IJAMEC

INTERNATIONAL JOURNAL OF APPLIED METHODS IN ELECTRONICS AND COMPUTERS

www.ijamec.org

International Open Access

Volume 12
Issue 04

December, 2024

Research Article

Political sentiment analysis using natural language processing on social media

Md Shakil Hossain [a], ***Md Rashidul Islam*** [b]*, ***Basheer Riskhan*** [a,b], ***Md Mehedi Hasan*** [a,b], ***Rabiul Islam*** [a,b]

School of Computing and Informatics, Albukhary International University, Jalan Tun Abdul Razaq 05200 Alor Setar, Kedah, Malaysia

ARTICLE INFO

Article history:
Received 12 September 2024
Accepted 6 November 2024

Keywords:
Sentiment Classification,
Sentiment Analysis Applications,
Sentiment analysis,
Natural Language Processing

ABSTRACT

In this contemporary era , social media has become an essential component of daily life as a result of the extensive use of the internet. This paper explores sentiment analysis of political topics through social media comments. We collected a large dataset of over 14,000 political comments and applied advanced machine learning models such as logistic regression , linear support vector classification , random forest, decision tree classification , and naive bayes to evaluate expressed sentiments. Performance metrics , including accuracy , precision , recall , and F1 scores , were utilized to assess these models , with Linear SVC achieving the highest accuracy at 91.18% , closely followed by Logistic Regression at 90%. This research not only evaluates model performance on political sentiment data but also addresses data imbalance, presenting actionable insights into each algorithm's suitability. Our study introduces a refined approach to political sentiment analysis by optimizing model selection for high accuracy and robustness, thus setting a foundation for effective political sentiment understanding on social media platforms.

1. Introduction

Sentiment analysis utilizing natural language processing is applied to categorize and evaluate textual data, especially product reviews, to determine whether the opinion is positive, negative or neutral. This study aims to extract subjective sentiments from textual data by applying natural language processing and computational linguistic techniques. The political discourse on social media implements this procedure, and sentiment analysis is vital for identifying public opinion, political trends, and sentiment dynamics [1].

The main aim of this study is to explore how sentiment analysis in the world of politics can be done with the help of natural language processing technology and social media data. The politics is an interesting object of the study of technology, politics, and society in the field of public opinion because of its diverse political situation and strong social media presence [2]. This task aims to extract opinions on politics from social media posts using a Natural Language Processing (NLP) algorithm. The purpose here is to discover the patterns, trends, and sentiments that will offer good recommendations for political campaigns, policy-making, and public communication. Analyzing user thoughts and emotions displayed on social media networks is an invaluable asset for analysts of politics, individuals in politics, and researchers in public opinion and mood fluctuations [3].

Social media has become a prominent means of political expression and activism in contemporary society. Social media, namely YouTube, has significantly impacted public discussions and political movements in the world. The influence of political violence on social media, particularly on YouTube, has prompted significant inquiries into the platforms' role in shaping public sentiment, spreading information, and impacting electoral results. Although political violence on social media, particularly YouTube, has recently received significant attention, this study contends that it is a more extensive issue . we must address instances of violence in the

* Corresponding author. E-mail address: mdrashidul.islam@student.aiu.edu.my
DOI: 10.58190/ijamec.2024.108

comments section of YouTube videos across all strata of society. The absence of good governance hinders the ability of the political system to effectively manage and prevent violence, both organized and unorganized, targeting various segments of society. Extended periods of military governance, persistent political intolerance, and political extremism increased this pattern [4].

This study focuses on enhancing comprehension regarding the utilization of sentiment analysis in social media, with a specific emphasis on YouTube. This is achieved by analyzing the sentimental politics associated with this topic in 2023. Sentiment analysis employs Natural Language Processing to extract, transform, and analyze opinions expressed in textual data, subsequently categorizing them as positive, negative, or neutral sentiment [5] . Most prior research has used sentiment analysis to analyze product or movie reviews, gain a deeper understanding of customers, and make informed decisions to enhance their offerings [6]. Researchers have studied sentiment analysis over the past decade, with most articles emerging and experiencing substantial growth since 2004 [7]. There are three distinct levels of sentiment analysis: sentence, document, and feature.

This study aimed to categorize opinions by analyzing sentences, documents, or attributes, distinguishing between positive and negative emotions. There are two primary methodologies for sentiment analysis: a machine-learning technique and a lexicon-based approach. The machine learning processing approach employs algorithms to extract and identify sentiments from a dataset.

In contrast, the lexicon-based approach operates by tallying the positive and negative words associated with the data. Researchers are actively developing novel and precise models for sentiment analysis. One problem when developing a model is the predominant focus on designing it for the English language [8].

2. Literature Review

Natural language processing (NLP) with sentiment analysis using pre-trained models has become one of the most prevalent techniques in analyzing political sentiment on social media. This survey implemented the first phase of word embedding and pre-trained models for the second phase of sentiment classification [9]. Das et al. (2024) employed natural language processing (NLP) techniques to analyze the mood of political talks on YouTube. The purpose of this study is to illustrate the correctness of the approach using techniques like transformer-based models, random forests, convolutional neural networks, and support vector machines. This technic focused on using algorithmic comparisons and adjustment to increase accuracy [10].

Islam et al. (2021) examined sentiment analysis in low-resource languages. They focused on how to overcome the necessity for in-depth algorithm comparisons, resource constraints, and the lack of train-test divides. Among the difficulties they have overcome are the absence of train-test divides, the scarcity of resources, and the need to do a comprehensive algorithm comparison in sentiment analysis. In order to increase the accuracy of sentiment analysis, this study employs algorithms such as SVM, Naïve Bayes, CNN, and transformers since they are classified as political sentiments. Thus far, studies have reported different accuracies. SVM attained an accuracy of 81.6% in bullying classification, but the transformer models excel in sentiment prediction for low-resource languages. Our work expands on these efforts by using reliable, publicly available models to political datasets with more resources, which yield insightful results, especially in resource-constrained environments[11].

In order to manage political sentiment data , Tusar et al. (2021) focused SVM, logistic regression, naive Bayes, and random forest models in conjunction with bag-of-words and TF-IDF techniques. A significant sentiment analysis model built on natural language processing methods and machine learning algorithms is required to comprehend the political attitudes on social media, particularly YouTube. They have had some woes, including dataset imbalance, feature extraction optimization, and pre-optimization for advanced ML algorithms. The results indicate the application of NLP technologies to text processing and ML classification algorithms, such as SVM and Logistic Regression, to classify sentiments efficiently with an accuracy rate of 77%. Although previous work tackled typical issues like feature optimization and dataset imbalance, our study improves on existing techniques by using more advanced feature extraction and model fine-tuning, with the goal of achieving greater accuracy and applicability across a range of political discourses[12]. Shafin et al. (2020) utilized sophisticated sentiment analysis techniques to determine people's emotions in political posts on social media platforms. Their accuracy rate was approximately 97%. Because the complex structure of the data cannot be handled within this framework, they have historically relied on human analysis of public mood on social networks, which is expensive, imprecise, and not scalable. Meanwhile, methodologies that involve text data preprocessing, feature extraction, and sentiment classification using NLP algorithms, such as Naive Bayes and SVM, as in previous research studies. The previous research reported accuracies that ranged between 54.24% and 85.58%. While their method worked well, it mostly depended on manual data processing, which is less scalable for big, complicated datasets. By automating text preparation, our study overcomes this constraint and enables great scalability and flexibility across a range of platforms and political circumstances[13].

Moon et al. ((2021) have focused on creating a sentiment analysis engine that is customized for politics on social media sites to serve a better purpose by incorporating the multi-source input to gain to analyze sentiments in online micro-media, including textual, audio, and visual ones, which is done using the intricate methods of machine learning that are apt to understand public opinions exactly. The research consists of obtaining, processing, and analyzing sentiments from different media sources, using schemes of sentiment dictionary, maximum entropy models, and semantic classifiers of the Naive Bayes, SVM, and ESME models. By focusing on multimedia sources, their approach broadened sentiment analysis' scope; however, our study centers on optimizing textual sentiment classification for political discussions, utilizing high-performing NLP algorithms tailored to this context [14]. Goel et al. (2018) have researched opinion analysis emphasizing the sequential importance of the words, especially in sentiment analysis using natural language processing on Twitter data. Research uses tree incorporation, recursion neural networks, embeddings based on information gain and bigrams, naive Bayes, naive Bayes, and SVM classification to conduct, and SVM classification to conduct sentiment analysis on social networks. Building on this, our work improves the precision and recall for political debate on social media by particularly fine-tuning feature engineering for political mood. We also take into account recent developments in embedding and classifier tuning that better capture the complex vocabulary of political discourse, which is consistent with Goel's findings on embedding approaches [15]. Developing potential solutions to future problems includes sophisticated classifiers , semi-supervised learning , and addressing limitations to increase precision and processing speed [16]. Wei Jin and Hung Hay Ho developed a robust machine-learning approach for web opinion mining, addressing server problems, and self-learning new vocabulary. Their model applied a novel bootstrapping method for managing large training sets, demonstrating its effectiveness in web opinion mining and product review extraction [17].

Guang Qiu, Bing Liu, Jiajun Bu, and Chun Chen focused on key tasks in opinion mining : lexicon enlargement and targeting extraction. The propagation method was developed for the iterative yielding of opinion words and targets based on a small seed lexicon.Consequently, a unique allocation approach for new opinion words and preconditions for noisy targets is proposed in this paper. By utilizing large-scale datasets to improve lexicon adaptability and sentiment recognition accuracy across political opinions, our method builds on existing work by applying targeted sentiment extraction particularly for political situations [18]. Mohammad Taher Pilehvar and Jose Camacho-Collados explore the impact of text preparation techniques on the accuracy of NLP. This study shows how important preprocessing consistency is for better results in text classification and sentiment analysis dataset [19]. Considering big Twitter data and based on co-occurrence statistics, Zhao et al. offered an unsupervised word embedding approach that exploits contextual semantic relationships in a hidden form. To predict the sentiment for various titles, they used a deep convolutional neural network using a sentiment feature set of n-grams and word polarity scores extracted from tweets. Their method's efficacy in sentiment analysis and word embedding on social media content was demonstrated when they compared it to a basic baseline model for five Twitter datasets. In order to improve model resilience over a range of language styles, our work goes one step further and adapts these techniques to the formal and informal combination frequently encountered in political discourse [20].

3. Methodology

3.1. Data Collection

This study used a dataset of 14,223 comments collected from the comment sections of political videos on popular YouTube channels such as Al Jazeera English, WION , TRT World , and others . YouTube was selected as a data source due to its broad reach and active user engagement on politically charged topics , allowing for a comprehensive analysis of public sentiment. The selected channels represent a variety of political perspectives from different global regions , providing a multitude of viewpoints essential to gain a detailed understanding of political sentiment.

This data ranges from different political aspects, comments, and conveys all sorts of opinions, which becomes imperative in analyzing political sentiment in a balanced manner. Comments were collected using methods designed to ensure reliability and minimize bias , enhancing the generalizability of the study's findings . Predominantly in English, the comments include a mixture of formal and informal language , incorporating slang , regional phrases , and colloquial expressions common on social media platforms. This further adds diversity to the dataset in terms of linguistics, making the communications on social media very realistic, which actually acts in favor of the models trained on sentiment analysis.

Reflecting the backgrounds of users from different regions, the dataset captures a wide range of perspectives, particularly valuable for studying political opinions and shifts in public sentiment across global audiences. The dataset’s linguistic and cultural diversity, along with the informal style of expression typical of social media, underscores its relevance in analyzing the dynamics of political discourse, especially in politically diverse regions like South Asia and the Middle East. This dataset serves as

a rich resource for understanding public sentiment on political issues, offering insights into the evolving landscape of political engagement on social media.
The process discusses below figure 1.

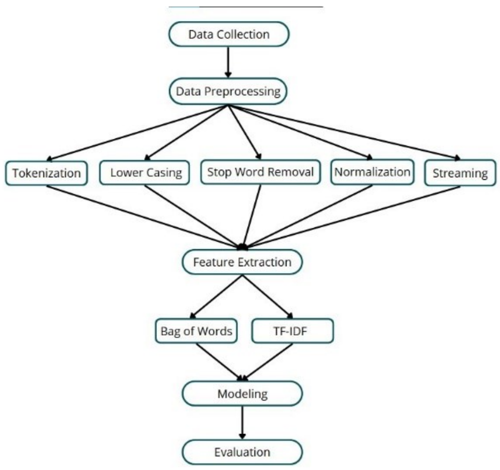

Figure 1. Methodology Process.

3.2 Data Preprocessing

Text preparation is an important phase in natural language processing (NLP), and methods like Tokenization, stop word removal, and lemmatization are essential. Preprocessing options have an important impact on text mining; suggestions change depending on the kind of text mining, the research objective, and the dataset features [21]. Preprocessing methods like stemming, Tokenization, and stop-word removal are essential for increasing the effectiveness of text classification [22]. It is contentious, however, whether preprocessing is necessary for categorizing casual texts; others contend that it might not provide suitable word embeddings as well as DNN architectures are used [23].

3.2.1 Tokenization

Tokenization is the method of dividing a large amount of text into smaller units known as tokens. Tokens can encompass words and characters. Tokens are the basis for natural language, a common method of processing raw text that focuses on the token level. Tokenization is a preprocessing step before applying natural language processing models to text input. The tokens are used to build a lexicon. This repository represents the collection of distinct tokens in the dataset [11]. It can be generated by considering each distinct token in the dataset or by selecting the top K most commonly occurring terms. Tokenization is an important stage in text processing, particularly for languages with complicated morphologies like Arabic, Thai, and Turkish [24].It significantly influences language systems' efficiency, with vocabulary size and tokenizer choice being crucial factors [25].Tokenization is a key idea in deep learning, and more theoretical and practical research is required to comprehend its use [26].

3.2.2 Lower Casing

Text lowercasing is the first step in data preprocessing to analyze political sentiments on social media. Such a methodology contributes to uniform sentiment lexicon mapping and prevents mistaking for a case-sensitive term. Hacohen-Kerner (2020) experiment has shown that lowercasing text could lead to better text classification precision [27].

Nevertheless, we should remember this nuance. Woo (2020) argues that a combination of lowercasing and lemmatization is appropriate for data preprocessing [28]. However, the social media texts usually contain non-formal language and slang words. For sentiment analysis, lowercasing alone could be more than enough even for larger datasets, as proven by the work done by Rosid (2020) regarding the Bahasa Indonesia documents. After all, whether to apply lowercase or not is subjective to the research question and the distinctive properties of your social media data [29].

3.2.3 Stop Word Removal

The stop words that don't affect the meaning of the comments are removed (for example and, or, still etc.). we have used the NLTK package for this reason, which checks our stop word removal [30]. The importance of stop-word removal in text processing is currently up for dispute, however, some research indicates that performance may be greatly enhanced by using stop-words tailored to a certain domain. This is especially important in software engineering, as it has been demonstrated that using domain-specific stop words performs better than using universal stop lists [23]. It is commonly known that stop word removal reduces corpus size and improves efficiency and accuracy in natural language processing applications like text classification and information recovery. But technological language processing additionally highlights the need for domain-specific stop words, as the conventional lists of stop words might not be enough in this context [31].

3.2.4 Normalization

The stop words that do not affect the meaning of the comments are removed (for example, and, or, still, etc.). We have used the NLTK natural language processing package for this reason, which checks our stop word removal [30]. The importance of stop-word removal in text processing is currently up for dispute. However, some research indicates that performance may be greatly enhanced using stop words tailored to a certain domain. This is especially important in software engineering, as it has been demonstrated that using domain-specific stop words performs better than using universal stop lists [23]. It is commonly known that stop word removal reduces corpus size and improves efficiency and accuracy in natural language processing applications like text classification and information recovery. However, technological language processing additionally highlights the need for domain-specific stop words, as more than the conventional lists of stop words might be needed in this context [31].

3.2.5 Streaming

A variety of research papers concerning text preprocessing in streaming applications were analyzed. Saloon (2021) utilized an NLP pipeline in real-time with imperceptible latency, operating on Apache Storm and Apache Kafka. In her paper of 2021, Khairunnisa investigated the application of sentiment analysis on Twitter data, which displayed remarkable results using cleaning, stemming, and normalization techniques [32]. Cabrera (2021) focused on overcoming the obstacles of streaming speech recognition and improving the proficiency of the RNNT model via language model fusion [33]. In the work of Chen (2022), the author introduced the consistency-based pretraining method, which regularized the shared speech and text representations by introducing consistency between real and synthetic speech [34]. These studies together promote reprocessing the text for streaming apps, which has results on NLP, sentiment analysis, speech recognition, etc.

One of the most important concepts for Porter stemmer is the concept of m in the equation 1:

$$[C][VC]^{m}[V] \qquad (1)$$

3.3 Feature Extraction

3.3.1 Bag of Words

A few studies have investigated the applicability of the BoW model in sentiment analysis. Barve (2022) proposed an attitude-based evolving BoW method for detecting misinformation by updating features if internet content evolved. Tabassum (2021) utilized BoW with Urdu and English tweets and reached a high accuracy degree based on machine learning techniques [35]. Mehanna (2021) proposed a semantic understanding approach via tagged BoW, which is efficient in sentiment analysis, especially for short messages like tweets. As in the study of Es-Sabery (2022) [36], the choice of feature extractors, such as BoW, played a major role in opinion mining, where it was found that feature selection significantly improved the model's performance. These research papers individually prove the applicability and efficacy of BoW for sentiment analysis done on different types of textual data sets.

3.3.2 TF-IDF

TF-IDF in sentiment analysis has made progress due to different methods. Jalilifard (2021) proposed STF-IDF, a semantic method that substantially improved TF-IDF's performance in grade-level informal documents [37]. Kalaivani (2023) detailed applying syntactic dependency features, including a supervised feature weighting scheme called delta TF-IDF, to enhance document-level sentiment analysis. Rahman (2020) studied the influence of TF-IDF and n-grams on the sentiment classification process, and according to his results, they performed so differently because of data size (Rahman, 2020). Arthamevia (2021) has incorporated TF-IDF in their aspect-based sentiment analysis for beauty product reviews, and by combining it with word unigram and SVM algorithm, they have attained high accuracy. With the concerted effort of these studies, the TF-IDF method has proven to be a very effective approach to sentiment classification in association with semantic and syntactic features, n-gram, and algorithms.

The TF is the approach used to measure term weights in a document, as seen in Eq. 2 . The IDF shows the number of words in more than one document and determines whether the word is a term (Stop Words).

TF-IDF score form term I in document j=TF(i,j)*IDF(i)

$$TF(i,j) = \frac{Term\ i\ frequncy\ in\ document\ j}{Total\ words\ in\ document\ j} \qquad (2)$$

For this purpose, the absolute value of the logarithm of the number of documents that contain the term must be divided by the number of documents, as shown in Eq 3.

$$IDF(i) = log\ log\ (\frac{Total\ documents}{documents\ with\ term\ i}) \qquad (3)$$

$$t = Term, j = Document$$

3.4 Modeling

We selected Logistic Regression, Linear SVC , Random Forest, Decision Tree , and Naive Bayes for their strengths in handling text-based sentiment classification. Logistic Regression and Linear SVC were chosen for accuracy in binary classification , Random Forest for its robust ensemble approach , Decision Tree for interpretability , and Naive Bayes as a quick baseline. Hyperparameters were fine-tuned via grid search , optimizing key parameters like C for SVC and n_estimators for Random Forest , to ensure each model achieved the best balance of accuracy and generalizability in analyzing political sentiment.

3.4.1 Logistic regression

The algorithm Logistic Regression is a probabilistic classification model for sentiment analysis. It predicts the likelihood of a sentiment-specific class according to input data features. It can enable sentiment analysis. We use the natural language processing (NLP) Python library for text processing and Tokenization. The algorithm calculates coefficients for each characteristic and builds the sigmoid function to produce binary sentiment expectations. Assessment metrics such as accuracy, Precision, and recall check its efficiency [38].

3.4.2 Linear SVC

The Linear SVC method is used as a sentiment analyzer. It optimizes a hyperplane to segregate sentiment categories in a high-dimensional space. The purpose of it is to increase the border between classes, consequently leading to higher classification precision. We employ a model-tuning process to optimize the hyperparameters, such as the strength of regularization, with a grid search algorithm [39].

We have used a supervised learning approach for sentiment classification. A machine learning model was trained on a labelled dataset containing social media, specifically YouTube, with corresponding sentiment labels (positive, negative, or neutral). Common algorithms for sentiment analysis include: Support Vector Machines (SVMs): Efficiently classify data points into separate categories. Model selection and hyperparameter tuning were performed to optimize performance on the training data. Evaluation metrics like accuracy, Precision, recall, and F1-score were used to assess the model's effectiveness.

3.4.3 Random Forest

The Random Forest technique is a hybrid learning algorithm for sentiment analysis that combines several independent decision trees. It builds up different decision trees and uses the combinations of their predictions to increase accuracy and prevent overfitting. each tree trains on a different subset of features and data samples. The hyperparameter tuning works to optimize tree depth and number of trees for better performance [40].

3.4.4 Decision tree classification

The methodology utilizes Decision Tree Classification for sentiment analysis. It constructs a tree-like structure to classify sentiment based on input features. The tree splits data into subsets using feature thresholds to minimize entropy or maximize information gain at each node. Pruning techniques prevent overfitting, enhancing generalization performance [41].

3.4.5 Naive Bayes

The methodology uses the approach of Naive Bayes for sentiment analysis. It uses Bayes' theorem to compute the probability of sentiment classes based on input features. The "naive" assumption of feature independence reduces computation complexity. Laplace smoothing deals with zero probabilities, thus improving robustness. Evaluation metrics such as accuracy and F1 score measure its performance [42].

3.5 Evaluation

In a research project such as political sentiment analysis on social media, evaluation metrics play a crucial role in evaluating the effectiveness of algorithms. Figure 2 shows the evaluation metrics predicated model vs true model heatmap.

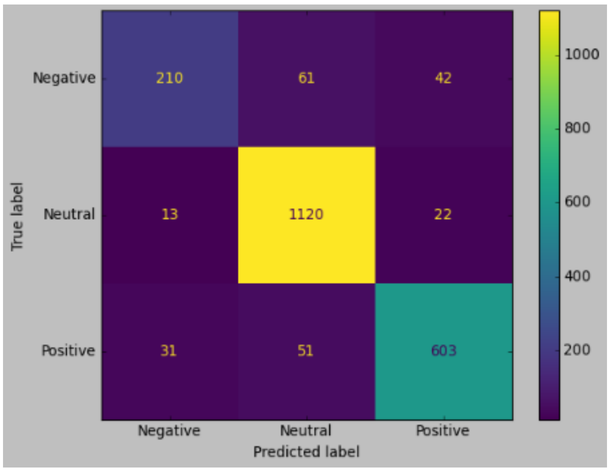

Figure 2. Predicted Model Vs True Model Heatmap.

The assessment metrics we suggested are summarized below.

3.5.1 Accuracy

This score assesses the accuracy of the model's predictions overall. The percentage of accurately predicted occurrences in all instances is computed. Although accuracy is crucial, unbalanced datasets with a predominance of one class over the others may not be the greatest fit for this statistic [43].

3.5.2 Precision

The ratio of accurately predicted positive observations to the total number of expected positives is known as Precision. It focuses on how accurate positive forecasts are. A high precision means that there are fewer false positives in the model [44].

3.5.3 Recall

The ratio of accurately predicted positive feedback to all observations made during the class is calculated as recall. It assesses how well the model can account for every good example. A high recall suggests that there are fewer false negatives in the model [45].

3.5.4 F1-score

The harmonic mean of recall and accuracy is known as the F1-score. It offers a harmony between recall and accuracy. This measure can be helpful when attempting to balance false positives and false negatives. When a dataset is unbalanced, the F1-score is frequently employed.

4. Result and Discussion

To assess the accuracy of our sentiment analysis models, various algorithms were evaluated, including Logistic Regression, Linear SVC, Random Forest, Decision Tree Classification, and Naive Bayes . Our results demonstrated that Linear SVC achieved the highest accuracy at 91.18% , with Logistic Regression following closely at 90%. Both models also scored well on precision and recall , indicating robust performance in accurately classifying political sentiment in social media posts . This aligns with findings in other studies where Linear SVC and Logistic Regression are noted for their effectiveness in sentiment classification, especially in high-dimensional text data.

Table 1. Performance Metrics of Sentiment Analysis Models Using Various Algorithms.

Algorithm	Accuracy	Precision	Recall	F1-score
Logistic Regression	90%	88%	84%	86%
Linear SVC	**91.18%**	**90%**	**86%**	**88%**
Random Forest	82.49%	85%	69%	72%
Decision Tree	89%	86%	84%	85%
Naive Bayes	68.56%	62%	63%	62%

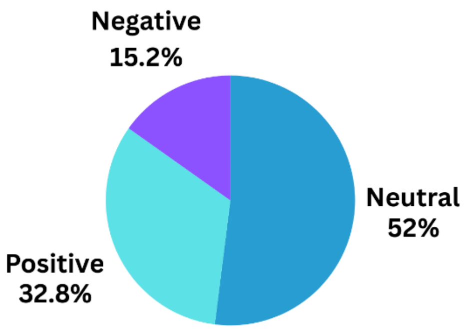

Figure 1. Comments Classification Pie Chart.

Comparatively, the Naive Bayes had considerably lower results in both accuracy and other performance metrics, which agrees with findings suggesting that Naive Bayes may be relatively unsuitable for complex, informal text data such as those seen on social media. The overall performance of Random Forest was also poor, given an accuracy of 82.49%, since the recall score was lower, which signifies a problem in recognizing rarer sentiments. Other studies also reported limitations of Random Forest when dealing with imbalanced datasets, and that is what we found. On the other hand, the Decision Tree Classification achieved excellent accuracy of 89% and had balanced precision and recall, which might make it a suitable choice as a simple yet interpretable model for sentiment analysis; this is in line with its established value in similar text classification tasks.

In summary, our study shows that evaluating Linear SVC and Logistic Regression for political sentiment analysis on platforms like YouTube compares favorably, or at worst marginally so, to benchmarks from previous related work, suggesting applicability to the problem. These findings suggest that a balance of accuracy, precision, recall, and efficiency should guide algorithm selection in political sentiment analysis on social media. Looking into this further, future research could refine these models, apply ensemble techniques, handle issues of data imbalance and feature extraction to learn better sentiment prediction performance in similar applications.

5. Conclusion

In conclusion, we have explained the space of political sentiment analysis through the use of natural language processing (NLP) methods in social media platforms, particularly focusing on YouTube comments related to political discussions. The main aim was to identify and study the sentiments emerging in the comments to know the dynamics of public opinion. The research process starts with an extensive review of literature that emphasizes the significance of sentiment analysis in understanding public sentiments, political trends, and sentiment dynamics. Numerous approaches and algorithms were used, including machine learning techniques like logistic regression, linear SVC, random forest, decision tree classification, and naive Bayes. After several thorough experiments and evaluations, the Linear SVC and the Logistic Regression algorithms had the most accuracy, precision, and recall. These models effectively and correctly categorize the sentiments of social media posts related to political communication. However, it is essential to point out that each algorithm has its strong sides and drawbacks, and the algorithm selection is based on the researchers' objectives, priorities, and computational efficiency. Difficulties like data bias and feature extraction optimization were also found. That is the place for improvements of sentimental analysis algorithms in the future. Overall, this research has added to the growing knowledge about political discourse analysis on social media. NLP techniques and machine learning algorithms are crucial in understanding public opinions, political discourse and trends for sentiments in modern society.

References

[1] R. S. Jagdale, V. S. Shirsat, and S. N. Deshmukh, “Sentiment analysis on product reviews using machine learning techniques,” in Cognitive Informatics and Soft Computing: Proceeding of CISC 2017, Springer, 2019, pp. 639–647.

[2] A. Ceron, L. Curini, S. M. Iacus, and G. Porro, “Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens’ political preferences with an application to Italy and France,” New media Soc., vol. 16, no. 2, pp. 340–358, 2014.

[3] M. Hassan et al., “Sentiment analysis on Bangla conversation using machine learning approach,” Int. J. Electr. Comput. Eng., vol. 12, no. 5, pp. 5562–5572, 2022.

[4] S. Datta, “Political violence in Bangladesh: trends and causes,” Strateg. Anal., vol. 29, no. 3, pp. 427–447, 2005.

[5] B. Agarwal, N. Mittal, P. Bansal, and S. Garg, “Sentiment analysis using common-sense and context information,” Comput. Intell. Neurosci., vol. 2015, p. 30, 2015.

[6] U. T. Gursoy, D. Bulut, and C. Yigit, “Social media mining and sentiment analysis for brand management,” Glob. J. Emerg. Trends e-Business, Mark. Consum. Psychol., vol. 3, no. 1, pp. 497–551, 2017.

[7] M. V Mäntylä, D. Graziotin, and M. Kuutila, “The evolution of sentiment analysis—A review of research topics, venues, and top cited papers,” Comput. Sci. Rev., vol. 27, pp. 16–32, 2018.

[8] N. Mishra and C. K. Jha, “Classification of opinion mining techniques,” Int. J. Comput. Appl., vol. 56, no. 13, 2012.

[9] L. Mathew and V. R. Bindu, “A review of natural language processing techniques for sentiment analysis using pre-trained models,” in 2020 Fourth international conference on computing methodologies and communication (ICCMC), IEEE, 2020, pp. 340–345.

[10] D. Das, S. Guha, J. Brubaker, and B. Semaan, “The" Colonial Impulse" of Natural Language Processing: An Audit of Bengali Sentiment Analysis Tools and Their Identity-based Biases,” arXiv Prepr. arXiv2401.10535, 2024.

[11] K. M. Islam, M. S. Reza, and M. D. Yeaser, “Sentiment analysis using Natural Language Processing (NLP) & deep learning.” Brac University, 2021.

[12] M. Tusar and M. T. Islam, “A Comparative Study of Sentiment Analysis Using NLP and Different Machine Learning Techniques on US Airline Twitter Data. arXiv 2021,” arXiv Prepr. arXiv2110.00859.

[13] M. A. Shafin, M. M. Hasan, M. R. Alam, M. A. Mithu, A. U. Nur, and M. O. Faruk, “Product review sentiment analysis by using nlp and machine learning in bangla language,” in 2020 23rd International Conference on Computer and Information Technology (ICCIT), IEEE, 2020, pp. 1–5.

[14] N. N. Moon et al., “Natural language processing based advanced method of unnecessary video detection,” Int. J. Electr. Comput. Eng., vol. 11, no. 6, pp. 5411–5419, 2021.

[15] G. Gautam and D. Yadav, “Sentiment analysis of twitter data using machine learning approaches and semantic analysis,” in 2014 Seventh international conference on contemporary computing (IC3), IEEE, 2014, pp. 437–442.

[16] V. Goel, A. K. Gupta, and N. Kumar, “Sentiment analysis of multilingual twitter data using natural language processing,” in 2018 8th International Conference on Communication Systems and Network Technologies (CSNT), IEEE, 2018, pp. 208–212.

[17] W. Jin, H. H. Ho, and R. K. Srihari, “A novel lexicalized HMM-based learning framework for web opinion mining,” in Proceedings of the 26th annual international conference on machine learning, Citeseer, 2009.

[18] G. Qiu, B. Liu, J. Bu, and C. Chen, “Opinion word expansion and target extraction through double propagation,” Comput. Linguist., vol. 37, no. 1, pp. 9–27, 2011.

[19] J. Camacho-Collados and M. T. Pilehvar, “On the role of text preprocessing in neural network architectures: An evaluation study on text categorization and sentiment analysis,” arXiv Prepr. arXiv1707.01780, 2017.

[20] Z. Jianqiang, G. Xiaolin, and Z. Xuejun, “Deep convolution neural networks for twitter sentiment analysis,” IEEE access, vol. 6, pp. 23253–23260, 2018.

[21] L. Hickman, S. Thapa, L. Tay, M. Cao, and P. Srinivasan, “Text preprocessing for text mining in organizational research: Review and recommendations,” Organ. Res. Methods, vol. 25, no. 1, pp. 114–146, 2022.

[22] S. Malik, S. A. Sani, A. Baqir, U. Ahmad, and F. ul Mustafa, “Preprocessing Techniques in Text Categorization: A Survey,” in Intelligent Technologies and Applications: Second International Conference, INTAP 2019, Bahawalpur, Pakistan, November 6–8, 2019, Revised Selected Papers 2, Springer, 2020, pp. 502–509.

[23] A. Kurniasih and L. P. Manik, “On the role of text preprocessing in BERT embedding-based DNNs for classifying informal texts,” Neuron, vol. 1024, no. 512, p. 256, 2022.

[24] P. Prakrankamanant and E. Chuangsuwanich, “Tokenization-based data augmentation for text classification,” in 2022 19th International Joint Conference on Computer Science and Software Engineering (JCSSE), IEEE, 2022, pp. 1–6.

[25] C. Toraman, E. H. Yilmaz, F. Şahinuç, and O. Ozcelik, “Impact of tokenization on language models: An analysis for turkish,” ACM Trans. Asian Low-Resource Lang. Inf. Process., vol. 22, no. 4, pp. 1–21, 2023.

[26] R. Friedman, “Tokenization in the Theory of Knowledge,” Encyclopedia, vol. 3, no. 1, pp. 380–386, 2023.

[27] Y. HaCohen-Kerner, D. Miller, and Y. Yigal, “The influence of preprocessing on text classification using a bag-of-words representation,” PLoS One, vol. 15, no. 5, p. e0232525, 2020.

[28] H. Woo, J. Kim, and W. Lee, “Validation of text data preprocessing using a neural network model,” Math. Probl. Eng., vol. 2020, pp. 1–9, 2020.

[29] M. A. Rosid, A. S. Fitrani, I. R. I. Astutik, N. I. Mulloh, and H. A. Gozali, “Improving text preprocessing for student complaint document classification using sastrawi,” in IOP Conference Series: Materials Science and Engineering, IOP Publishing, 2020, p. 12017.

[30] B. Gupta, M. Negi, K. Vishwakarma, G. Rawat, P. Badhani, and B. Tech, “Study of Twitter sentiment analysis using machine learning algorithms on Python,” Int. J. Comput. Appl., vol. 165, no. 9, pp. 29–34, 2017.

[31] S. Sarica and J. Luo, “Stopwords in technical language processing,” PLoS One, vol. 16, no. 8, p. e0254937, 2021.

[32] M. Arshi Saloot and D. Nghia Pham, “Real-time Text Stream Processing: A Dynamic and Distributed NLP Pipeline,” in 2021 International Symposium on Electrical, Electronics and Information Engineering, 2021, pp. 575–584.

[33] R. Cabrera, X. Liu, M. Ghodsi, Z. Matteson, E. Weinstein, and A. Kannan, “Language model fusion for streaming end to end speech recognition,” arXiv Prepr. arXiv2104.04487, 2021.

[34] Z. Chen, Y. Zhang, A. Rosenberg, B. Ramabhadran, P. Moreno, and G. Wang, “Tts4pretrain 2.0: Advancing the use of text and speech in ASR pretraining with consistency and contrastive losses,” in ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2022, pp. 7677–7681.

[35] N. Tabassum et al., “Semantic analysis of Urdu english tweets empowered by machine learning,” Intell. Autom. Soft Comput., vol. 30, no. 1, pp. 175–186, 2021.

[36] F. Es-sabery, K. Es-sabery, H. Garmani, J. Qadir, and A. Hair, “Evaluation of different extractors of features at the level of sentiment analysis,” INFOCOMMUNICATIONS J. A Publ. Sci. Assoc. INFOCOMMUNICATIONS, vol. 14, no. 2, pp. 85–96, 2022.

[37] A. Jalilifard, V. F. Caridá, A. F. Mansano, R. S. Cristo, and F. P. C. da Fonseca, “Semantic sensitive TF-IDF to determine word relevance in documents,” in Advances in Computing and Network Communications: Proceedings of CoCoNet 2020, Volume 2, Springer, 2021, pp. 327–337.

[38] M. Kwemou, M.-L. Taupin, and A.-S. Tocquet, “Model selection in logistic regression,” arXiv Prepr. arXiv1508.07537, 2015.

[39] E. Elgeldawi, A. Sayed, A. R. Galal, and A. M. Zaki, “Hyperparameter tuning for machine learning algorithms used for arabic sentiment analysis,” in Informatics, MDPI, 2021, p. 79.

[40] D. Makariou, P. Barrieu, and Y. Chen, “A random forest based approach for predicting spreads in the primary catastrophe bond market,” Insur. Math. Econ., vol. 101, pp. 140–162, 2021.
[41] Y.-Y. Song and L. U. Ying, “Decision tree methods: applications for classification and prediction,” Shanghai Arch. psychiatry, vol. 27, no. 2, p. 130, 2015.
[42] S. Taheri and M. Mammadov, “Learning the naive Bayes classifier with optimization models,” Int. J. Appl. Math. Comput. Sci., vol. 23, no. 4, pp. 787–795, 2013.
[43] D. Probst, “Aiming beyond slight increases in accuracy,” Nat. Rev. Chem., vol. 7, no. 4, pp. 227–228, 2023.
[44] R. Yacouby and D. Axman, “Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models,” in Proceedings of the first workshop on evaluation and comparison of NLP systems, 2020, pp. 79–91.
[45] D. Park and S. Kim, “Probabilistic Precision and Recall Towards Reliable Evaluation of Generative Models,” in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 20099–20109.