

Article

Not peer-reviewed version

Twitter Sentiment Analysis of Lagos State 2023 Gubernatorial Election using BERT

[Ashiribo Senapon Wusu](#) , [Olusola Aanu Olabanjo](#) ^{*} , Rebecca Maulome Padonu , Manuel Mazzara

Posted Date: 20 March 2023

doi: 10.20944/preprints202303.0335.v1

Keywords: NLP; NLU; Twitter; Sentiment Analysis; Opinion Mining; Nigeria; Election; Machine Learning; BERT; LSTM; SVM



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Twitter Sentiment Analysis of Lagos State 2023 Gubernatorial Election using BERT

Ashiribo WUSU ¹, Olusola OLABANJO ^{2,3,*}, Rebecca PADONU ⁴ and Manuel MAZZARA ⁵

¹ Department of Mathematics, Lagos State University, Ojo, Lagos, Nigeria

² Department of Mathematics, Morgan State University, Maryland, USA

³ Department of Computer Science, Lagos State University, Ojo, Lagos, Nigeria

⁴ Department of Theatre Arts, Ahmadu Bello University, Zaria, Kaduna, Nigeria

⁵ Institute of Software Development and Engineering, Innopolis University, 420500 Innopolis, Russia

* Correspondence: olola57@morgan.edu

Abstract: Election outcomes have been predicted in the past with the help of various state-of-the-art language models. Sentiment analysis helps in establishing the opinions of the public about a particular subject, a popular experiment known as opinion mining. Twitter has grown in popularity and proven to be a key tool in mining people's sentiments concerning election and other trending subjects of interest. The outcome of the just concluded Presidential election in Nigeria shifts the focus on Lagos State governorship election. In this study, we propose a Bidirectional Encoder Representations from Transformers (BERT) model for the sentiment analysis of governorship election in Lagos State Nigeria using Twitter data. A total of 800,000 personal and public tweets were scraped from twitter concerning the three prominent contesting candidates using carefully selected search queries. The tweets were preprocessed to avoid noise and inconsistencies. The preprocessed tweets were passed into the pretrained and finetuned BERT model. The result was analyzed to establish the sentiments of the public about the candidates. The social networks of the candidates were also analyzed. The parameter-tuning yield different results with different learning rates (LR). Results showed that the learning rate at 1e-7 gave the best performance and that the smaller the learning rate, the higher the accuracy but the larger the epoch size, the higher the accuracy.

Keywords: NLP; NLU; Twitter; Sentiment Analysis; Opinion Mining; Nigeria; Election; Machine Learning; BERT; LSTM; SVM

I. Introduction

Sentiment analysis models using public political discourse have gained a wide popularity in the field of political methodology [1]; an approach for the study of quantitative and qualitative approaches used to understand politics and political systems [2]. Several studies have presented methods often combine statistics, mathematics and formal theory for the understanding of politics. Election results have usually been predicted in the past using analytical and statistical techniques where the methodology relies on surveys and qualitative methods, and analyzing political party manifestos while observing the trends in the mainstream media [3,4]. With an increased contest and opposition in the government, particularly in nations with multi-party system like Nigeria [5,6], it has become more difficult but interesting to predict elections. The history of election forecasting goes back in time with a number of reasons such as aiding politicians to adapt their strategy/ideas, helping scientists improve their forecasting models, assisting media to inform readers/viewers, and satisfying the curiosity of citizens among other things [7].

The vastness, veracity, velocity and variety (often known as the four Vs [8]) of the social media space and how greatly it has been established to affect elections and electoral systems motivated this study. The influence of social media on elections became noticeable first in the early year 2000 [6]. In his first presidential campaign, Barack Obama made use of social media to mobilize the public and win the 2008 election. Statistics revealed that 74% of internet users—or 55% of the adult population—looked for election news online during Obama's first campaign. Another famous example is when Beto O'Rourke came dangerously close to unseating Senator Ted Cruz in 2018. The 2018 Texas Senate race broke the record for the most money spent in a U.S. Senate election, according to the Center for Responsive Politics, spending \$93 million, a large portion of which was raised from and used for social media events and advertisements. The social media is, no doubt, a key tool to understanding and analyzing public political opinion and predicting election outcomes. Several studies proved that messages on Twitter validly mirror the political landscape off-line, therefore, it

can be used to predict election results. Overtime, results have shown that citizens hold political conversations on Twitter which accurately reflects political party affiliations and directly affects the election result [9].

Before and after the social media revolution, surveys have been used frequently to identify opinions prior to elections; however, this approach has been limited by the difficulty in constructing an appropriate sampling procedure, hence making it difficult to obtain a representative sample of political viewpoints. Social media platforms such as Twitter, Facebook, Reddit and Instagram have helped in some way to overcome the “systematically inappropriate” sampling procedure in survey studies and have been the prominent tools of political campaigns and activism during elections. The recent advancements in the fields of Natural Language Understanding (NLU) and Natural Language Processing (NLP) have improved the reliability of prediction models built for unstructured and unsupervised dataset. The social media has radically upended the traditional campaign norms and tactics in national elections vis-à-vis its volume, velocity, scope and tactics of use [10,11]. Studies also showed that even if it cannot be categorically said that social media singlehandedly elected Donald Trump, his social media campaign strategies changed the way social media will be used in elections in the future [11,12]. Specifically, users of Twitter commonly seek information to digest or disseminate. Being a unique platform where trends are easily formed around topics via generation of hashtags as against individual posts which is prominent across other social media outlets. It is therefore easier to track users’ moods and directions as regards selected topics. Advertently, in a study to examine the potential of Twittersphere, authors found that out of all sampled categories, political hashtags and therefore discussions were the most persistent on Twitter [13].

There has been a growing interest in the use of NLP and other artificial intelligence techniques to predict election results in recent years with social media datasets. These techniques include sentiment analysis and topic modeling in addition to more advanced models that incorporate deep learning and fundamental statistical techniques. Two important aspects of NLP are sentiment analysis and opinion mining, which aid in classifying and investigating the behavior and approach of social media users with regards to brands, events, companies, customer services as well as elections. There are algorithms for automating the process of extracting emotions from user’s posts by processing unstructured texts and preparing models that extract knowledge from it. In this study, we consider Twitter; one of the prominent social network sites, to analyze the upcoming 2023 Lagos State Gubernatorial election. Given the statistics, Twitter has an active monthly user of 316 million and, on the average, about 500 million tweets are posted daily. A review of related works in subsequent section shows that Twitter is one of the most powerful tools in political analysis. In the current fast-paced global network system, information spreads digitally between users and shapes the way these users feel about an event, thereby making it crucial to understand the thought polarity, emotions and sentiment of the masses.

This study has been conducted to understand opinions on 2023 Lagos State gubernatorial election using Twitter dataset. It aims to capture, process and evaluate the public opinion from three main perspectives: sentiment and timeline analysis of the personal accounts of contesting candidates; sentiment and general tweet analysis of the public on the three candidates and sentiment and general tweet analysis of the public on the Lagos State 2023 elections. The study would therefore concentrate on the following methodology:

- i. Identifying keywords, hashtags and accounts which wholistically explain the subject matter as well as capture the top three contestants for Lagos State 2023 gubernatorial elections.
- ii. Scraping the tweets through Twitter API using Python programming.
- iii. Data cleaning (removing white spaces, links, punctuations, stop words, tokenization, retweet).
- iv. Developing, pretraining and tuning a BERT model and training with existing annotated IMDB dataset.
- v. Using Natural Language Understanding techniques such as sentence polarity, topic modeling, entity extraction, word frequency, word cloud etc. to understand the personal profiles of the three strongest Lagos gubernatorial candidates.
- vi. Analyzing the result to predict the direction of Lagos State election to be conducted on March 18, 2023.

The study is further structured into sections: the first section discusses existing works related to our study; the next section which covers methodology explains in detail the technique, models construction and evaluation as well as other analyses done in this to achieve our main aim and objectives, the next section discusses the results obtained in the various experimental setups while, finally we discuss the results and concludes with our findings and direction for future work.

2. Study Premise

Lagos State happens to be the smallest of the thirty-six states in Nigeria, Federal Capital Territory inclusive, but with the highest urban population consisting 27.4% of the National population. Lagos is Africa’s leading New Partnership for Africa’s Development (NEPAD) city and world’s sixth megacity which attained its megacity status in 1995. It is essentially a Yoruba environment but with diverse culture and traditions and inhabited by the Yorubas, Aworis and the Ogus, with other pioneer immigrant settlers (Edos, Saros, Brazilians, Kannike/Tapa) collectively called Lagosians but also referred to as the Ekos. The megacity status of Lagos can be attributed to its sound economic base, strategic maritime location and socio-political importance which induced a high rate of migration to the State.

Politically, Lagos State consists of 20 Local Government Areas (LGAs) and 37 Local Council Development Areas (LCDAs). The prominent ruling political party in Lagos is the All Progressives Congress (APC) which is one of the three major ruling parties in Nigeria, alongside People’s Democratic Party (PDP) and Labour Party (LP). Unlike in the past where the Presidential election results were dominated by APC and PDP only, the Presidential election which held on the 25th February, 2023 was keenly contested by three political parties with the winning distribution shown in Figure 1. This is as rightly predicted by authors in previous study [14].

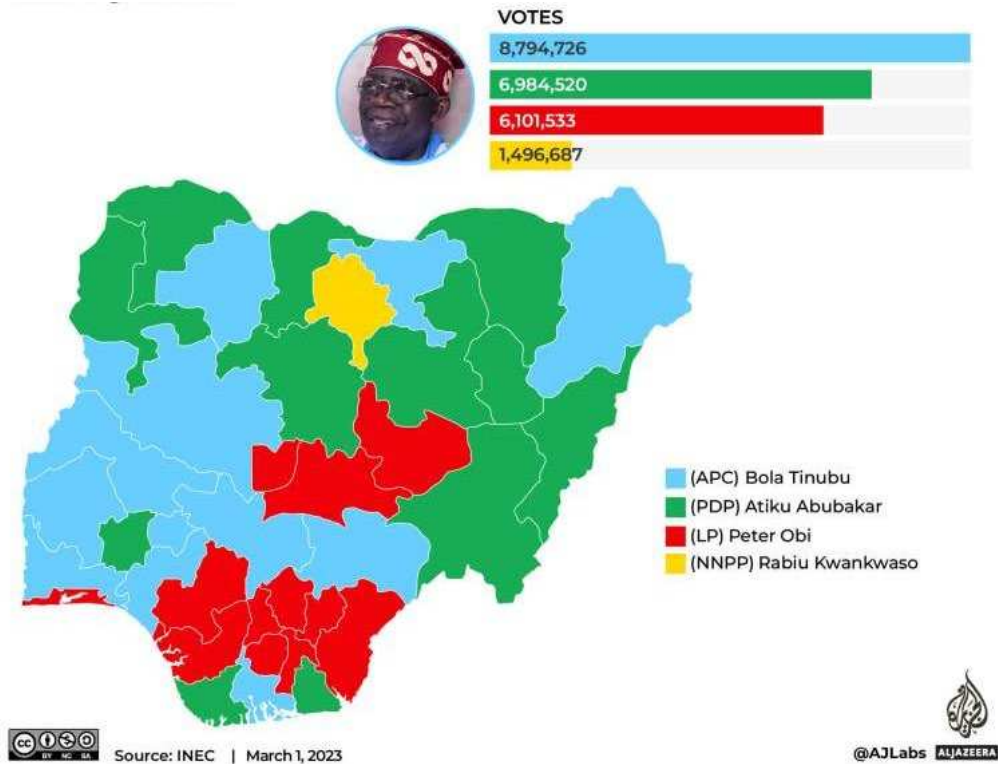


Figure 1. Geographical Distribution of Candidate’s Wins by States (image adopted from Aljazeera).

After the Independent National Electoral Commission (INEC)–the election conducting body–announced the results of Presidential election, there have been allegations of rigging, misconducts and misappropriation in the electoral process, improper conducts, unacceptable result transmission from polling booths, amongst others, thereby electrifying the Twitter space. This agitation, with the fact that the 2023 winner party APC did not win Lagos, brings a major paradigm shift from the status-quo of the APC party always winning Lagos State gubernatorial seat. This implies that the result of the forthcoming governorship election in Lagos State will be as interestingly dicey as the Presidential election; hence, the justification for this study.

3. Related Works

Sentiment analysis has been used to predict the opinions of the citizens on US election using Twitter data [15]. The authors used 17,000 tweets to train their model (Naïve Bayes) and the model achieved a less than 60% prediction accuracy by classifying the tweets into positive, negative, neutral and not-sure and this assisted in analyzing real time tweets

from the people which gave great insights about public opinions on each candidate. Another study [16] used Twitter dataset to analyze tweets to get international opinions concerning a protest that happened in India conducted by the farmers and about 20,000 tweets was scraped to analyze and categorize tweets into positive, negative and neutral. Bag of words and TF-IDF were used to conduct the analysis and Bag of words outperformed TF-IDF. Other authors [17,18] also established that Twitter can be used as an election indicator and has the ability to predict the favorite candidate of the people to emerge as the winner before the election is conducted. The people gave positive opinions about Donald Trump in almost all states in the United States prior to election. 1,000,000 tweets were collected from various users from different states and sentiment analysis was conducted. A study [19] used Twitter data to check and conduct how different countries affected by the Corona virus coped with the situations. Tweets posted in English were analyzed to give the opinion and emotion of the people concerning the pandemic in their respective countries and 50,000 tweets were used in the study.

Studies combined different features and used ensemble models to increase the detection of polarity of tweets for sentiment analysis and established that unsupervised and ensemble machine learning models outperformed other classical machine learning models in detecting opinions from text [20,21]. Authors have used Twitter data to examine political homophily in American Presidential Elections in 2016 where 4.6 million tweets were collected for analysis [22]. Authors [23] motivated by the prospect of a forecasting relationship between politics and market performance and the possibility of evaluating Twitter as a social media tool in a political context, concluded that that correlation does exist between public sentiment in regards to the local elections and the FTSE movements, in the form of correlation between MOOD and CLOSE. They agreed that sentiment analytics can be used to determine user's behavior and predict election results.

In some studies, authors used the ratio of positive message rate and negative message rate to predict the likely winner of a forthcoming election using twitter data and it showed that these opinions could be used to predict candidate that will emerge as the winner [24]. Twitter data has been used to establish that social media is not only used to express opinions, but it is also used to share ideas and opinions among other users. Authors used 100,000 tweets to predict German federal election in 2009 which could serve as a political indicator for the election [25] and another study [26] also predicted the outcome of German presidential election of 2021 using 58,000 tweets which they established that traditional machine learning methods like Naive Bayes performed less than transformer-based models like the bidirectional encoder from transformers (BERT).

Authors [27] performed social media sentiments about political parties to study and forecast Pakistan's general election. They used supervised machine learning algorithms to classify tweets into positive, negative, and neutral. The findings of their experiment shows that social media content can be a useful indication for identifying political behavior of various parties. In another study [28], 90154 tweets were analyzed and the results were compared with the actual election results, their model predicted the winning party accurately. Twitter is indeed used as a platform for political deliberation. The political preferences shown via basic tweets comes close to traditional election polls. Simply, the number of tweets collected per party reflects the votes they receive in the election. Therefore, it's concluded that Twitter can be considered a good indication of real-time political sentiment [9].

4. Materials and Methods

This study seeks to analyze tweets of the political discourse around the upcoming 2023 Gubernatorial election in Lagos State and its prominent contesting candidates and this section describes in detail the approach used in discovering the sentiments of people around this area of public interest. Figure 2 shows the workflow of the BERT framework designed for this study. Every tweet with a political content either contains a neutral, positive or negative sentiment for or against a party or candidate. The sentiments contained in tweets especially when it is specific to a candidate is not easy to compute with algorithms since emotion expression varies with the personality, region and cultural background of each person. Since this is an unstructured (unlabelled) dataset, sentiment analysis is often challenging because of the features, context and semantics peculiarity of each tweet.

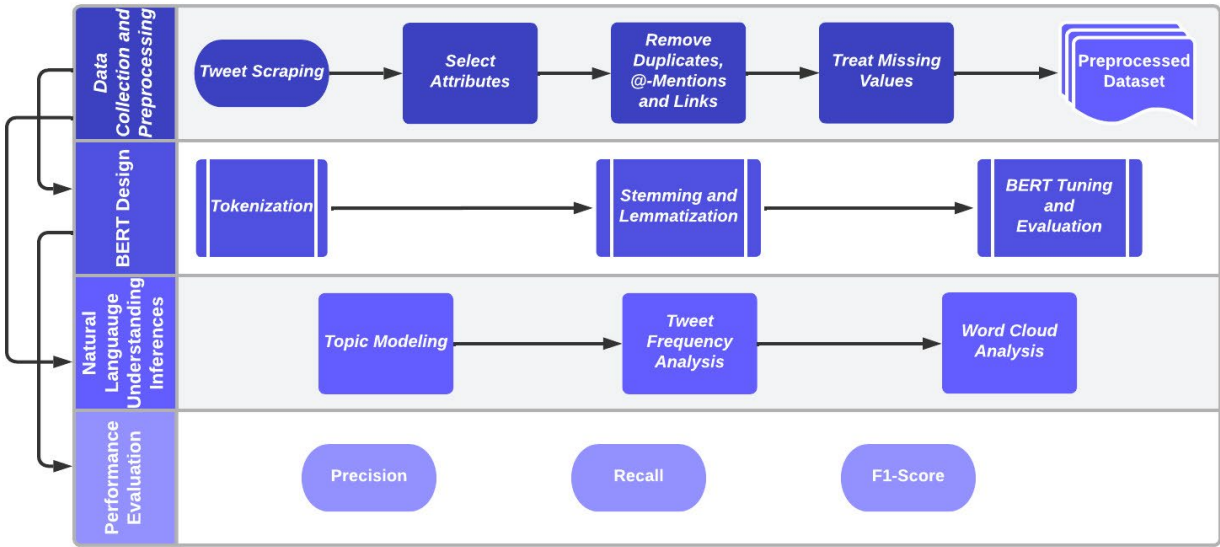


Figure 2. BERT Sentiment Analysis Framework.

The study framework has been stratified into four phases namely: the data collection and preprocessing phase, the BERT design phase, the Natural Language Understanding Inference (NLUI) phase and the performance evaluation phase.

Data Collection and Preprocessing Stage: This phase is concerned with the systematic extraction of relevant tweets from Twitter databases concerning Lagos State gubernatorial election. Two categories of dataset were extracted from Twitter using the Python snsrape library: personal tweets of three most prominent candidates in Lagos election race and the public tweets showing the interest of the potential voters in each of these candidates. After scraping, only the attributes of interest were retained for the study. The preprocessing (Algorithm 1) was done to remove noisy, irrelevant and inconsistent contents from the dataset.

Algorithm 1: Algorithm for Preprocessing our Twitter Dataset

Input: Twitter comments or text data
Output: Pre-processed text data
For each comment in the Twitter data file
Initialize temporary empty string processedTweet to store result of output.
1. Replace all URLs or https://links with the word 'URL' using regular expression methods and store the result in processedTweet.
2. Replace all '@username' with the word 'AT_USER' and store the result in processedTweet.
3. Filter All #Hashtags and RT from the comment and store the result in processedTweet.
4. Look for repetitions of two or more characters and replace with the character itself. Store result in processedTweet.
5. Filter all additional special characters (: n [] ; : { } - + () < > ? ! @ # % *,,) from the comment. Store result in processedTweet.

6. Remove the word 'URL' which was replaced in step 1 and store the result in processedTweet.

7. Remove the word 'AT_USER' which was replaced in step 1 and store the result in processedTweet.

Return processedTweet.

The output of this phase is a fully preprocessed dataset ready to be passed into the BERT model.

BERT Design Phase: This involves the tokenization of the tweets, stemming and lemmatization and the BERT fine-tuning and evaluation. Tokenization is a term which describes separating a corpus into smallest, syntactically meaningful, units. It is a fundamental step in modeling text data which aids in understanding the meaning behind the text by analyzing the sequence of the words. Porter stemmer was used to reduce the inflection towards their root forms. This was done by stripping the suffix to produce stems [29]. This is then passed into the BERT pipeline (Figure 3) for sentiment analysis.

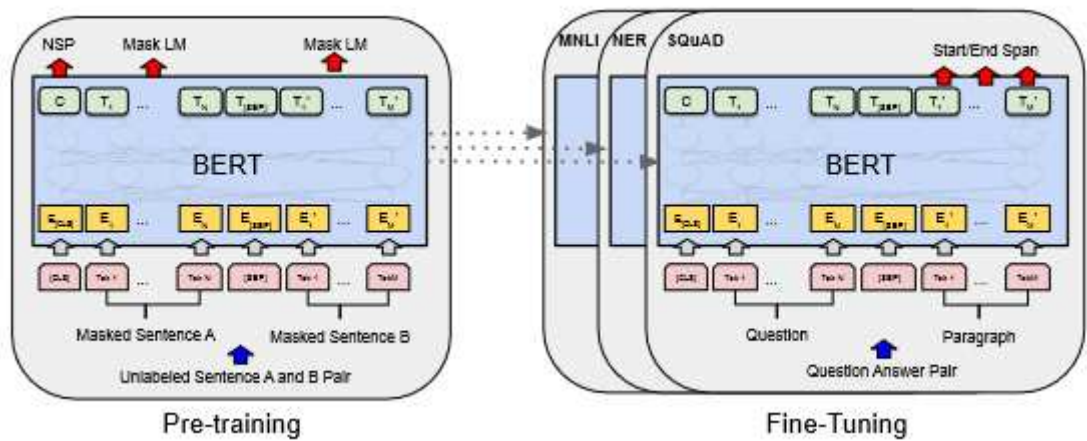


Figure 3. Overall pre-training and fine-tuning procedures for BERT [30].

The Bidirectional Encoder Representations from Transformers (BERT) model is a Google AI Language masterpiece which has been widely applied in a wide range of NLP tasks [30, 31]. The main strength of the BERT model is its application of bidirectional training of Transformer which is a popular attention model to language modeling. Earlier similar models looked at text sequences from either a left-to-right or a combined left-to-right and right-to-left training perspective. BERT models have demonstrated that bidirectionally trained language models can comprehend context and flow of language more deeply than single-direction language models [32]. Classification tasks such as sentiment analysis are done similarly to Next Sentence classification, by adding a classification layer on top of the Transformer output for each token. The Linear Support Vector Classifier was used to test the accuracy of our deep models. SVM has been one of the most robust prediction techniques which is based on a statistical learning framework [33]. Its workings and model development strategies have been broadly explained in some existing works [34–36].

The BERT pipeline as shown in Figure 3 has two main phases: the pre-training and the finetuning; the first concerned with the pretraining of unlabeled data while the second deals with modifying the learning parameters of the model. This is done until the optimum performance is achieved. BERT models use a multi-layer bidirectional transformer encoder with 6 layers in the encoder. For each word entered into the attention layer of the model, it is converted into Query (Q), Key (K) and Value (V) representing the actual word, keyword or meaning of the word and the value of the intent, purpose or polarity of the word respectively [37,38]. The finetuning and evaluation process deals basically with the selection of hyperparameters and compares experimentally the results of training and testing. The major hyperparameters which require tuning in a BERT model are the learning rate, epoch and the batch size.

Natural Language Understanding Phase: Three tasks are embedded in this phase to aid our understanding of the dataset being studied especially as it relates to the three contesting candidates under study. The time-series-frequency

analysis was used to identify the tweet pattern and interest growth of voters regarding the candidates. Topic modeling is a technique for unsupervised categorization of the twitter documents which helps to identify natural groups of words even when we are not certain what the outcome will be. In this study, the Latent Dirichlet Allocation (LDA), a particularly popular algorithm for achieving this task, was used [39,40]. A n -dimensional Dirichlet random variable θ takes values in the $(k-1)$ -simplex, that is, a k -vector θ lies in the $k-1$ simplex iff $\theta_i \geq 0, \sum_{i=1}^k \theta_i = 1$ and has the probability density function on this simplex as given in Equation (1)

$$p(\theta|\alpha) = \frac{\Gamma(\sum_{i=1}^k \alpha_i)}{\prod_{i=1}^k \Gamma(\alpha_i)} \theta_1^{\alpha_1-1} \dots \theta_k^{\alpha_k-1} \dots \dots \dots (1)$$

where α is a k -vector with component $\alpha_i > 0$ and $\Gamma(x)$ is the Gamma function. We used WordCloud, also called TagCloud to visually represent our Twitter corpus. Tags are tokens, the importance of which is represented with font size or color as a depiction of word significance and word co-occurrences. The size of each word in our WordCloud is given in Equation (2).

$$s_i = \left\lceil \frac{f_{max} \cdot (t_i - t_{min})}{t_{max} - t_{min}} \right\rceil \quad \forall t_i > t_{min} \text{ else } s_i = 1 \dots \dots \dots (2)$$

where s_i is display font size

f_{max} is the maximum font size

t_i is the count

t_{min} is the minimum count

t_{max} is the maximum count

Words in our WordCloud appear bigger the more often they are mentioned and are great for visualizing unstructured text and getting insights on trends and patterns [41,42].

Performance Evaluation: The sentiment models were assessed using recall, precision and F-measure as seen in Equations (3)–(5). Recall is defined by Equation (3). Precision is the proportion of the predicted positive cases that were correct. The precision can be calculated using Equation (4). The F-Measure score is the harmonic mean of the precision and recall. This evaluates the equivalency between the sensitivity (recall) and the precision (correctness) of the data. This gives us the interpretation of how the measure recall and precision values behave for the dataset. The F-measure can be calculated using Equation (5).

Let TP_+ = true positive for all positive tweets, TP_- = true positive for all negative tweets, TP_ϕ = true positive for all neutral tweets. Then:

$$TPR, Sensitivity, Recall = \frac{\sum TP_+}{\sum \text{All Actual Positives}} \dots \dots \dots (3)$$

$$Precision = \frac{TP}{TP + FP} \dots \dots \dots (4)$$

$$F - Measure = \frac{2 * Precision * Recall}{Precision + Recall} \dots \dots \dots (5)$$

5. Results

In this section, we present and discuss the results obtained in this study with their implications for the set goals. Results are recorded in terms of the dataset, processing, parameter tuning, inferences and performance evaluation of the BERT sentiment model.

a. Dataset

Our dataset was extracted from immediately after the announcement of the Presidential election results (March 1, 2023) till two days before the gubernatorial elections. There is a keen interest of Twitter users in the Lagos State election,

hence the vast number of tweets scraped for this study. Figure 4 shows the size of the dataset as it passed through each preprocessing tasks.



Figure 4. File size variation in Megabytes at each preprocessing task.

Figure 3 is a representation of the change in the size of dataset during each preprocessing task. The raw dataset was reduced from 302MB to 201MB after the preprocessing tasks were completed.

b. Personal Tweet Analysis

In this section, we examined the personal Twitter accounts (@jidesanwoolu, @GRVLagos and @ officialjandor) of each of the three Lagos Gubernatorial election candidates of interest from March 1, 2023 which happens to be the day attention shifted from the national election to the state election to March 17, 2023 (the eve of the state election). The first dataset contained the personal tweets of the candidates and Table 4 shows their tweet summarization, tweet frequency and impressions made by their personal tweets. The percentage total impression is given in Figure 6. Figure 5 shows the tweet counts of the respective candidates per day for the seventeen days study period. Results also show that Jandor has the least percentage total impressions (near zero), tweets and followers.

Table 1. Personal tweet summaries of the Governorship Candidates from March 1st – March 17th 2023 excluding retweets and replies.

	<i>Sanwoolu (APC)</i>	<i>GRVLagos(LP)</i>	<i>Jandor (PDP)</i>
Handles	@jidesanwoolu	@GRVLagos	@officialjandor
Total tweets	68	102	18
Retweets	40754	446963	987
Favorited	188907	1390358	3400
Replies	58634	49788	415
Quoted	233400	14537	101
Total followers	1908972	228521	13195

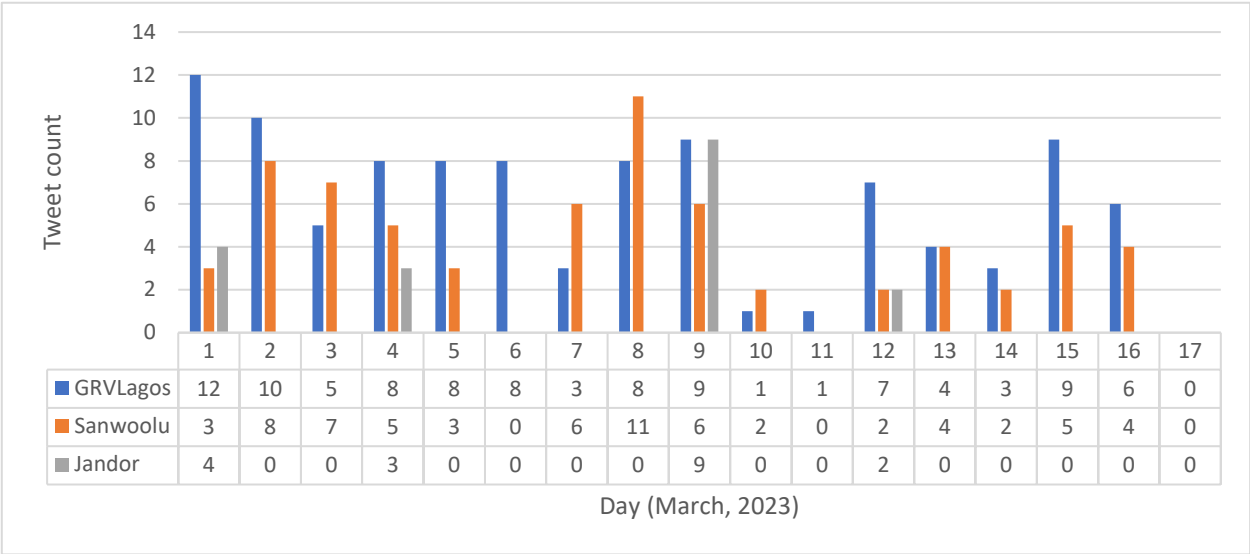


Figure 5. Daily tweet pattern of the three prominent Lagos State governorship candidates.

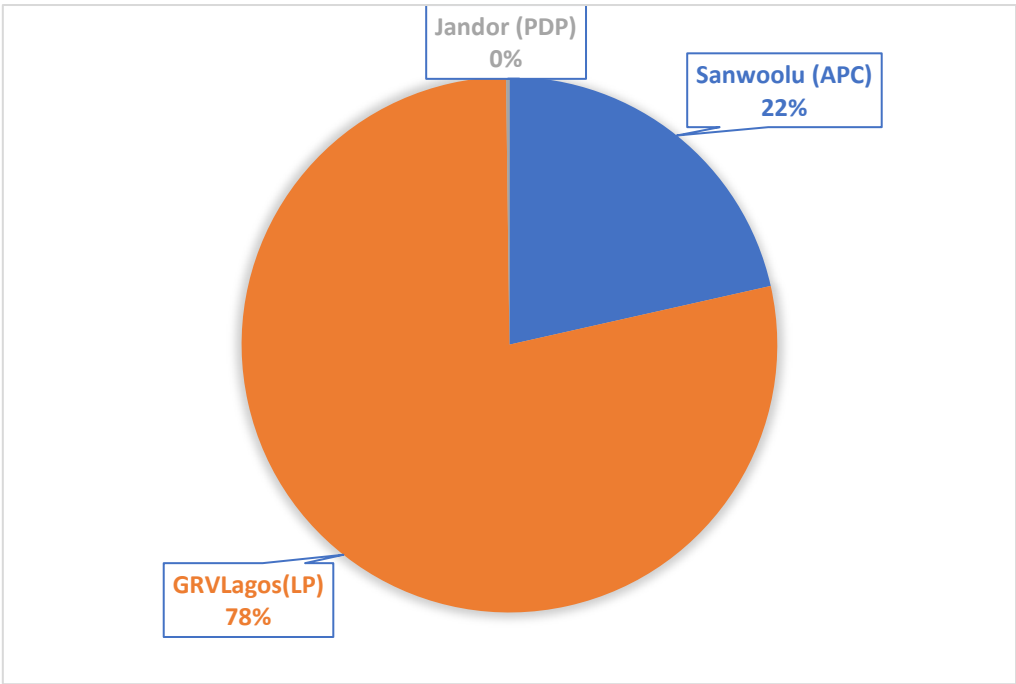


Figure 6. Total Impressions (Retweets + Replies + Quotes + Likes) of the three Lagos governorship candidates.

c. Social Network Analysis

Our social network analysis on each of the three candidates for the period covered in this study produced Figure 7. This was used understand the Twitter activities of the verified friends of the gubernatorial election candidates in terms of retweets and mentions. It shows that Sanwoolu maintained the strongest and most active verified network of Twitter friends relative to GRVLagos and Jandor. Jandor, on the other hand, has many friends who are dormant and inactive to his posts. Sanwoolu has prominently active verified friends including the President-elect, Bola Ahmed Tinubu, Babatunde Raji Fashola (minister for Works and Housing of Nigeria), etc.

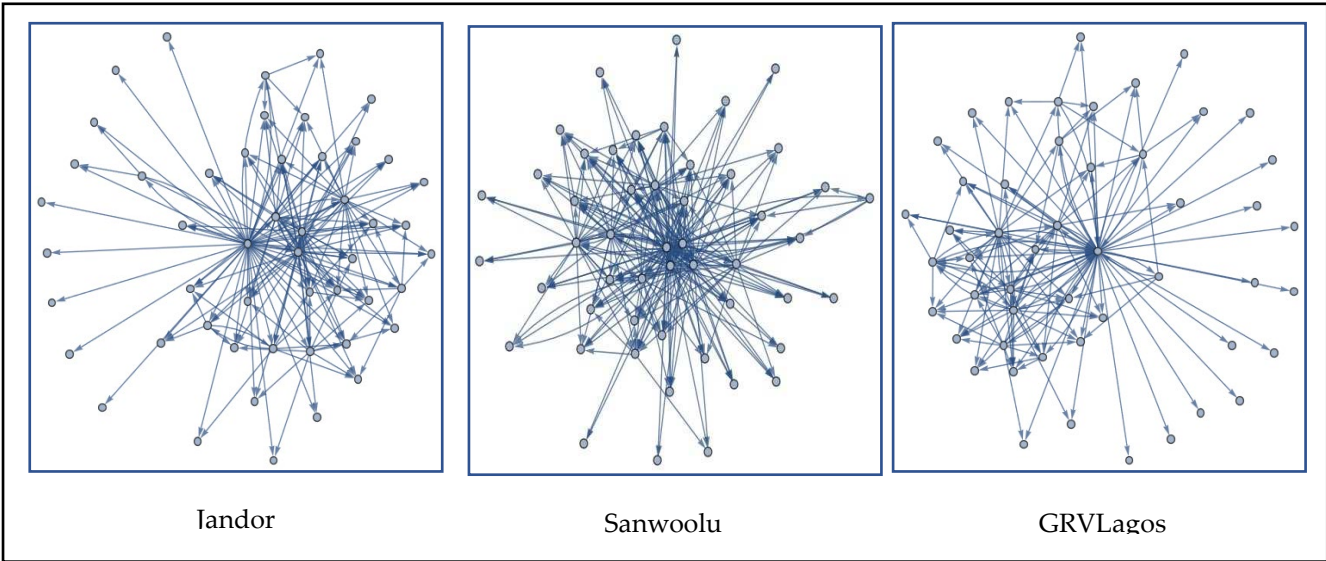


Figure 7. Social Network of the Lagos Governorship Candidates.

d. Public Tweet Analysis

The BERT model was applied to the public tweets of the three candidates to determine the sentiments of the public concerning them. Tweets were categorized into Positive, Neutral and Negative represented by 1, 0 and -1 respectively. Table 2 presents the public tweets as scraped from Twitter. It shows that GRVLagos of the Labour Party amassed most public tweets and impressions. Figure 9 shows the sentiments of the tweet concerning the candidates as determined by the pretrained BERT model. This figure also shows that Jandor made the least total public impressions so far in this gubernatorial election contest. Also GRVLagos has made the most wave on Twitter with the highest tweets and highest overall public positive sentiments as adjudged by our BERT model.

Table 2. Public tweet summaries of the Governorship Candidates from March 1st – March 17th 2023 excluding retweets and replies.

	<i>Sanwoolu (APC)</i>	<i>GRVLagos(LP)</i>	<i>Jandor (PDP)</i>
Query/Search Term	(APC OR Sanwoolu OR Jidesanwoolu OR Babajide Sanwoolu) AND (Lagos OR Governor)	(GRV OR GRVLagos OR LP OR Labour Party) AND (Lagos OR Governor)	(Jandor OR PDP) AND (Lagos OR Governor)
Total tweets	190188	356259	92569
Retweets	1881512	3066151	1080255
Favorited	5119776	8495579	3048653
Replies	463942	620994	301962

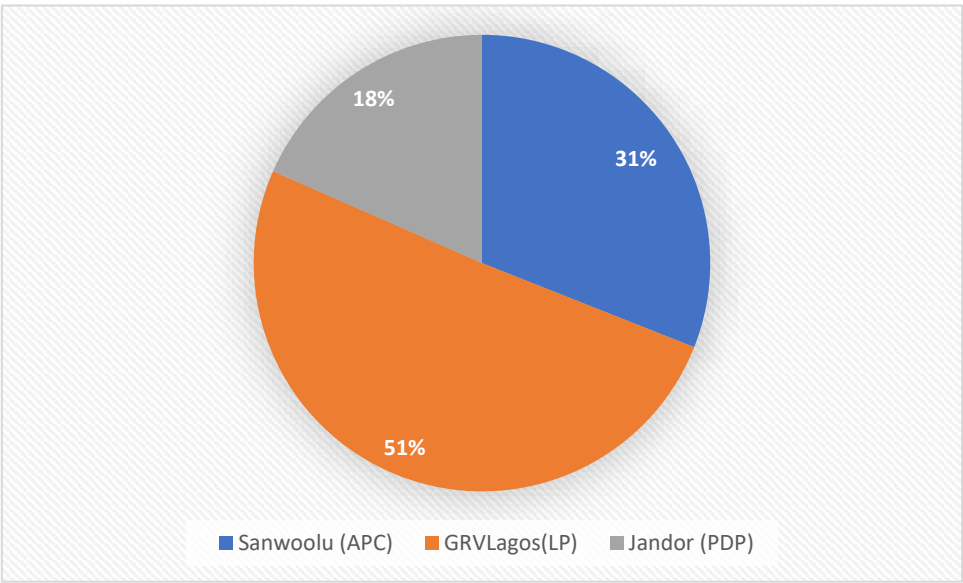


Figure 8. Total impressions from public tweets.

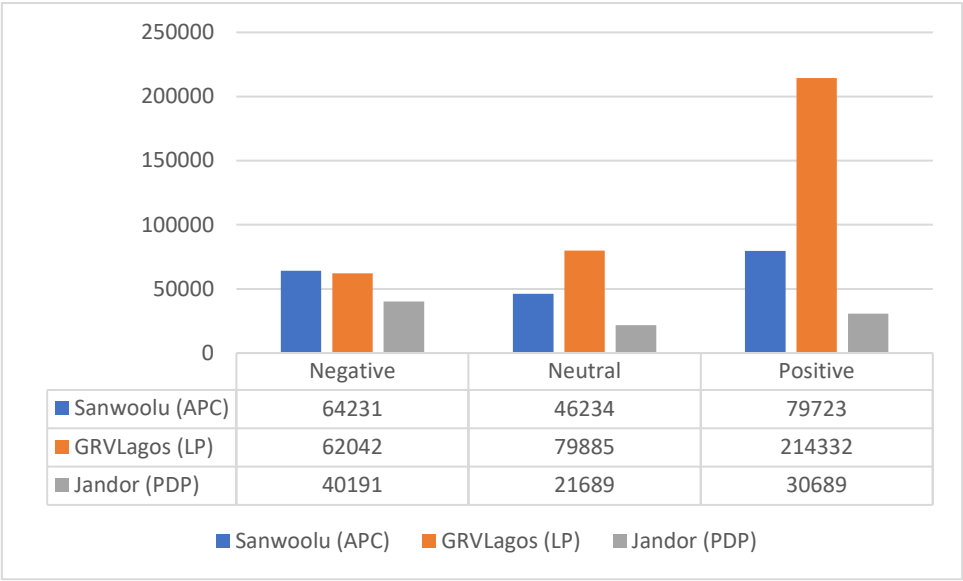


Figure 9. Summarization of Public Tweet Sentiments.

e. Parameter Tuning and Performance Analysis

The tuned parameters yield different results given the various experimental stages of parameter tuning with different learning rates (LR). The results of the precision, recall and f1-measure as reflected by the variations in the learning rate are explained in Figure 9. It also shows that the learning rate at 1e-7 gave the best performance. This shows that the smaller the learning rate, the higher the accuracy but the larger the epoch size, the higher the accuracy.

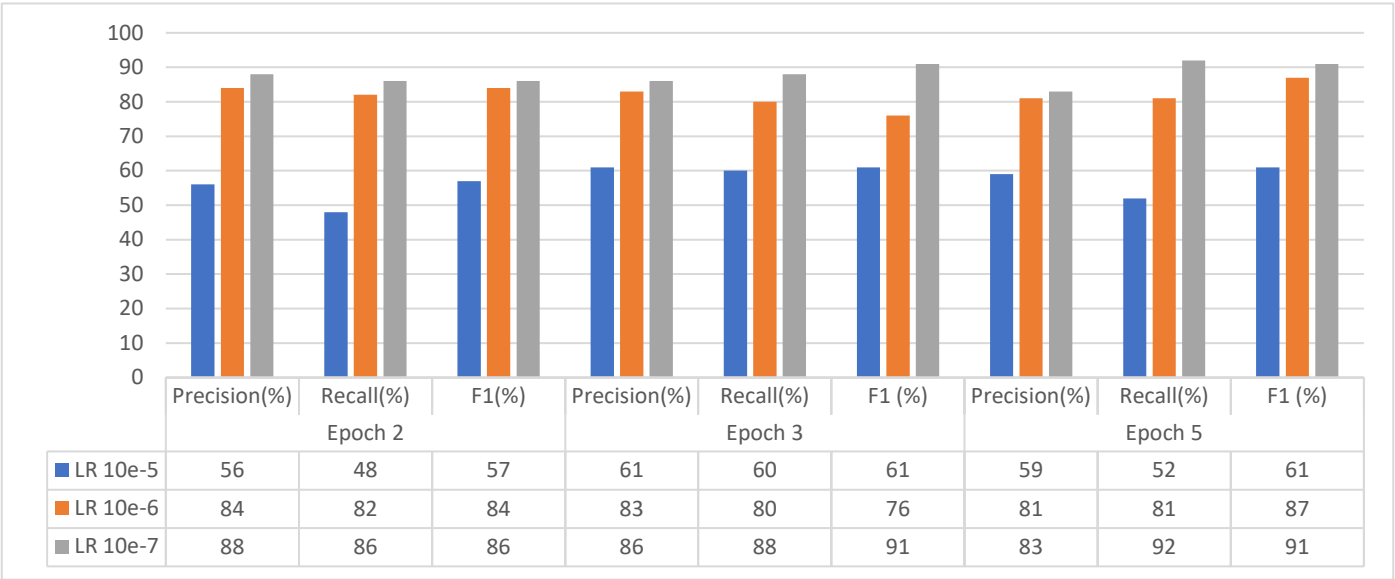


Figure 9. Results of the parameter tuning per epoch.

6. Discussion

In this study, we proposed the BERT model for the understanding the discourse around Lagos State 2023 gubernatorial election. Lagos State is a megacity with the highest population as well as the highest internal revenue generating state when compared with other 35 states in Nigeria. Agitations arising from the results of the just concluded Presidential election have given an unprecedented attention to the State election. In order to fully understand the Twitter crowd behind each of the three candidates from the three prominent candidates, personal and public tweets were examined from March 1, 2023 to March 17, 2023. BERT model was used to establish the sentiments of the public tweets. Our experiments confirmed that:

- I. Sanwoolu, being an active government official, maintained the highest number of followers.
- II. The dominant discourse in the negative Sanwoolu tweets is the occurrence of the popular EndSARS protest which allegedly took the lives of peaceful protesters in Lagos.
- III. The dominant word in the GRVLagos negative tweets is the “no man’s land” phrase which is a discussion of whether he is actually the son of the soil or qualified to rule Lagos given his maternal East-Nigeria’s descent.
- IV. Sanwoolu, however, maintained the strongest social media networks with many in-power personnel in his active circle.
- V. GRVLagos has the highest twitter impressions and the highest positive impressions of the three candidates.
- VI. Jandor does not pull the twitter crowd, hence, making the outcome of the election to become a potential two-horse race between Sanwoolu and GRVLagos.

7. Conclusion

The BERT model has also proven to be an effective sentiment model. The reason BERT model is powerful is because instead of predicting the subsequent word in the sequence, the model accounts for all words in the sequence and comes up with a deeper understanding of the context and its potency has been proven in many literature.

The actual outcome of the election under study may vary with the results in this study because it is assumed that not all eligible voters air their opinions on Twitter or even have Twitter accounts at all. Also, tweets used in this study were not location-specific, therefore, it is difficult to tell if the scraped tweets were from accounts of Lagosians. However, this study has aided our overall understanding of the Lagos State governorship election by providing insights into the networks, impressions and public sentiments of the three prominent candidates vying for the Governorship seat. In the future, authors shall consider the modified version of the BERT model such as RoBERT.

References

1. Roberts, M.E., *What is political methodology?* PS: Political Science & Politics, 2018. **51**(3): p. 597-601.
2. Achen, C.H., *Toward a new political methodology: Microfoundations and ART*. Annual review of political science, 2002. **5**(1): p. 423-450.
3. Ascher, W., *Political forecasting: The missing link*. Journal of Forecasting, 1982. **1**(3): p. 227-239.
4. Remmer, K.L., *The political economy of elections in Latin America, 1980–1991*. American Political Science Review, 1993. **87**(2): p. 393-407.
5. Oluro, M.J. and J.O. Bamigbose, *Legislative Cross-carpeting, Multiparty System and the Challenges of Democratic Good Governance in Nigeria*. Journal of Public Administration and Governance, 2021. **11**(1): p. 2639-2639.
6. Cameron, M.P., P. Barrett, and B. Stewardson, *Can social media predict election results? Evidence from New Zealand*. Journal of Political Marketing, 2016. **15**(4): p. 416-432.
7. Sanders, E., *Vox populi. On Forecasting Elections with Twitter*. 2023.
8. Bizer, C., et al., *The meaningful use of big data: four perspectives--four challenges*. ACM Sigmod Record, 2012. **40**(4): p. 56-60.
9. Tumasjan, A., et al., *Election forecasts with Twitter: How 140 characters reflect the political landscape*. Social science computer review, 2011. **29**(4): p. 402-418.
10. Allcott, H. and M. Gentzkow, *Social media and fake news in the 2016 election*. Journal of economic perspectives, 2017. **31**(2): p. 211-36.
11. Hendricks, J.A. and D. Schill, *The social media election of 2016*. The 2016 US presidential campaign, 2017: p. 121-150.
12. Wani, G. and N. Alone, *A survey on impact of social media on election system*. International journal of computer science and information technologies, 2014. **5**(6): p. 7363-7366.
13. Kwak, H., et al. *What is Twitter, a social network or a news media?* in *Proceedings of the 19th international conference on World wide web*. 2010.
14. Olabanjo, O., et al., *From Twitter to Aso-Rock: A Natural Language Processing Spotlight for Understanding Nigeria 2023 Presidential Election*. 2022.
15. Wang, H., et al. *A system for real-time twitter sentiment analysis of 2012 us presidential election cycle*. in *Proceedings of the ACL 2012 system demonstrations*. 2012.
16. Neogi, A.S., et al., *Sentiment analysis and classification of Indian farmers' protest using twitter data*. 2021. **1**(2): p. 100019.
17. Somula, R., et al., *Twitter sentiment analysis based on US presidential election 2016, in smart intelligent computing and applications*. 2020, Springer. p. 363-373.
18. Chandra, R. and R.J.I.A. Saini, *Biden vs trump: Modeling US general elections using BERT language model*. 2021. **9**: p. 128494-128505.
19. Kausar, M.A., et al., *Public sentiment analysis on Twitter data during COVID-19 outbreak*. 2021. **12**(2).
20. Carvalho, J. and A.J.A.I.R. Plastino, *On the evaluation and combination of state-of-the-art features in Twitter sentiment analysis*. 2021. **54**(3): p. 1887-1936.
21. Bibi, M., et al., *A novel unsupervised ensemble framework using concept-based linguistic methods and machine learning for twitter sentiment analysis*. 2022. **158**: p. 80-86.
22. Caetano, J.A., et al., *Using sentiment analysis to define twitter political users' classes and their homophily during the 2016 American presidential election*. 2018. **9**(1): p. 1-15.

23. Nisar, T.M. and M. Yeung, *Twitter as a tool for forecasting stock market movements: A short-window event study*. The journal of finance and data science, 2018. **4**(2): p. 101-119.
24. Yavari, A., et al., *Election Prediction Based on Sentiment Analysis using Twitter Data*. 2022. **35**(2): p. 372-379.
25. Tumasjan, A., et al. *Predicting elections with twitter: What 140 characters reveal about political sentiment*. in *Proceedings of the International AAAI Conference on Web and Social Media*. 2010.
26. Schmidt, T., et al. *Sentiment Analysis on Twitter for the Major German Parties during the 2021 German Federal Election*. in *Proceedings of the 18th Conference on Natural Language Processing (KONVENS 2022)*. 2022.
27. Razzaq, M.A., A.M. Qamar, and H.S.M. Bilal. *Prediction and analysis of Pakistan election 2013 based on sentiment analysis*. in *2014 IEEE/ACM International conference on advances in social networks analysis and mining (ASONAM 2014)*. 2014. IEEE.
28. Singh, P., R.S. Sawhney, and K.S. Kahlon. *Predicting the outcome of Spanish general elections 2016 using Twitter as a tool*. in *International Conference on Advanced Informatics for Computing Research*. 2017. Springer.
29. Jivani, A.G., *A comparative study of stemming algorithms*. Int. J. Comp. Tech. Appl, 2011. **2**(6): p. 1930-1938.
30. Devlin, J., et al., *Bert: Pre-training of deep bidirectional transformers for language understanding*. arXiv preprint arXiv:1810.04805, 2018.
31. Ravichandiran, S., *Getting Started with Google BERT: Build and train state-of-the-art natural language processing models using BERT*. 2021: Packt Publishing Ltd.
32. Lee, J.-S. and J. Hsiang, *Patent classification by fine-tuning BERT language model*. World Patent Information, 2020. **61**: p. 101965.
33. Noble, W.S., *What is a support vector machine?* Nature biotechnology, 2006. **24**(12): p. 1565-1567.
34. Suthaharan, S., *Support vector machine*, in *Machine learning models and algorithms for big data classification*. 2016, Springer. p. 207-235.
35. Pisner, D.A. and D.M. Schnyer, *Support vector machine*, in *Machine learning*. 2020, Elsevier. p. 101-121.
36. Widodo, A. and B.-S. Yang, *Support vector machine in machine condition monitoring and fault diagnosis*. Mechanical systems and signal processing, 2007. **21**(6): p. 2560-2574.
37. Fan, T.-K. and C.-H. Chang, *Sentiment-oriented contextual advertising*. Knowledge and information systems, 2010. **23**: p. 321-344.
38. Di Caro, L. and M. Grella, *Sentiment analysis via dependency parsing*. Computer Standards & Interfaces, 2013. **35**(5): p. 442-453.
39. Yu, H. and J. Yang, *A direct LDA algorithm for high-dimensional data—with application to face recognition*. Pattern recognition, 2001. **34**(10): p. 2067-2070.
40. Wei, X. and W.B. Croft. *LDA-based document models for ad-hoc retrieval*. in *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*. 2006.
41. Heimerl, F., et al. *Word cloud explorer: Text analytics based on word clouds*. in *2014 47th Hawaii international conference on system sciences*. 2014. IEEE.
42. Jin, Y., *Development of word cloud generator software based on python*. Procedia engineering, 2017. **174**: p. 788-792.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.