Human-Centric Intelligent Systems (2024) 4:437–446
https://doi.org/10.1007/s44230-024-00078-y

RESEARCH ARTICLE

Social Media Profiling for Political Affiliation Detection

Ihsan Ullah Khan[1] · Muhammad U. S. Khan[2]

Received: 31 January 2024 / Accepted: 10 July 2024 / Published online: 20 July 2024

Abstract
The notion of discerning political affiliations from users' social media behavior instills a sense of unease in many. Democracy necessitates that individuals' political affiliations remain private, and social media challenges this foundational principle of democracy. This study uses BERT, a pre-trained language model to analyze X's (formally Twitter) users and their political affiliations to understand that how much it is easy now to find the political affiliation of people. We collect posts in both English and Urdu languages from different political leaders and their followers, which are used to fine-tune the BERT model. The model classifies the users' profiles into Pro, Neutral, or Anti-government classes. To assess the performance of the proposed method, experiments are conducted to evaluate its accuracy, efficiency, and effectiveness. The findings of this study confirm the hypothesis that it is easy to detect the political affiliation of individuals using social media with high accuracy (69% for English and 94% for Urdu language) and it can undermine democracy.

Keywords Sentiment analysis · Political affiliation · Social Media · BERT · Text classifications · Social media monitoring · Opinion mining

1 Introduction

Social media platforms, such as X (former Twitter) have effects on politics in democratic and authoritarian countries, such as Pakistan. The debates on the internet amplify global economic, political, and cultural frustrations. Social media platforms enable people to express opinions on politics, leading to political affiliations and online discussions. It has been observed that social media has been a key tool in coordinating protests and providing a platform for the opposition in autocratic regimes [1]. Social media can lead to backlash, criticism, discrimination, and exclusion. Anti-government views can result in exile, job difficulties, and reputation damage. Misinformation can also cause mistrust and legal consequences, such as charges or defamation lawsuits.

It is crucial to be cautious and mindful of social media platforms. In volatile political environments, governments may face opposition or challenges from individuals or groups with opposing views. In such situations, individuals who express their political views publicly may be subject to biases or challenges in various contexts. Frequently on social media, individuals express themselves without thoughtful consideration, leading to instances of trolling and conflicts.

Identifying the political leanings of users also serves as a valuable tool for uncovering the automated accounts (Bots [2, 3]) deployed by political parties to prop- agate their viewpoints. Distinguishing these bots is essential for refining an accurate representation of the genuine public sentiments toward various political parties.

This study examines the hypothesis that political affiliations can be readily identified through social media profiles, with political parties potentially exploiting this information to their advantage. The primary objective of this research is to alert users to exercise caution in their social media posts,[1] particularly in countries such as Pakistan, where the protection of individual rights may be incomplete, highlighting the potential negative consequences.

M. U. S. Khan prepares the idea of the paper whereas I. Khan conducted experiments and prepared the results.

✉ Muhammad U. S. Khan
usman.shahid@sines.nust.edu.pk

Ihsan Ullah Khan
ihsankhattak770@gmail.com

1 Computer Science, COMSATS University Islamabad, Abbottabad 22044, KPK, Pakistan

2 School of Interdisciplinary Engineering & Sciences (SINES), National University of Sciences & Technology (NUST), Islamabad 44000, Pakistan

[1] Note: The terms Tweet and Post are used interchangeably in this work.

Table 1 Comparative analysis of different approaches

Paper	Methodology	Data Source	Results
[4]	KGSN	Social network data	87%
[5]	RI and LSH	Twitter	87% Recall
[6]	LSTM	Twitter (Political tweets)	74% Accuracy
[7]	DNN	E-commerce platforms	88 Accuracy
[8]	CNNs	Chinese conversation corpus	CNN outperformed
[9]	Naive Bayes, SVM	Twitter (Hindi tweets)	78.4% Accuracy
[10]	Na¨ıve Bayes, SVM, Logistic Regression, Random Forest	Twitter	95% accuracy
[11]	BERT	Online news articles	64.4% Accuracy
[12]	LLM's	Twitter	63% Accuracy
[13]	NgramCNN	Twitter	91.2% Accuracy
[14]	KGSN	Social network data	87% Accuracy
[15]	ML Models	Twitter	91.44% F1-Score
[16]	CNN	Twitter	84% Accuracy

This study uses a BERT (Transformer architecture) to categorize X's profiles that frequently express views with respect to Pakistani politics as pro-government, neutral, or anti-government based on their tweet content. While this study focuses solely on the social media landscape of one country, its findings are applicable universally to all countries.

This study involves collecting 15,000 English language posts (tweets) and 11,000 Urdu language posts from the X website. These posts are categorized into pro- government, anti-government, and neutral groups based on their content. The models are then trained on this dataset to predict the label of any given tweet. The results show that the BERT model is able to predict the affiliation of the users better than the other state-of-the-art models. Moreover, the graph of users' profiles is also made in this project to group similar profiles that are working together to spread the same political ideology. The contribution of this paper are as follows:

- The study shows that the detection of political affiliations has become more accessible with the use of Artificial Intelligence.
- The BERT models performs better than CNN, LSTM, KNN, Decision Tree, ANN, and SVM on the detection of political affiliation of users on social media both in English and Urdu language.
- A dataset of "pro-government", "anti-government", and "neutral" with respect to the government's stance on social media is collected and made available for further research
- The BERT model performs with 69% and 94% validation accuracy results on English and Urdu datasets for political affiliation detection.

Apart from the Introduction Section, this paper is organized as follows: In Section "Literature Review", we provide a literature review of previous works on sentiment analysis and opinion mining. The methods employed in this study are outlined in the "Methodology" Section. The experimental design is detailed in "Experimental Setup" Section, and the results of the study are presented and analyzed in "Results and Discussion" Section. The conclusions of our research are presented in the last section. The video demo of the work is presented at the link.[2]

2 Literature Review

Table 1 presents the methodologies, data used, and results of different related studies. Sentiment analysis is a technique that uses natural language processing to detect emotions and attitudes in a given event or political situation. In the past, various techniques have been proposed by researchers in an effort to improve the accuracy of predicting social opinions and behavior through text analysis. This addresses the challenges inherent in using natural language for sentiment analysis and develops more effective methods for understanding and interpreting social attitudes and behaviors [6, 17]. Research on analyzing movie and electronic product reviews can be found in [4, 7]. Moreover, numerous studies use sentiment analysis to detect events and news from Twitter data [4, 18, 19].

M Hasan et al. [17] propose a system that can efficiently identify both major and minor news in real-time, while also being computationally efficient. This system could be useful

[2] Video Demo Link.

for quickly and accurately identifying significant events and news on Twitter. There have been several studies that have analyzed the sentiment of political posts, including [3, 6, 9, 10, 16, 20–24]. Studies on political tweet sentiment analysis have emphasized its value in understanding emotions and attitudes in natural language text, with ongoing research aiming to improve the accuracy and performance of sentiment analysis models.

Nedjah et al. [25] conducted a study on the impact of hyperparameters on model performance and found that their results showed an 18 percent improvement over base models. However, they also noted the need for further research to fully understand the model's performance with a larger dataset and more classes of polarity. In another research, Wilfrido et al.[26] suggest using Vector Space Models (VSM) with a high number of dimensions as a way to analyze the effectiveness of a lexicon-based approach for sentiment analysis. They applied this approach to the task of determining the sentiment (i.e. polarity) of posts related to politics in Mexico.

Alfina et al. [10] utilize hashtags in posts to identify the sentiment (i.e., polarity) of political posts. To do this, they separated the hashtags from the posts and manually assigned them to one of four categories based on their sentiment. These categories were then used to classify the sentiment of the posts.

Recent advancements in sentiment analysis have seen the exploration of topic- specific political stances in social networks, as demonstrated by Wu et al. [14]. Demidov et al [11] delves into the classification of political bias within news content using advanced techniques like BERT. Mets et al. (2024) [12] tackle the challenging task of automated stance detection in complex topics, particularly focusing on immigration within polarizing news media. Thapa et al. [15] address the detection of stances and hate events related to climate activism in tweets, showcasing the capabilities of machine learning models.

In another study [9] the authors use sentiment analysis to predict the popularity of political parties in India using a dataset of 42,000 posts. The data were classified into positive, negative, and neutral categories using dictionary-based, naive Bayes, and SVM algorithms. Their results show an accuracy of 78% in predicting the sentiments of posts from five political parties. While previous research has applied various machine- learning techniques to classify posts, few have examined how hyperparameters impact prediction accuracy. A recent study sought to predict the political views of users in Pakistan based on their Twitter posts. Using a dataset of 10,000 posts, the researchers employed a CNN (Convolutional Neural Network) to classify the sentiments of the posts as positive and negative. The analysis achieved an accuracy rate of 84% when predicting the sentiments of posts [16].

Diverging from the previously cited studies, this research introduces an approach utilizing the Transformer (BERT) model, a cutting-edge natural language processing (NLP) model developed by Google. The study assesses the efficacy of the BERT model in detecting political affiliation by analyzing historical posts from user profiles, aiming to discern the profile's political inclination toward or against the government. The BERT model is compared with the techniques presented in the above-mentioned literature and the outcomes indicate that BERT outperforms the other models.

3 Methodology

In this section, we outline the process of preparing the data, building the model, and generating prediction results. Our model for classification tasks includes four stages: data collection, data preprocessing, word embedding, and BERT Model (Bidirectional Encoder Representations from Transformers). To begin, we collect relevant data for our analysis. We then preprocess the data by removing special symbols, extra space, punctuation, and emojis. We apply word-embedding techniques to represent the words in the data as numerical vectors, which allows us to input them into the BERT model. Finally, we use the BERT model to classify the data and generate prediction results (Fig. 1).

3.1 Data Collection

This research uses the Twitter Streaming API to gather data from 80 political leaders in Pakistan. The data is categorized into "pro-government", "anti-government", and "neutral" with respect to the government's stance on social media. The decision to manually label the data, rather than relying on automated methods, underscores the importance of accuracy and context in the classification process. The datasets of posts contain both the English and Urdu language posts. Furthermore, to facilitate further research and replication of results, the entire dataset has been made publicly available [27]. Interested researchers can access the dataset through the provided link, promoting transparency and collaboration within the academic community. The avail- ability of this dataset is expected to contribute significantly to future studies in the field, providing a valuable resource for both linguistic and social media research. The detail of the posts is mentioned in Table 2.

3.2 Pre-processing

Pre-processing is a crucial step in preparing a dataset downloaded from X (Twitter) for use with a BERT model, as it involves cleaning and preparing the data in a suitable format for the model to process. This includes removing

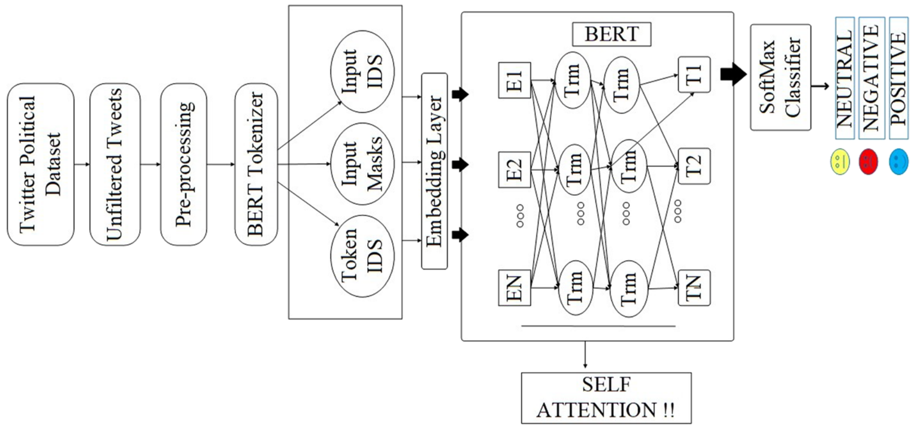


Fig. 1 Methodology diagram to classify the posts

any irrelevant or spammy content, normalizing the text by lowercasing words, expanding contractions, removing punctuation, and special characters, tokenizing the text into individual words and encoding them using a suitable vocabulary, and adding special tokens such as separation or padding tokens to ensure that all sequences in the dataset have the same length.

3.3 Word-Embedding

Word embedding is crucial in natural language processing (NLP) tasks as it allows models to represent words in numerical form. BERT uses contextualized word embeddings, generating embeddings based on the context in which they appear. These embeddings can be fine-tuned for various NLP tasks. BERT's contextualized embeddings are generated using a multi-layer bidirectional transformer architecture.

3.4 Model

BERT [28] is composed of multiple layers of attention-based transformer blocks, which are a type of neural network architecture that allows the model to process input sequences flexibly and efficiently. One key feature of BERT is its ability to perform "masked language modeling" a self-supervised learning technique that involves randomly masking a portion of the input tokens and then predicting the masked tokens based on the context provided by the rest of the sequence. This allows BERT to learn rich and nuanced representations of words and their relationships to one another, which is essential for many NLP tasks. In addition to its transformer blocks, BERT also includes embedding layers that map each token in the input sequence to a fixed-length vector representation. These embeddings are used as input to the transformer blocks and are updated during training to capture the context and dependencies between words in a sentence.

Table 2 Datasets of political affiliation [27]

Dataset	Language	Size
EnglishPoliticalPosts	English	15,000
UrduPoliticalPosts	Urdu	11,000

In this study, the BERT model is fine-tuned on a dataset (See, Table 2) of 15000 English and 11000 Urdu language posts, in addition to its original training on a dataset of approximately 3.3 billion words and 12 layers with hidden states of size 768. Finetuning pre-trained models on smaller, task-specific datasets is a common method for improving their performance on a specific task or domain. In this case, additional training on posts enhances the BERT model's ability to comprehend the linguistic patterns and characteristics specific to posts. This enhances its capacity to categorize social media posts as pro-government, neutral, or anti-government. Table 3, presents the summarized hyperparameter values. The specifics can be found in Section IV.

3.5 Predictions

After successfully training and testing, the BERT model is utilized to classify social media profiles on X (Twitter) based on user's posts as either pro-government, neutral, or anti-government. Moreover, the model also classifies the political sentiment of the user's followers, enabling it to determine the political orientation of the user's followings.

Table 3 Hyperparameters of the BERT Model

Hyperparameter	Value
Learning rate	2e-5
Batch size	16
Number of epochs	30
Dropout rate	0.5
Max sequence length	308
Optimizer	Adam
Activation function	Softmax

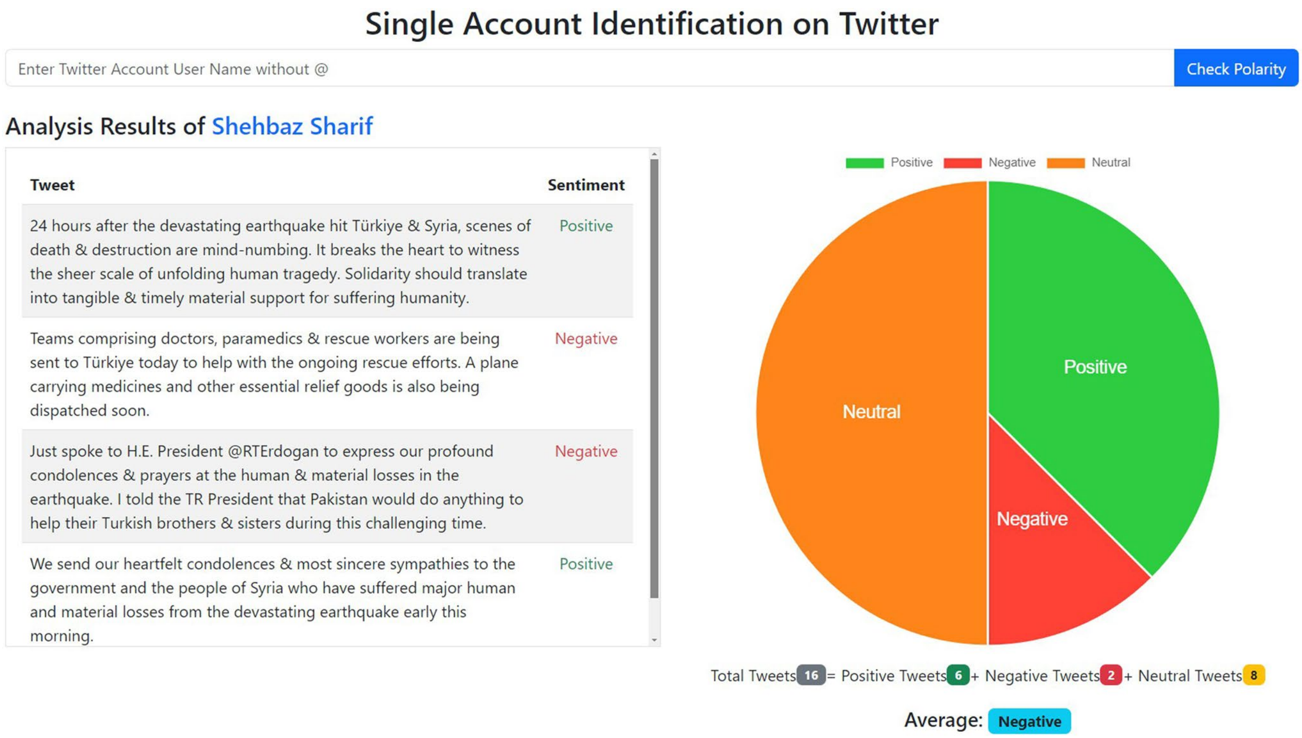


Fig. 2 User Profile Identification: The system uses a predictive model to detect whether a user profile is engaging in pro-government, neutral, or anti-government activities

The input posts are fed into the model, which uses the information learned during training to make predictions about the political sentiment of the posts on a social media user profile. The model's output is a classification of the profiles as either government-pro, neutral, or anti-government based on its analysis of the posts' content. This provides valuable insight into the political leanings of both individual users and their followers on social media. Figure 2 shows the political alignment of one user (political leader of one party) whereas Fig. 3 shows its followers.

4 Experimental Setup

The experiments are conducted on an HP VICTUS 15 machine with a 12th Gen Core i5 Octa-Core Processor, 16GB of RAM, a 512GB SSD, and a 4GB NVIDIA GeForce GTX1650 graphics card. A range of hyperparameter settings are tested, including learning rates from 1 to 5e–5, batch sizes from 16, and 30 epochs. The model used in the experiment is made up of 12 transformer blocks with a hidden size of 768 and 12 self-attention heads and has approximately 110 million trainable parameters.

The dataset (See, Table 2) contains 15000 English and 11000 Urdu language posts. The training and validation data are split in a 90/10 ratio. After the validation analysis, new posts are collected and they are used as a test set. The test set includes 100 positive, 100 neutral, and 100 negative posts against the government opinion both in English and Urdu. The models are evaluated manually on the test set.

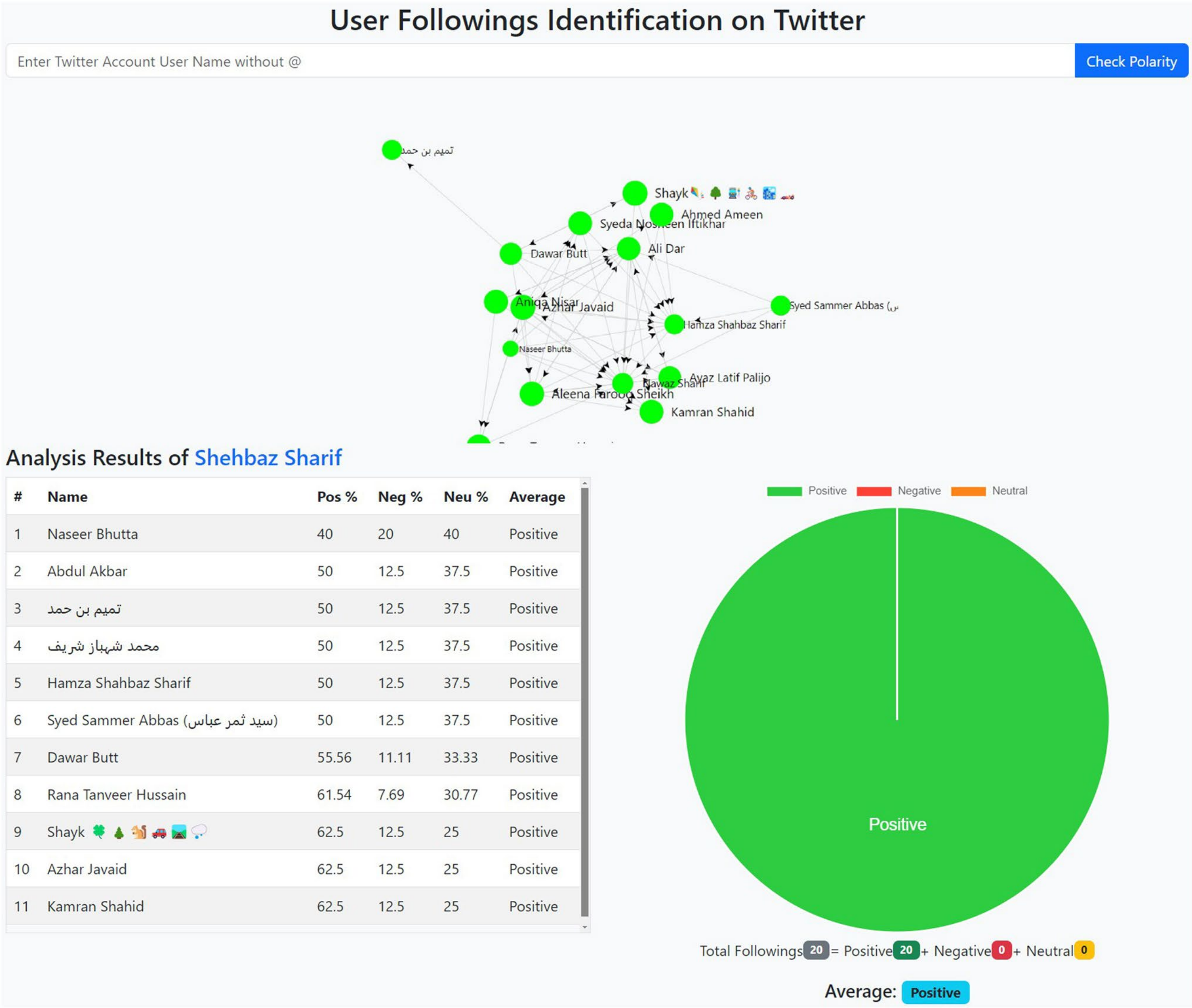
5 Results and Discussion

To determine the effectiveness of the proposed BERT model, experiment is conducted to answer the following questions:

- How accurately does the BERT model detect the polarity of political posts?
- How does the choice of hyperparameters impact the accuracy of the model?

The settings of hyperparameters allow us to quantify the effectiveness of the model and identify any areas for improvement. Through these experiments, we aim to gain a better understanding of the performance of the BERT model and identify the most effective settings for its hyperparameters.

The experiments are inspired by previous research such as [10, 13, 29]. We use accuracy, precision, recall, and



#	Name	Pos %	Neg %	Neu %	Average
1	Naseer Bhutta	40	20	40	Positive
2	Abdul Akbar	50	12.5	37.5	Positive
3	تمیم بن حمد	50	12.5	37.5	Positive
4	محمد شہباز شریف	50	12.5	37.5	Positive
5	Hamza Shahbaz Sharif	50	12.5	37.5	Positive
6	Syed Sammer Abbas (سید ثمر عباس)	50	12.5	37.5	Positive
7	Dawar Butt	55.56	11.11	33.33	Positive
8	Rana Tanveer Hussain	61.54	7.69	30.77	Positive
9	Shayk	62.5	12.5	25	Positive
10	Azhar Javaid	62.5	12.5	25	Positive
11	Kamran Shahid	62.5	12.5	25	Positive

Fig. 3 User Followings Profiles Identification: Identifying pro-govt, neutral profiles, and anti-govt profiles on social media and grouping the connected profiles. Profile graphs show how people with similar attention are grouped together

F-measure metrics to evaluate the performance of the model. We conduct experiments with different learning rates, dropout rates, and epoch settings to examine how these hyperparameters affect the model's performance. Overall, these experiments allowed us to assess the effectiveness and robustness of the proposed technique in detecting the polarity of political posts.

6 Experiment

In the experiment, a BERT model is used to perform political sentiment analysis on English and Urdu language datasets. By using a combination of 30 epochs, a learning rate 3e-5 in the English language model, and a batch size of 16, the model for the English language achieves an accuracy of 69% along with 74% precision, 74% recall, and 70% F-measure as shown in Fig 4. While on the Urdu language dataset, the model achieves a validation accuracy

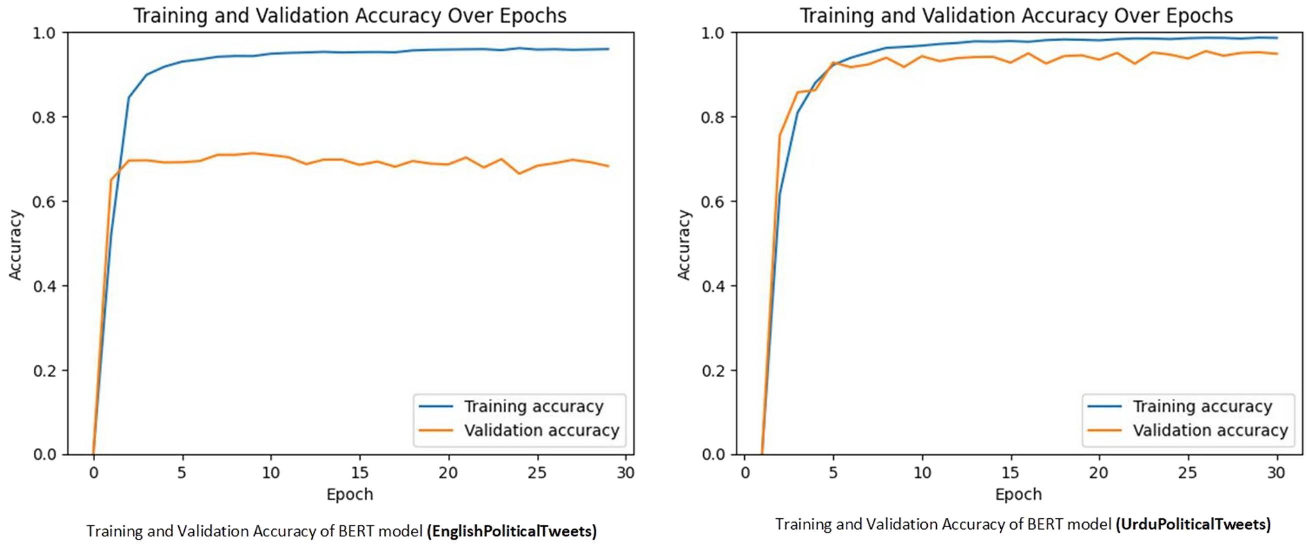


Fig. 4 Training and Validation Accuracy of BERT model English and Urdu language posts

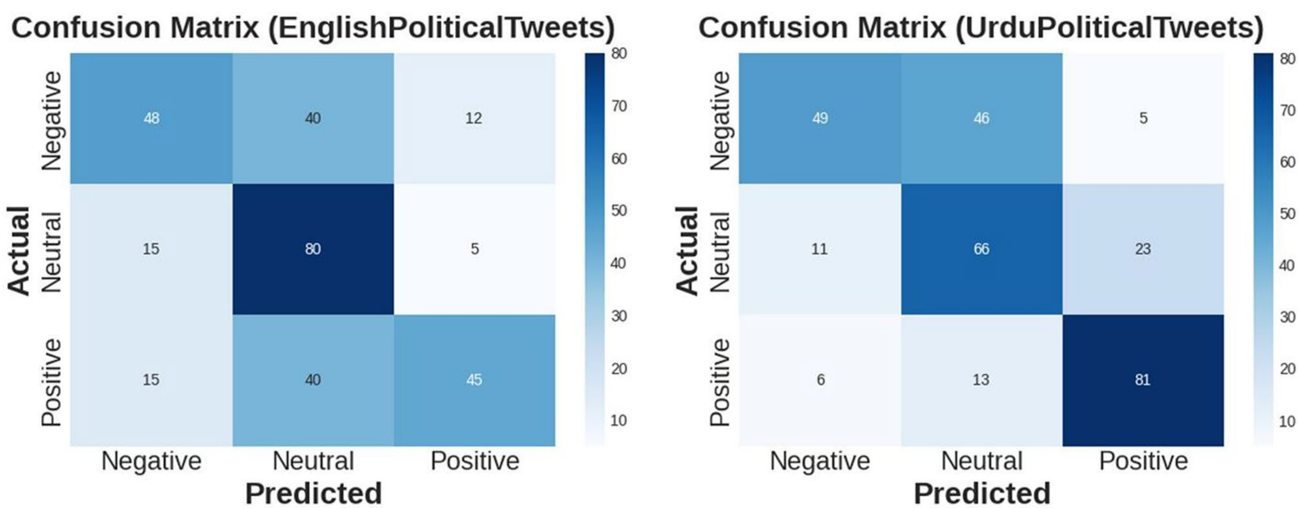


Fig. 5 Confusion matrix English and Urdu language posts

of 94% with 74% precision, 74% recall, and 73% f1-score with a learning rate of 2e–5.

After the validation analysis, new posts are collected and they are used as test set. The models are evaluated manually on test sets for positivity, neutrality, and negativity of posts in favor, neutral, or against the government opinion. The confusion matrix of the model on test set both in English and Urdu posts is presented in Fig. 5.

The English language model achieves testing accuracy of 45%, 80%, and 40% in identifying positive, neutral, and negative sentiment. Similarly, the Urdu language model achieves an accuracy of 81%, 66%, and 49% in identifying positive, neutral, and negative sentiments. The results demonstrate the proficiency of these simple models in detecting and classifying political sentiment in posts.

The results have shown that the BERT model can detect political sentiment in social media posts with high accuracy (Fig. 6). Similarly, hyperparameter settings play an important role in improving the performance of the model. Table 3, presents the summarized hyperparameter values that help in achieving the highest accuracy of the model.

6.1 Experiment with Baseline Models

Several machine learning models are also trained as baselines to compare their performance against the BERT model.

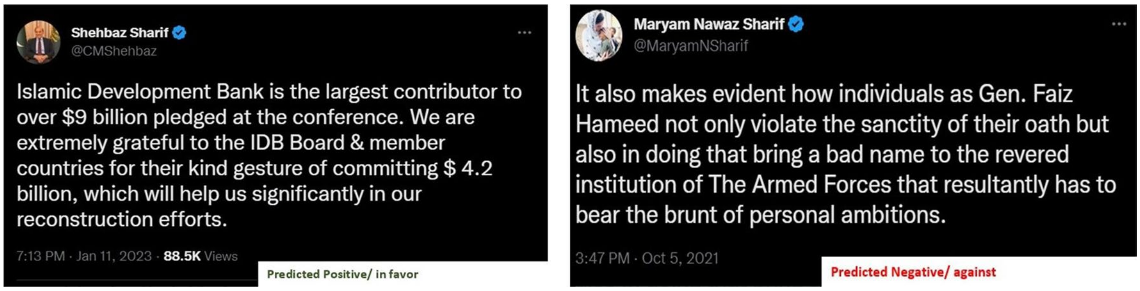


Fig. 6 Predicting the sentiment of a tweet using the BERT model

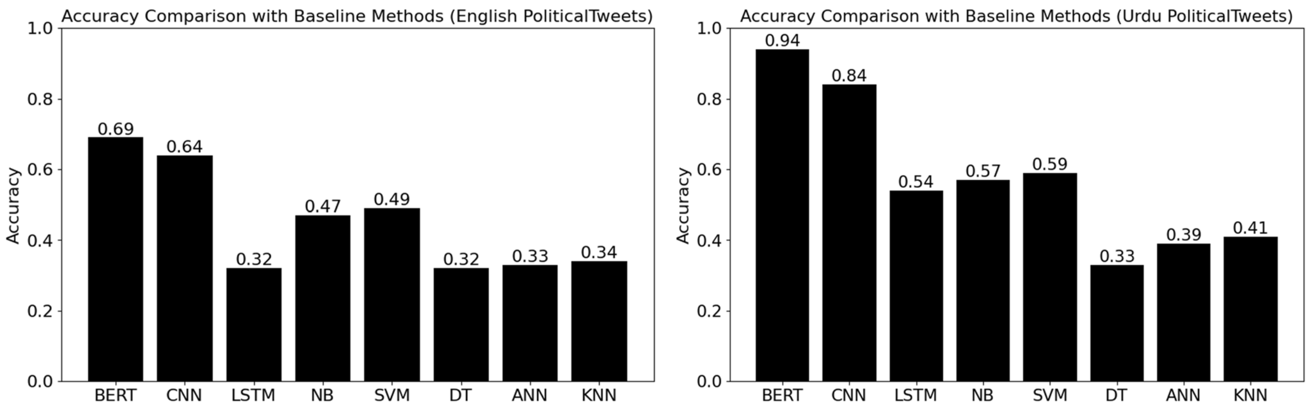


Fig. 7 Accuracy Comparison with Baseline Methods in English and Urdu languages

These models included the CNN, Naive Bayes, Support Vector Machine (SVM), Decision Tree, K-Nearest Neighbors (KNN), Artificial Neural Network (ANN), and Long Short-Term Memory (LSTM).

1. CNN: The Convolutional Neural Network (CNN) is a popular choice for tasks involving images or video data. A CNN model is trained on a training dataset to classify the sentiment of posts as positive, negative, and neutral. The CNN model achieves an accuracy of 64% in English and 84% in Urdu Language posts classification.
2. The Long Short-Term Memory (LSTM) model is a type of Recurrent Neural Net- work (RNN) that can remember information over a longer period. The model achieves an accuracy of 32% on English and 54% on Urdu Language posts.
3. Naive Bayes: The Naive Bayes Classifier is a simple and effective algorithm used for classification tasks. It is based on Bayes' theorem. The classifier calculates the probability of each token for both classes and makes predictions based on the conditional probability of each class given the tokens in each tweet. The model achieves an accuracy of 47% on English and 56% on Urdu Language posts.
4. K-nearest Neighbor: The K-Nearest Neighbors (KNN) algorithm is a classification method that uses the K-Nearest neighbor from a set of training examples to make predictions. The model achieves an accuracy of 34% on English and 42% on Urdu Language posts.
5. Decision Tree: The Decision Tree is a machine-learning algorithm that can be used for both regression and classification tasks. It works by creating a hierarchical model that makes predictions based on decision rules learned from training data. In this experiment, a Decision Tree achieves an accuracy of 32% on English and 33% Urdu Language posts.
6. ANN: The Artificial Neural Network (ANN) is a machine-learning algorithm that can be used in classification tasks. The model achieves an accuracy of 33% on English and 39% on Urdu Language posts.
7. Support Vector Machine: The Support Vector Machine (SVM) is a supervised learning algorithm that uses a kernel function to draw a hyperplane boundary between different output classes. The model achieves the highest

accuracy of 49% on English and 59% on Urdu Language posts.

The comparison has shown that BERT models perform better than the above- mentioned models in detecting the political affiliation of the users on social media both in English and Urdu languages (see Fig. 7).

7 Conclusions

Democracy relies on the privacy of people's political beliefs, and social media poses a serious challenge to this fundamental democratic principle. This research employs BERT, a pre-trained language model, to examine users on X and their political association, aiming to gauge the ease with which one can now determine people's political affiliations. The works aim to caution individuals to use social media carefully as authoritarian governments can easily use social media to understand the political affiliation of people even before elections and can use that information for their own interests. We gather posts both in English and Urdu languages from various Pakistani political leaders and their followers to fine-tune the BERT model. The model categorizes users into Pro, Neutral, or Anti-government classes. The study's findings affirm the hypothesis that the detection of political affiliations has become more accessible with the use of Artificial Intelligence.

Acknowledgements We extend our gratitude to Nasir Khan and Tehnees Sheraz from the Department of Computer Science, COMSATS Abbottabad, Pakistan for their invaluable assistance in both the data collection and post-labeling processes. Additionally, we are thankful to the reviewers for their insightful feedback, which significantly enhanced the quality of this paper.

Data availability Data is already made public and its link provided in the paper https://ieee-dataport.org/documents/political-tweets-sentiment-anayslis-pro-neutral-and-anti-govt-activities.

References

1. Zhuravskaya E, Petrova M, Enikolopov R. Political effects of the internet and social media. Annu Rev Econ. 2020;12:415–38.
2. Mohammad S, Khan MU, Ali M, Liu L, Shardlow M, Nawaz R. Bot detection using a single post on social media. In: 2019 Third World Conference on smart trends in systems security and sustainability (WorldS4), 2019; pp. 215–220. IEEE.
3. Khan MUS, Ali M, Abbas A, Khan SU, Zomaya AY. Segregating spammers and unsolicited bloggers from genuine experts on twitter. IEEE Trans Depend Secure Comput. 2016;15(4):551–60.
4. Joshi R, Tekchandani R. Comparative analysis of twitter data using supervised classifiers. In: 2016 International Conference on Inventive Computation Technologies (ICICT), 2016; vol. 3, pp. 1–6. IEEE.
5. Hasan M, Orgun MA, Schwitter R. Twitternews: real time event detection from the twitter data stream; 2016. Available at: https://peerj.com/preprints/2297/.
6. Ansari MZ, Aziz M-B, Siddiqui MO, Mehra H, Singh KP. Analysis of political sentiment orientations on twitter. Proc Comput Sci. 2020;167:1821–8.
7. Hu Z, Hu J, Ding W, Zheng X. Review sentiment analysis based on deep learning. In: 2015 IEEE 12th International Conference on e-Business Engineering, 2015; pp. 87–94. IEEE.
8. Bandpey R, Zoriyatkha A. Sentiment classification with convolutional neural networks; 2017. https://doi.org/10.13140/RG.2.2.25355.11040.
9. Sharma P, Moh T-S. Prediction of Indian election using sentiment analysis on Hindi twitter. In: 2016 IEEE International Conference on big data (big Data), 2016; pp. 1966–1971. IEEE.
10. Alfina I, Sigmawaty D, Nurhidayati F, Hidayanto AN. Utilizing hashtags for sentiment analysis of tweets in the political domain. In: Proceedings of the 9th International Conference on machine learning and computing, 2017; pp. 43–47.
11. Demidov D. Political Bias of News Content; 2023. Available at: https://www.researchgate.net/publication/377074277_Political_Bias_of_News_Content_Classification_based_on_Individual_Articles_and_Media/citations.
12. Mets M, Karjus A, Ibrus I, Schich M. Automated stance detection in complex topics and small languages: the challenging case of immigration in polarizing news media. PLoS ONE. 2024;19(4):0302380.
13. C¸ ano E, Morisio M. A deep learning architecture for sentiment analysis. In: Proceedings of the International Conference on geoinformatics and data analysis, 2018; pp. 122–126.
14. Wu K, Zhou Y, Ma J, Guo X. Topic-specific political stance inference in social networks with case studies. IEEE Access. 2024;12:21921–35.
15. Thapa S, Rauniyar K, Jafri F, Shiwakoti S, Veeramani H, Jain R, Kohli GS, Hu¨rriyeto˘glu A, Naseem U. Stance and hate event detection in tweets related to climate activism-shared task at case 2024. In: Proceedings of the 7th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE 2024), 2024; pp. 234–247.
16. Bukhari SM, Khan MU, Khan T, Ali M. Social media, ai, and political affiliation. In: 2021 International Conference on innovative computing (ICIC), 2021; pp. 1–6. IEEE.
17. Hasan M, Orgun MA, Schwitter R. Real-time event detection from the Twitter data stream using the Twitternews+ framework. Inf Process Manag. 2019;56(3):1146–65.
18. Chen J, Hossain MS, Zhang H. Analyzing the sentiment correlation between regular tweets and retweets. Soc Netw Anal Min. 2020;10:1–9.
19. Jiang B, Zhang H, Lv C, Yang C. Sentiment classification based on clause polarity and fusion via convolutional neural network. In:

2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBD-Com/IOP/SCI), 2018; pp. 1039–1044. IEEE.
20. Afzal H, Khan U, Awan I, Mufti M. A simple greedy algorithm for cognitive radio networks. School of Computing, Information and Media, University of Bradford; 2013.
21. Kannangara S. Mining twitter for fine-grained political opinion polarity classification, ideology detection and sarcasm detection. In: Proceedings of the Eleventh ACM International Conference on web search and data mining, 2018; pp. 751–752.
22. Bilal M, Asif S, Yousuf S, Afzal U. 2018 Pakistan general election: understanding the predictive power of social media. In: 2018 12th International Conference on mathematics, actuarial science, computer science and statistics (MACS), 2018; pp. 1–6. IEEE.
23. Khan MU, Jawad M, Khan SU. Adadb: ADAPTIVE diff-batch optimization technique for gradient descent. IEEE Access. 2021;9:99581–8.
24. Khan W, Ali S, Muhammad UK, Jawad M, Ali M, Nawaz R. Adadiff-grad: an adaptive batch size implementation technique for diffgrad optimization method. In: 2020 14th International Conference on innovations in information technology (IIT), 2020; pp. 209–214. IEEE.
25. Nedjah N, Santos I, Macedo Mourelle L. Sentiment analysis using convolutional neural network via word embeddings. Evol Intell. 2019;15:1–25.
26. Paredes-Garc´ıa, WJ, G´omez-Loenzo RA. Polarity classification of political tweets in Mexico using a lexicon-based approach. In: 2018 XIV International Engineering Congress (CONIIN), 2018 pp. 1–7. IEEE.
27. Ullah Khan I. Political tweets sentiment analysis for pro, neutral and anti-govt activities. https://doi.org/10.21227/rf7t-6g73.
28. Kenton JDM-WC, Toutanova LK. Bert: Pre-training of deep bidirectional transformers for language understanding. In: Proceedings of NAACL-HLT, 2019; pp. 4171–4186.
29. Zhang X, Zhang L. Topics extraction in incremental short texts based on lstm. Soc Netw Anal Min. 2020;10(1):83.