Twitter Sentiment Analysis on Election Using Machine Learning Techniques: A Case Study of the 2023 Presidential Elections in Nigeria

[1]Muhammad, U.,[2] Mbanaso, U. M. and [2]Kulugh, V. E.

1 - Department of Computer Science,
Faculty of Natural and Applied Science,
Nasarawa State University Keffi, Nigeria.

2 - Centre for Cyberspace Studies,
Nasarawa State University Keffi, Nigeria.

Abstract

The 2023 presidential election in Nigeria presented a contest that differed from previous elections as it presented four leading candidates that all had huge potentials to win, thus making prediction of the outcomes by mere speculations difficult. Consequently, this article implemented sentiment analysis using machine learning (ML) techniques to harvest data on Twitter and analyzed it to provide insights into the outcome of the election. The Logistics Regression and Naïve Bayes ML algorithm were used. The results showed that Logistic Regression has more promise for accuracy and performance as it provides a 97% accuracy and between 94% to 98% performance on precision, recall and F1-Score metrics. The observable pattern on the data analyzed was that each candidate had a day that they reached their summit in Twitter popularity. Findings showed that these peak days were associated with candidates' key activities during their campaigns. For instance, two of the candidates had their peak days when they spoke at the Chatham House in the UK. The other two had theirs when they engaged some influential groups within the country. A comparison of the prediction results with the actual election results however, showed a 25% accuracy as only one candidate maintained the same position in the prediction and the actual results. The implication is that other factors would have influenced the outcome of the election as data collection for the prediction was completed three weeks before the election.

Key words: presidential election, prediction, Twitter, Sentiment Analysis

Introduction

Social media data has been used in recent years for the monitoring and predicting users' sentiments towards products, brands, trending issues or events. According to Ariwala (2023), social media data has been used extensively in brand monitoring; customer service; finance and stock monitoring; business intelligence; and market research and analysis. In recent years, the application of this data in election monitoring and prediction is gaining momentum. Amongst several social media platforms, Twitter has emerged as a major micro-blogging website, with over 360 million users generating over 500 million tweets everyday (Aslam, 2023). Thus, Twitter has consistently attracted users to convey their opinions and perspectives about any trending issue, brand, company, or any other topics of interest, (Aslam, 2023). On Twitter, users share opinions in the form of tweets with a maximum of 280 characters including white spaces, (this has been recently upgraded to 4000 characters). With this tweets' length limitation, users are compelled to compress their opinions using slang, abbreviations, emoticons, short forms, etc., sometimes in combination with sarcasm and polysemy. These Characterization makes Twitter data unstructured; thus, analyzing it requires natural language processing tools. During electioneering campaigns, parties, candidates, and citizens (in this case, netizens) express their opinions on social media platforms (Twitter). These activities generate huge amount of unstructured data that has potentials for providing insights into the outcome of the elections when analyzed.

Sentiment analysis is an area of natural language processing (NLP) that studies the measurement of a sentiments (which could be positive, negative, or neutral) of a piece of text by analyzing the words used in it, (Feldman, 2013). It is often used to understand the favourability towards a product, a figure, or an event, (Mukherjee and Bhattacharyya, 2012). These days, sentiment analysis is gaining importance in the research study of text mining and NLP. The major role of sentiment classification is to analyze an online document, such as a blog, comment, review, and new items as a comprehensive sentiment and categorize it as positive, negative, or neutral (Dave, *et al*., 2021). This capability is making the study of sentiment analysis popular among scholars. Thus, several studies are being conducted on the subject. It is also known as opinion mining and sentiment classification. The sentimental analysis constitutes text classification and segregates sentiments for subjective texts, which are mainly related to consumer's reviews on products and services.

In Nigeria, previous presidential elections were easy to predict due to the nature of the contest usually presenting two main candidates under two main political party platforms. However, the 2023 general election was termed the most unpredictable election with 18 candidates from different parties vying for the office of president. However, four candidates and political parties (Atiku Abubakar/PDP, Bola Ahmed Tinubu/APC, Peter Obi/LP and Rabiu Musa Kwankwaso/NNPP) appeared as key gladiators for the presidential contest. Consequently, huge unstructured data was generated on social media and particularly Twitter by citizens stating their opinions in a bid to protect the interest of their respective candidates. Amid this unpredictability was the desire by political actors to have insights into the outcome or performance of their candidates and parties. This posed a huge challenge; thus, this paper explores the use of machine learning techniques for conducting sentiment analysis of data generated from Twitter to provide insights into the outcome of the election. Two machine learning algorithms were deployed to classify the data that was collected from twitter within the period of January 1 to January 31, 2023. Since this work extended beyond the time when the actual election was conducted, the paper attempts to compare the actual results of the election with the sentiments expressed on twitter. The rest of the paper is organized thus: section 2 highlights the related works; section 3 presents the methodology while sections 5 and 6 present data and discussed the findings respectively as section 6 concludes the paper.

Machine Learning

Machine learning (ML) is the application of artificial intelligence (AI) that aims at automating the task of analytical model building to perform cognitive tasks like object detection or natural language translation. This is achieved by applying algorithms that iteratively learn from problem-specific training data, which allows computers to find hidden insights and complex patterns without explicitly being programmed (Bishop, 2006). Three key Components of ML algorithms are identified as representation – how knowledge is represented (e.g decision trees, set of rules, etc.); evaluation – the metrics that will support the testing of ML algorithms (e.g. accuracy, precision, recall, likelihood); optimization – the search process (e.g. combinatorial, convex, and constrained optimizations). In sentiment analysis, sentiment classification or polarity classification is the binary classification task of labeling an opinionated document as declaring either a totally positive or a totally negative opinion and a technique for analyzing subjective information in many texts. A standard approach for sentiment classification is to use machine learning algorithms. These algorithms are described by Buche *et al*. (2013) to include: supervised learning; unsupervised learning; and natural language processing (NLP). NLP has several approaches, text preprocessing; lexical analysis; semantic analysis; natural language generation; and application of sentiment analysis.

In NLP, sentiment analysis has become one of the many fields of computational studies (Bollen *et al*., 2021). In general, sentiment analysis deals with the mining of information related to sentiments or opinions from a group for a specified topic. Sentiment analysis involves a series of tasks, namely, sentiment classification, subjective analysis, opinion holder extraction, and aspect or object-based extraction. The subjective analysis involves evaluating a text document or a sentence to label them as subjective or objective. After this step, these documents or sentences tagged as objective are immediately discarded, as they are not much of use for the Sentiment Analysis process.

In Oyebode and Orji, (2019) both lexicon-based and supervised learning techniques were used in analyzing public sentiments on two popular candidates and political parties during 2019 presidential election in Nigeria. On the three lexicon-based techniques used, VADER-EXT outperformed VADER and TextBlob while Logistics Regression out-performed the supervised learning techniques used in the study. The key drawback on this study is that it focused on only two popular presidential candidates as contrasted to our study that deals with four popular presidential candidates. Secondly, the dataset was mined from a less popular social media platform – Nairaland, with far less reach when compared with twitter. However, this study mined data from twitter, a more robust social media platform. Barnaghi et al., (2016) use unigrams and bigrams and apply Term Frequency Inverse Document Frequency (TF-IDF) to find the weight of a particular feature in a text and hence filter the features having the maximum weight. The TF-IDF is a very efficient approach and is widely used in text classification and data mining. Bouazizi and Ohtsuki (2019), proposed an approach where besides reliance on vocabulary, the expressions and sentence structures used in different conditions were considered. They classified features into four classes: sentiment-based; punctuation and syntax-based; unigram-based; and pattern-based features. However, in multi-class classification, different sentiments have much in common e.g. fun and happy. Thus, classification accuracy decreases with an increase in number of sentiments.

The work of Mamgain *et al*., (2016) is a bit different as the authors did not focus on a particular topic or event but proposed to find trending topics in a region. They divided the features extracted in two categories: Common Features and Tweet Specific Features. The former is combination of common sentiment words while the later includes @-network features, user sentiment features and emoticons. Based on the post time of each user, a feature vector is built. However, the work Mamgain *et al*., (2016) focused on sentiments of colleges/institutions and not political party/candidates. Chung and Mustafaraj (2011) performed sentiment analysis of tweets collected from MAsen10 campaign. They used Opinion-Finder technique to find sentiment analysis and achieved overall accuracy of 41% which is not reliable. To improve accuracy, they further used SentiWordNet a lexical resource with 207, 000 pair of words, thus, the accuracy improved to 47.19%. They introduced sentiment analysis to tweets but failed to increase the accuracy with methods they adapted due to low coverage of dataset used for the study.

Avello, Metaxas and Mustafaraj (2011) used the dataset belonging to MAsen10 and USsen10 in studying predictive power of social media against several senate races by conducting several sentiment analysis experiments. They argued that election results cannot be predicted using simple tweet share, sentiment analysis must be performed for achieving better results. Avello at al., (2011) concluded pre-election tweets volume for MAsen10 seemed to be good for prediction of elections. They calculated Mean Absolute Error (MAE) 17:1% following Tumasjan (2010) method and sentiment analysis MAE as 7:6%. Predictions were computed according to Sentiment Analysis and Twitter Chatter volume. No new technique was deployed, but rather, a replication of previous research (Chung and Mustafaraj, 2011) with the aim of getting a repeated or same results.

Another detailed study for US Midterm Elections 2010 was carried out by Livne who also found correlation in the outcomes of US midterm elections 2010 and tweets (Simmons, 2011). In their study, they analyzed around 460,000 tweets over three years for 687 candidates contesting for National House, State Governor, or Senate seats. Model built by them based on graph structure, content and election results was able to predict winner or loser candidate with accuracy of around 88.0%. The Comparison of 2012, French and US Presidential elections was made by Farhad (2013) who adopted the approach of performing time series sentiment analysis by restricting their findings to only two candidates, Barack Obama, and Mitt Romney in case for US. They performed time series sentiment analysis for US candidates Farhad (2013) and for French candidates by computing scores, based on three different scoring functions, polarity, sentiment, and affinity. They claimed twitter to be the best medium for judging the candidates and responding to their voters and vice versa.

Materials and Methods

Proposed Architectural Model

The sentiment classification model classifies tweets as positive or negative. Its implementation is based on a system that uses features derived from the data collected from twitter and the architecture of the proposed model for sentiment classification is presented in the Figure 1. The framework comprises of four stages which includes text collection and cleaning; pre-processing; sentiment analysis and finally evaluation.

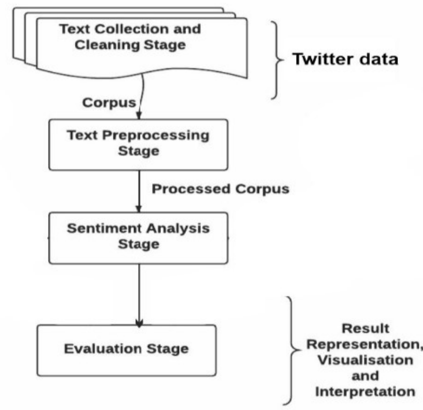

Figure. 1: Architecture of the Proposed Model

Data Collection

A total of 1,675,020 tweets were collected based on four key candidates of the presidential election using snscrape python library in twitter between 1st January 2023 to 31st January, 2023. This is presented in Table 1. Dates were mapped to filter the daily trends of the tweets collected, and tweet attributes like locations, tweet_id, usernames were filtered out. The collected tweets were generally between 25 and 50 words. For sentiment classification, all the extracted data were arranged based on candidate's names.

Table 1: Tweets Collected before Pre-processing

#	**Candidate**	**Number of Tweets**
1.	Peter Obi	740, 991
2.	Tinubu	541, 746
3.	Atiku	344, 968
4.	Kwankwaso	47, 315
	Total	**1, 675, 020**

Data Cleaning

Texts from social media platforms, like Twitter, tend to be noisy with non-text entities, which decrease the accuracy of the sentiment scores assigned to them. Consequently, excessive punctuation, emoji, and acronyms were ignored during the sentiment model implementation. This was to minimize the effect slangs, acronyms, and misspellings on the overall sentiment as emphasized by (Agarwal *et al*., 2011). Term Frequency, Inverse Document Frequency (TF-IDF) sentiment analysis tool was used for the study, as it is specifically tuned for sentiments expressed in social media posts, (Hutto and Gilbert, 2022).

Data Preprocessing

Raw tweets, often referred to as noisy data are collected from Twitter. This requires further pre-processing for machine learning analysis. This is achieved by filtering the data using python libraries to retrieve meaningful parts of the tweets and excluding unnecessary contents which can affect the accuracy of the model. Attributes such as links, urls, retweet counts, usernames, stopwords, emoticons and locations were not considered useful in this study as they potentially magnify the noise in the data. Thus, lower case conversion, punctuation, removal of hashtags and other Twitter notations (@ RT), URLs, stop words, tokenization, and stemming were all part of the pre-processing activity. In this work the following 10 steps for Data Pre-processing were applied:

i. Lower case conversion,
ii. Removal of punctuations,
iii. Removal of hashtags,
iv. @RT and URL,
v. Removal of stop words,
vi. Stemming of words, tokenization of words,
vii. Removal of missing values,
viii. Removal of emoticons,
ix. Removal of duplicate tweets, and
x. Removal of neutral sentiments.

Feature Extraction

TF-IDF is a well-known technique in NLP for obtaining useful words and their scores from the given corpus (Havrlant and Kreinovich, 2017). TF represents how many times a particular word has appeared in the corpus. Another significant measure of the importance of a word is document frequency (DF); it describes how many documents contain a specific term. IDF is the multiplicative inverse of DF and TF, it provides a measure of the occurrence of certain words. The term frequency $tf(i,\delta)$ is given in Equation (1):

$$tf(i,\delta) = \frac{f\delta(i)}{f_{\delta}\,(w)} \qquad (1)$$

where $f\delta(i)$ is the term frequency of i in the document δ, while $f_{\delta}(w)$ is the total words in document δ. Similarly, the IDF of ith word Δ can be expressed as given in Equation (2).

Where:

$$idf(i,\Delta) = ln\, ln\left(\frac{|\Delta|}{|\gamma|}\right) \qquad (2)$$

where Δ is the total documents and γ represents documents with term i. The TF-IDF might not always be a suitable means of extracting emotions and sentiments from the data. In the case of sarcasm, the frequency score might reflect wrong sentiments. For example, "it is okay if you do not like me, not everyone has good taste", or "I do not have the energy to pretend to like you today." In both cases, TF-IDF might reflect positive sentiments while they are intended to be negative. Thus, the study applied bi-grams and tri-grams techniques in text processing to correlate words with neighboring words to understand the context in a better way (Jurafsky, 2000).

Classification Algorithms

Logistics Regression

Logistic regression is a supervised learning algorithm which is used to calculate or predict the probability of a binary (yes/no) event occurring. Logistics Regression is applied to this research to determine whether a tweet is positive sentiment or a negative sentiment. Since there are two possible outcomes to this question - positive, or negative, this is called binary classification. The logistic regression is presented in equation (3)

$$h0(x) = \frac{1}{1 + e - (\beta_0 + \beta_1 X)} \qquad (3)$$

Where:

$h0(x)$ = output of logistic function, where $0 \leq h0(x) \geq 1$
β_1 = the slope
β_0 = the y-intercept
X = I independent variable

Naïve Bayes Algorithm

Naive Bayes is a well-known technique for classification; it uses Bayesian statistics while assuming that features are statistically independent of each other. In a recent study, Sunil (2023) argued that this assumption is often not true in real-world scenarios. Naive Bayes can learn high-dimensional data with minimal training. We selected Naive Bayes for classification because it is scalable and very lightweight. Since tweet data grow steadily with time, it is the most suitable classifier with stable and predictable results. A study shows the theoretical basis behind the excellent performance of Naive Bayes classification, (Zhang, 2004). This algorithm is expressed mathematically in equation (4)

$$P(x_1, x_2, x_3, \ldots x_n) = \frac{P(y)P(y)}{P(x_1, x_2, x_3, \ldots x_n)} \qquad (4)$$

$P(y)$ = labels
$P(x_1, x_2, x_3, \dots x_n)$ = Comments
$P(y|x_1, x_2, x_3, \dots x_n)$ = the probability of hypothesis(labels) given the observed evidence
$P(x_1, x_2, x_3, \dots x_n|y)$ = the probability of evidence given that hypothesis is true
$P(y)$ = the probability of hypothesis before observation
$P(x_1, x_2, x_3, \dots x_n)$= the probability of new evidence given all possible hypotheses

Training and Testing

After labelling the collected data using text polarity score, the tweets were inspected to cross-check sarcasm, and identify sentiments. To keep the dataset more balanced, tweets labelled as neutral were filtered out because, some were sarcasm, made in other language or full of adverts and promotions so they were crosschecked and filtered out to avoid overfitting the model. Consequently, the 1,675,020 tweets collected, only 807,699 were utilized for positive and negative sentiment analysis in this study. The remaining 867,321 tweets were filtered out because they contain low volume of needed information as majority were labelled as neutral, some are duplicates while others contain adverts and promotions or made with other local languages. The 807,699 tweets (dataset) were divided in the ratio of 60:40 for training and testing respectively. Cross-validation and confusion matrix were used for evaluating the model.

Results

Figure 2 presents the breakdown of tweets labelled positive and negative sentiments. The pie chart in Figure 2 shows that majority of the tweets (68.87) were labelled as positive while (31.13%) were labelled negative, based on the text polarity score.

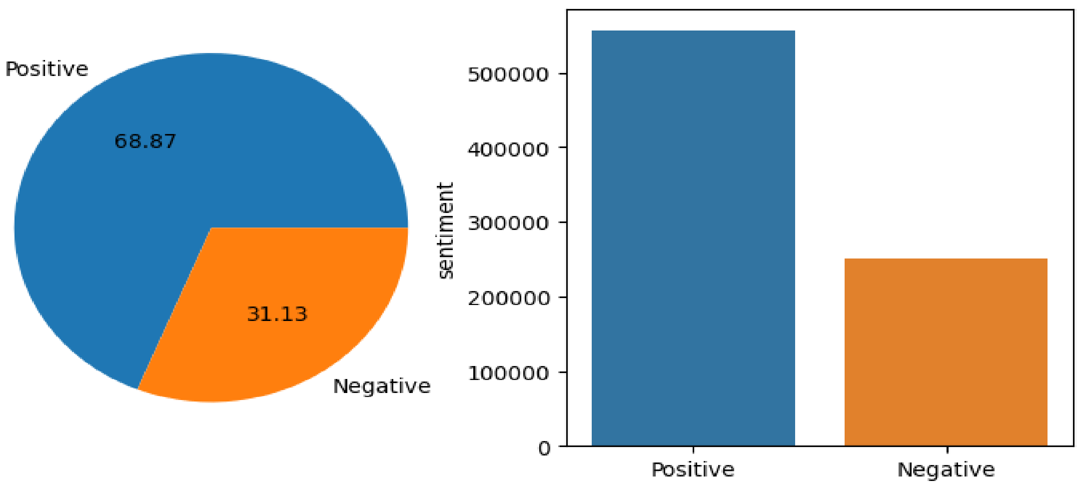

Figure 2: Distribution of Tweets Labelled Positive and Negative Sentiment

Figure 3 presents the number of tweets collected per candidate per day within the period in which data was collected, i.e. January 1 to January 31, 2023, for each of the four candidates.

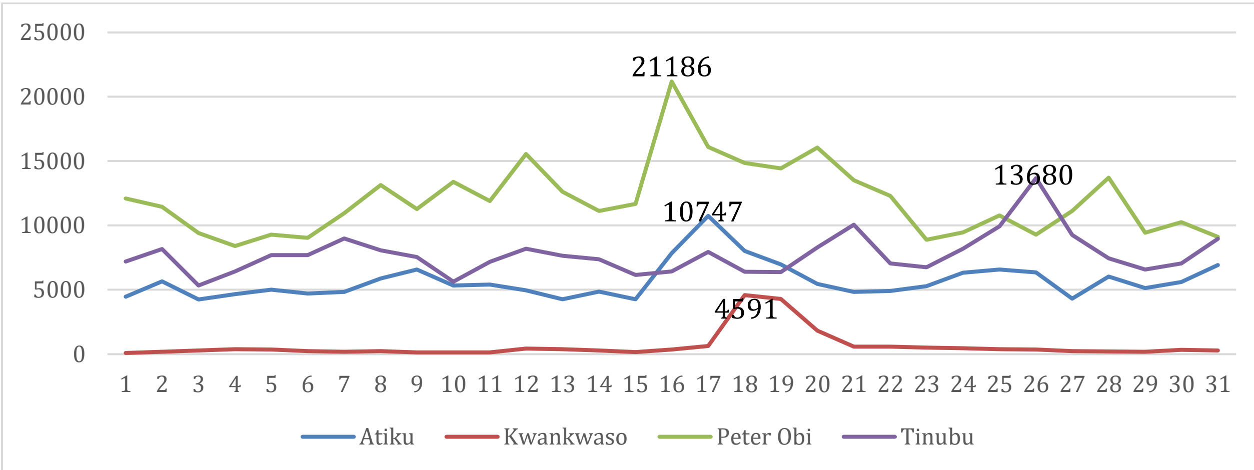

Figure 3: Tweets per Day Per Candidate From 1st to 31st January 2023

The tweets were categorized into four based on the leading presidential candidates. It can be observed that each candidates/party have their peak tweet day(s) with most of them summitting between the 3rd and 4th weeks of January 2023. For instance, NNPP's Kwankwaso peaked on the 18th of January 2023 with 4,591 tweets. PDP's Atiku had his summit point of tweets on 17th January 2023 with 10,747 tweets. Peter Obi and Tinubu had their peaks on 16th and 26th January 2023 with 21,186 and 13,680 tweets respectively. The peak presentation showed that Peter Obi has 42.20% of peak tweets followed by Tinubu with 27.25%. Atiku and Kwankwaso has 21.41% and 9.14% of tweets peak scores respectively.

Figures 4 and 5 presents word cloud visualization that was used to examine how tweets are distributed, with size of each word indicating the frequency. Additionally, it is also used to emphasize related datapoints. Figures 4 and 5 show the frequency of words mentioned in both positive and negative tweets in the dataset.

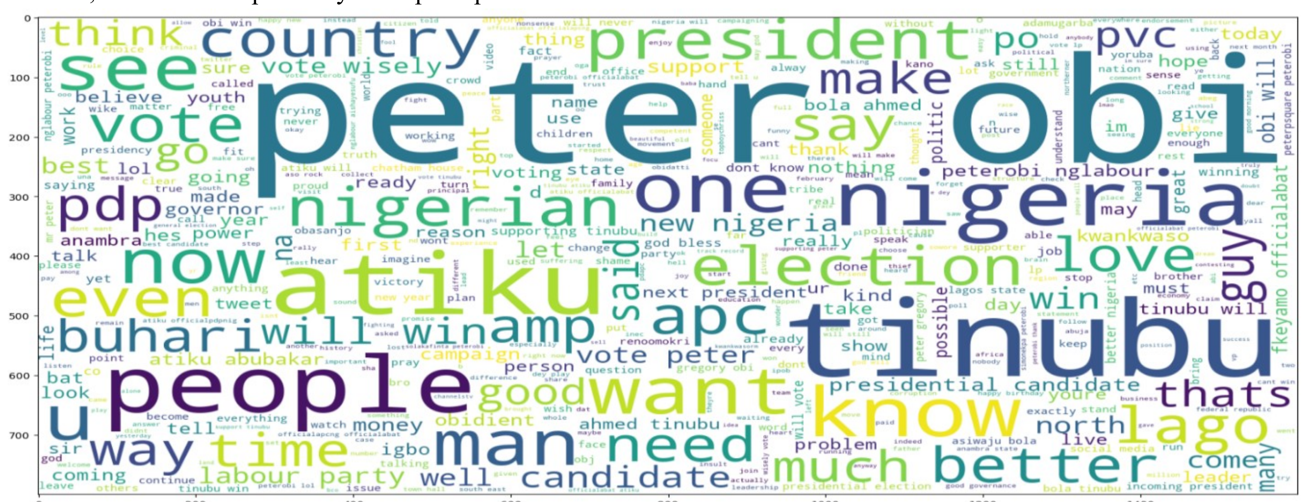

Figure 4: Word Cloud for Positive Tweets

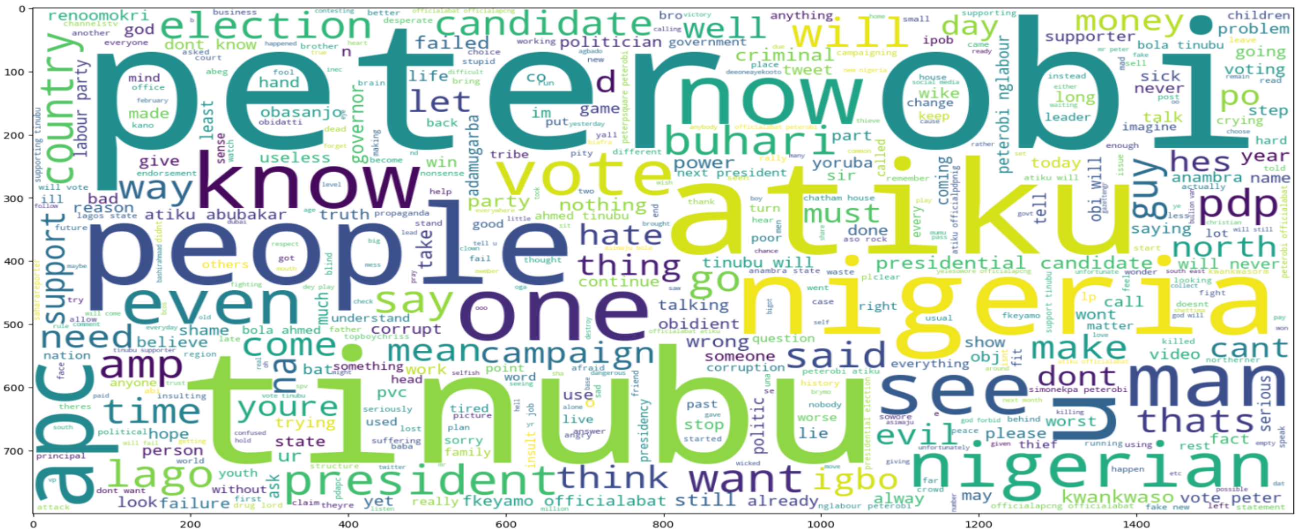

Figure 5: Word Cloud for Negative Tweet

Twitter Popularity and Elections Results

In this article, the authors attempt to compare the twitter popularity based on the sentiment analysis with the actual elections results as published by the independent national electoral commission (INEC). Figure 6 shows the popularity of each of the candidate on twitter during the electioneering campaigned as captured in the study timeframe of January 1 to 31, 2023. Figure 7 on the other hand shows the number of votes gained by candidates/parties during the elections.

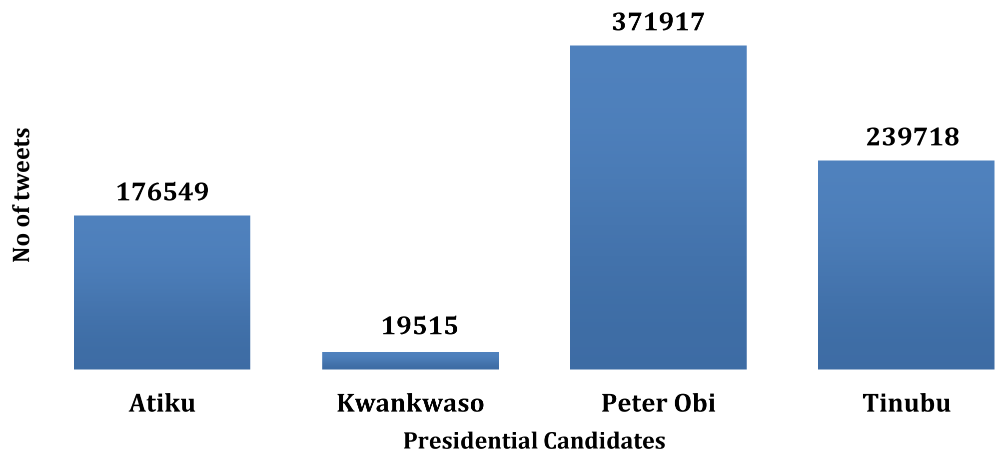

Figure 6: Candidates' Twitter Popularity Data

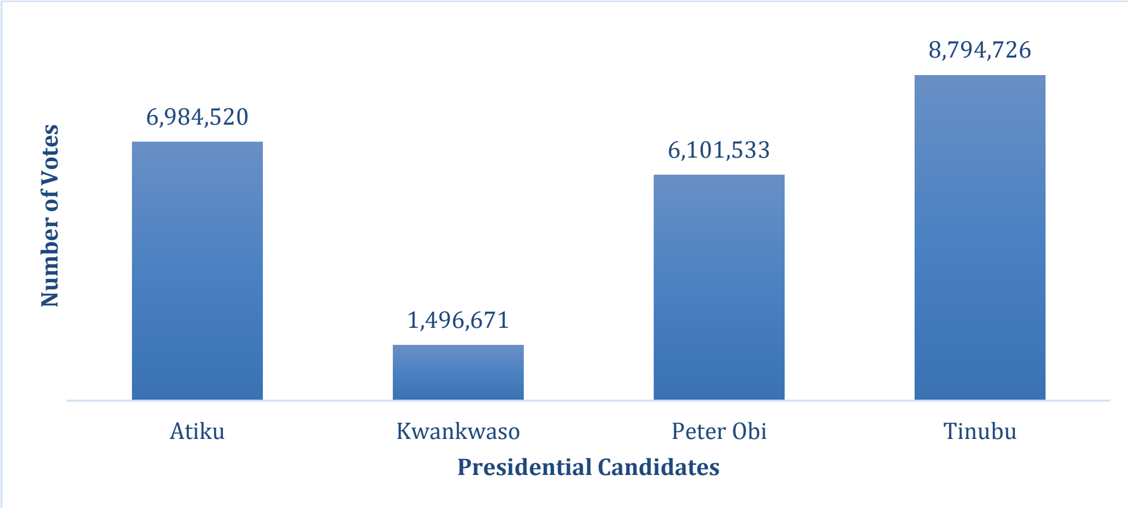

Figure 7: Votes Scored by Candidates

The data in Figures 6 and 7 shows that Peter Obi – Labour Party maintained twitter popularity that was higher than those of the other three candidates within the time frame of data collection as he gained 46.05% of tweets as compared to Tinubu (29.68%), Atiku (21.86%) and Kwankwaso (2.42%). However, the votes scores in the elections presented a near reverse trend by placing Peter Obi – Labour party third; Bola Ahmed Tinubu - All Progressives Congress, who maintained second position on twitter popularity placed first. Atiku Abubakar of the PDP on the other hand placed third on twitter popularity but placed second in the election results. The candidate of the NNPP – Rabiu Musa Kwankwaso, however, maintained a consistent outlook in both twitter popularity and election results as he placed fourth position in both cases.

Accuracy and Performance Evaluation

807,699 tweets were used to create a training and testing dataset. The dataset was partitioned into a 60% training and 40% testing dataset.

Table 2: Accuracy and Performance Evaluation for Logistics Regression

	Precision	**Recall**	**F1-Score**
0	0.96	0.94	0.95
1	0.98	0.98	0.98

Accuracy = 97%

After training the Logistics Regression with 60% of the dataset, the system was tested with 40% test dataset. The results showed an accuracy of 97%, with a precision of 96%, Recall of 94% and F1 score of 95% for the negative sentiments and 98% as precision, recall and F1 score for the positive sentiment as shown in the Table 2. The performance was further evaluated using the Confusion Matrix in Figure 8 shows that 191,015 tweets were classified as true positive (TP) which stands for the true label being positive and predicted label being positive, 3,794 as false negative (FN) the true label being positive while the predicted label is negative. 4858 as false positive (FP) denotes that the true label is negative while the predicted label is negative. Finally, 83,028 for true negative (TN), the true label is negative while the predicted label is negative.

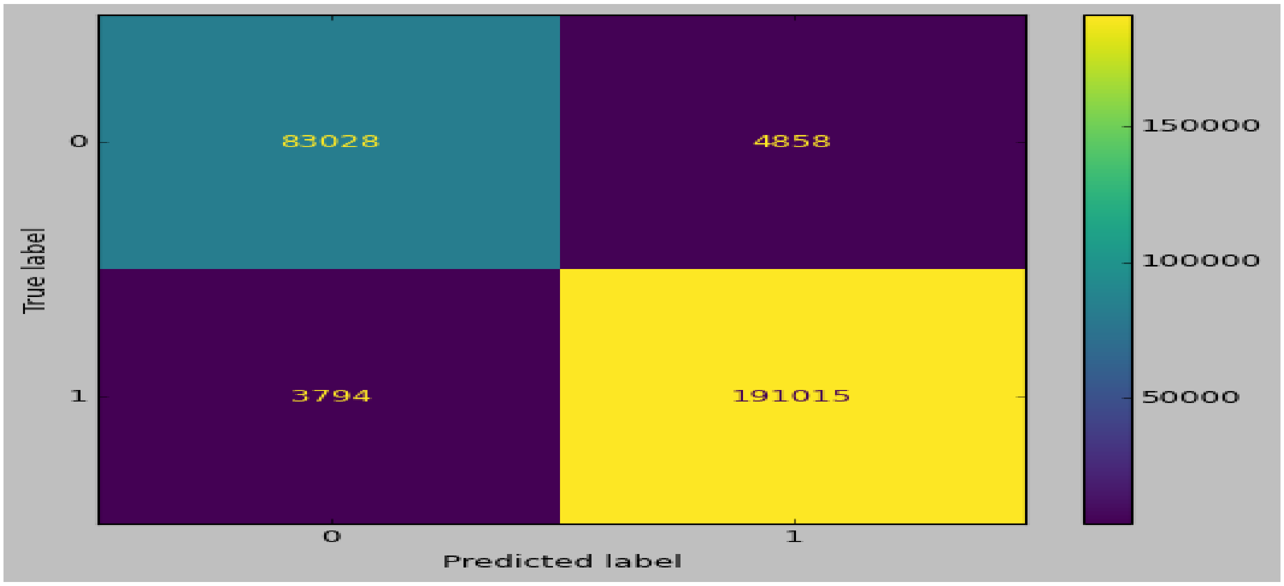

Figure 8: Confusion Matrix of Logistics Regression

Table 3: Accuracy and Performance Evaluation for Naïve Bayes

	Precision	Recall	F1-Score
0	0.84	0.63	0.72
1	0.85	0.94	0.90

Accuracy = 85%

After training the model using Naïve Bayes classifier, the results show the accuracy of 85%, with the precision of 84%, Recall of 63% and F1 score of 72% for the negative sentiments and 85%, 94%, and 90% as the precision, recall and F1 score respectively for the positive sentiment as shown in the Table 3. The performance was further evaluated using the Confusion Matrix in figure below, which shows that 184,081 tweets were classified as true positive (TP) which stands for the true label being positive and predicted label being positive, 10,728 as false negative (FN) the true label being positive while the predicted label is negative. 32,360 as false positive (FP) denotes that the true label is negative while the predicted label is negative. Finally, 55,526for true negative (TN), the true label is negative while the predicted label is negative.

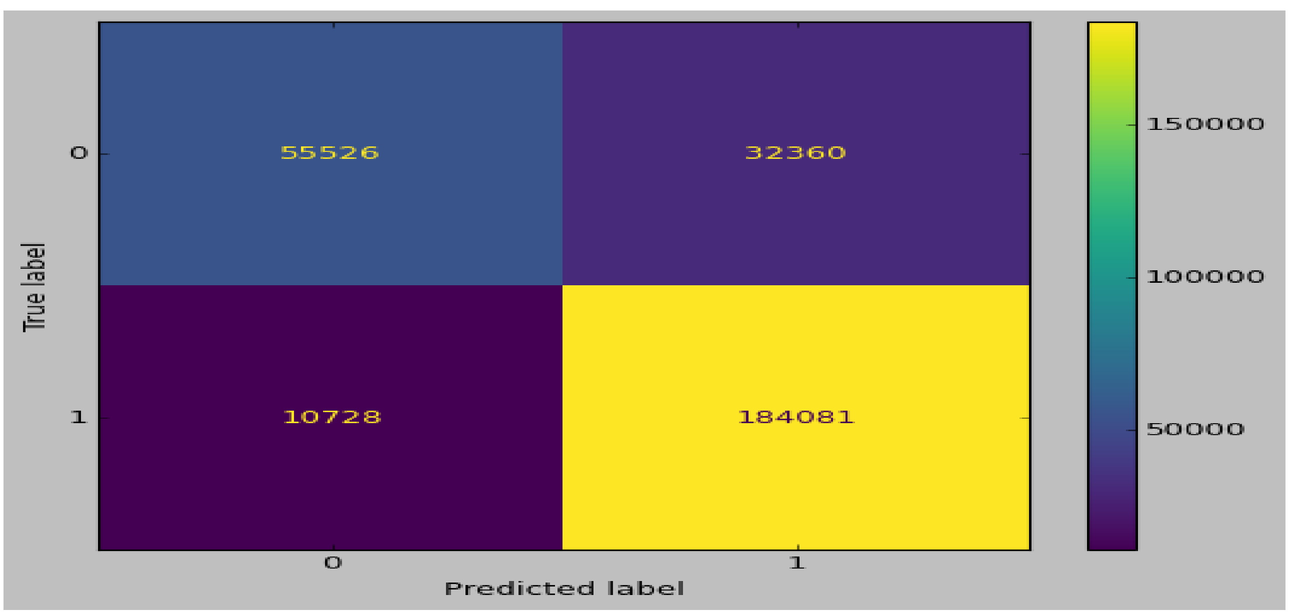

Figure 9: Confusion Matrix of Naïve Bayes Classifier

Discussion

The 2023 presidential election in Nigeria presented a more sophisticated race that was hard to predict because four leading presidential candidates were in the race with potentials to win. Thus, the desire by political actors to predict the winner of the election could best be achieved by machine learning techniques. Findings in this study reveal that the peak tweet popularity achieved by candidates as shown in Figure 3 were associated with certain key events during the campaign within the period of data collection. For instance, Peter Obi and Rabiu Musa Kwankwaso got to their peak popularity on 16th and 18th January 2023 respectively. These coincided with the dates they both spoke at Chatham House in the United Kingdom. Although Tinubu spoke at the Chatham House on an earlier date (5th December 2023, which did not cover the reference period for the data collection for this study), his peak popularity within the reference period coincided with the date he spoke at Abeokuta in Ogun state – that was 25th January 2023. Atiku's tweet popularity peak relates to the day he had an engagement with Nigerian Economic Summit Group, in Lagos, Nigeria – 18th January 2023.

Although, the works of (Simmons, 2011; Chung and Mustafaraj, 2011) in sentiment analysis predicted election results with 88% accuracy as compared to the election results in the USA election of 2012 and 47% respectively; the prediction based on sentiment analysis in our article as presented in Figures 6 and 7 presents a near reverse trend and transposition of positions of the candidates. This saw the lead candidate in tweet popularity move to third position, second position in tweets popularity moved the first and the third position moved to second. The only exception is the candidate of the NNPP – Kwankwaso, who maintained fourth position in both prediction and actual results of the election.

When the accuracy and performance of the Logistic Regression and Naïve Bayes are compared; Logistic Regression outperformed Naïve Bayes algorithm with an accuracy of 97% as against Naïve Bayes' 85% accuracy. Also, Logistic Regression did better than Naïve Bayes on precision, recall and F1-Score as its lowest value on these metrics is 94% and highest value 98%. Naïve Bayes on the other hand has its lowest value as 63% and highest value as 90%. Thus, Logistics Regression presents as the best algorithm for this sentiment analysis.

The authors consider that the application of more ML and DL models have potentials to improve the prediction of future elections in Nigeria. Additionally, the adoption of more social media data sources like Facebook, Instagram and even national dailies could broaden the prediction and allow the comparison of data from various sources, thus, provide more insights into the future.

Conclusions

The 2023 presidential election Nigeria presented four leading contenders which is a sharp departure from previous contests that had two leading contenders and make prediction of winners relatively easier even by manual methods. Thus, in this paper, the authors argued that the application of machine learning techniques through natural language processing using opinions expressed by citizens and other political actors will provide the needed data and tools for prediction. Consequently, data was generated on Twitter from the 1st of January to 31st January 2023. This data was analyzed using two ML algorithms, the results showed that ML can reasonably predict elections. Key findings showed that candidates had the peak periods in tweets popularity on certain events in their campaigns during the reference period. The results of the Twitter popularity as analyzed when compared to the actual results of the elections, however revealed that the prediction was only 25% accurate as only one out of the four candidates maintain the same position in prediction and the actual election results.

References

Agarwal A, Xie B. Vovsha, I., Rambow, O. and Passonneau, R. J. (2011): Sentiment analysis of twitter data. In: Proceedings of the Workshop on Language in social media. 2(1): 30–38.

Ariwala, P. (2023): Introduction to Sentiment Analysis: Concept, Working, and Application. Maruti techlabs. 8(9): 45-53. https://marutitech.com/introduction-to-sentiment-analysis/

Aslam, S. (2023): Twitter by the numbers: Stats, Demographics and Fun Facts. Retrieved from https://www.ombicoreagency.com/witter-statistics/#:~:text=Statistics%20-(editor's%20choice)-,1., Titer%20(Twitter%2C%202022

Avello, D. G., Metaxas, P. T., Mustafaraj, E. (2011): Limits of Electoral Predictions Using Social Media Data, Fifth International AAAI Conference on Weblogs and social media, July 11-13, 11-16.

Barnaghi, P., Bermudez-edo, M. and Tonjes, R. (2016): Challenges for Quality of Data in Smart Cities. ACM Journal of Data and Information quality. 3(2): 1-3.

Bishop, C. M. (2006): Pattern recognition and machine learning (Information science and statistics). Springer-Verlag New York, Inc. ISBN: 0387310738

Bollen, J., Mao, H. and Zeng, X. (2021): Twitter mood predicts the stock market. Journal of Computational Science. 2(1): 1-8.

Bouazizi, M. and Ohtsuki, T. (2019): Multi-Class Sentiment Analysis on Twitter: Classification Performance and Challenges. Big Data Mining and Analytics ISSN 2096-0654, 2(3). 181–194. DOI: 10.26599/BDMA.2019.9020002.

Buche, A., Chandak, M. B. and Zadgaonkar, A. (2013): Opinion mining and analysis: A survey, International Journal on Natural Language Computing (IJNLC). 2(3): 41-52.

Chung, J. and Mustafaraj, E. (2011): Can Collective Sentiment Expressed on Twitter Predict Political Elections? Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence July 11-13, 1770-1771

Dave, K., Lawrence, S. and Pennock, D. M. (2021): Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In Proceedings of the 12th international conference on World Wide Web, 519-528.

Farhad, N., Viswanathan, A. and Chiru, C. (2012): 2012 Presidential Elections on Twitter -- An Analysis of How the US and French Election were Reflected in Tweets. Control Systems and Computer Science (CSCS), 2013 19th International Conference on. DOI: 10.1109/CSCS.2013.72, 1149-1156

Feldman, R. (2013): Techniques and applications for sentiment analysis. Communicate ACM. 56(4): 82–89.

Havrlant, L. and Kreinovich, V. (2017): A simple probabilistic explanation of term frequency-inverse document frequency (tf-idf) heuristic (and variations motivated by this explanation). International-Journal-of-General-Systems, 4(2), 1563-5104

Hutto, C. and Gilbert, E. (2022) Vader: A parsimonious rule-based model for sentiment analysis of social media text. In: Proceedings of the International AAAI Conference on Web and social media; DOI:10.1609/icwsm.v8i1.14550, 147-157

Jurafsky, D. and Martin, J, H. (2000): Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice Hall. New Jersey 07458

Mukherjee, S. and Bhattacharyya, P. (2012): Feature specific sentiment analysis for product reviews. In: International Conference on Intelligent Text Processing and Computational Linguistics, Springer. 5(1): 475–487.

Mamgain, N., Mehta, E., Mittal, A. and Bhatt, G. (2016): Sentiment analysis of top colleges in India using Twitter data, *2016 International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT)*, New Delhi, India, 525-530, doi: 10.1109/ICCTICT.2016.7514636.

Oyebode, O. and Orji, R. (2019). Social Media and Sentiment Analysis: The Nigeria Presidential Election 2019, *2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Vancouver, BC, Canada, 2019, 0140-0146, doi: 10.1109/IEMCON.2019.8936139.

Simmons, A. L, M. P., Adar, E. and Adamic, L. A. (2011): The Party is Over Here: Structure and Content in the 2010 Election, Fifth International AAAI Conference on Weblogs and social media. 2011, Association for the Advancement of Artificial Intelligence (www.aaai.org), 201-208

Sunil, R. (2023): Naive Bayes Classifier Explained: Applications and Practice Problems of Naive Bayes Classifier. Analytics Vidhya. 3(1): 67-75. https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/

Tumasjan, A., Sprenger, T., Sandner, P. and Welpe, I. (2010): Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, Washington DC, 23-26 May 2010, 178-185.

Zhang, T. (2004): Statistical analysis of some multi-category large margin classification methods. Journal of Machine Learning Research. 5(3): 1225-1251.