/S0140-6736(20)30185-9).
- [27] J. Zhou, J. X. Huang, Q. V. Hu, and L. He, "SK-GCN: Modeling syntax and knowledge via graph convolutional network for aspect-level sentiment classification," *Knowl.-Based Syst.*, vol. 205, Art. no. 106292, 2020. <https://doi.org/10.1016/j.knsys.2020.106292>.
- [28] K. Bijari, H. Zare, E. Kebriaei, and H. Veisi, "Leveraging deep graph-based text representation for sentiment polarity applications," *Expert Syst. with Appl.*, vol. 144, Art. no. 113090, 2020. <https://doi.org/10.1016/j.eswa.2019.113090>.
- [29] H. Cai, Y. Tu, X. Zhou, J. Yu, and R. Xia, "Aspect-category based sentiment analysis with hierarchical graph convolutional network," *Proc. 28th Int. Conf. Comput. Linguistics*, 2020, pp. 833-843. <https://doi.org/10.18653/v1/2020.coling-main.72>.
- [30] J. van der Geer, J.A.J. Hanraads, R.A. Lupton, "The art of writing a scientific article," *Journal of Science Communication*, vol. 16, issue 3, pp. 51-59, 2000.
- [31] J. E. Campbell, *The American Campaign: U.S. Presidential Campaigns and the National Vote*, 2nd ed. Texas A&M University Press, 2008.
- [32] M. A. Al-Garadi et al., "Analyzing voter behavior via BERT-based topic modeling," *Proc. ACM Web Sci. Conf.*, 2021, pp. 1-9.
- [33] P. G. Ferreira et al., "Cross-domain sentiment analysis for political campaigns using adversarial BERT," *Knowl.-Based Syst.*, vol. 238, Art. no. 107842, 2022.
- [34] T. Chen and D. Wijaya, "Multilingual election sentiment analysis with BERT: A comparative study," *Proc. EMNLP*, 2021, pp. 456-470.



Dr. G UMA MAHESWARI holds a Ph.D. in Software Defined Networking, an M.E. in Network Engineering, and a B.E. in Computer Science and Engineering, all with first-class distinctions. Her academic journey reflects a strong foundation in computer science and networking. Her research interests are centered around advanced networking technologies, including Software Defined Networking (SDN), network security, and optimization techniques. She has also explored applications of artificial

intelligence in networking, as evidenced by her publications on quantum machine learning and data analytics.



SANGEETHA V is currently pursuing Ph.D. in the domain of Detecting social media threats using Machine Learning. She is currently an Assistant Professor at SRM Easwari College of Engineering and Technology, bringing over 5 years of teaching experience in computer science and engineering. Her expertise spans network security, machine learning, and Big Data.



Dr MEENAKSHI A is a Professor and Head of the Department of Computer Science and Engineering at Kamaraj College of Engineering and Technology. With over 24 years of distinguished teaching experience and 3 years in industry, she is a seasoned academic and researcher. She holds a Ph.D. and her core areas of specialization include Machine Learning, Big Data, and Quantum Computing, with a strong research focus on cybersecurity applications for smart environments and the Internet of Things (IoT). She has an extensive publication record in Scopus-indexed journals, with recent work focused on the detection of cyber intrusions and network attacks using advanced deep learning models, optimal learning mechanisms, and ensemble transfer learning techniques. She has successfully led several institutional research projects as Chief Coordinator, focusing on practical IoT applications and cloud-based learning systems.



Dr. VASUDEVAN KARUPPIAH is an Associate Professor in the Department of Electronics and Communication Engineering at Thiagarajar College of Engineering, Madurai, India. His research expertise lies in the areas of electromagnetic bandgap structures, metamaterials, MIMO antennas, 5G/6G communication systems, and signal integrity for high-speed digital circuits. He has an extensive publication record, with over 38 works including journal articles, conference papers, and book chapters, primarily in high-quality, Scopus-indexed venues such as AEU – International Journal of Electronics and Communications, Optics Communications, and the International Journal of Microwave and Wireless Technologies. His recent research focuses on designing compact, high-performance antennas for next-generation applications, including 5G/6G terminals, body area networks (WBAN) for biotelemetry, and V2X communication.

...