UNIABUJA Journal of Engineering and Technology
https://ujet.uniabuja.edu.ng/
ISSN: 2714-3236 (Online); 2714-3228 (Print)
Volume 3, Issue 2, 2026; 59-72

Deep Learning-Driven Sentiment Analysis for Electoral Outcome Prediction

Nurudeen JIBRIN[1], Chaku E. SHAMMAH[2], Ahmed IBRAHIM[3] Abdulhamid ABUBAKAR[4], Okorie S. ONYEDIKACHI[5]

[1]*Department of Computer Science, Baze University, Abuja, Nigeria*
[2,3,4]*Centre for Cyberspace Studies, Nasarawa State University, Keffi, Nasarawa State, Nigeria*
[5] *Economic Community of West African States (ECOWAS)*

[1*]*nurudeen.jibrin@bazeuniversity.edu.ng*, [2]*chakushammah@nsuk.edu.ng*, [3]*ibrahimloko@nsuk.edu.ng*, [4]*abdulhamid@ab-bkr.com*, [5]*sokorie@ecowas.int*

Abstract
Social media platforms offer a unique, real-time window into public opinion, creating novel opportunities for measuring electoral sentiment. This study employs a mixed-methods approach to critically evaluate the predictive power of X (formerly Twitter) sentiment analysis for electoral forecasting, using Nigeria's 2023 presidential election as a case study. We analysed a dataset of 136,500 tweets and conducted semi-structured interviews with 15 domain experts. Long Short-Term Memory Recurrent Neural Network (LSTM-RNN) was benchmarked against traditional classifiers, Naïve Bayes, and Support Vector Machines (SVM). The optimised LSTM-RNN model achieved superior performance (93.8% accuracy, 95.0% F1-score). However, a moderately strong yet statistically insignificant correlation with official election results underscores the limitations of sentiment analysis as a standalone predictive tool. Qualitative insights contextualize this finding, highlighting the volatility of online discourse, challenges of multilingualism, and critical issues of representativeness, including urban bias and the influence of misinformation. While demonstrating the considerable promise of deep learning, our results reveal significant pitfalls, advocating for hybrid predictive frameworks that integrate real-time sentiment tracking with demographic weighting and multimodal data validation. Beyond forecasting, this research underscores the utility of social media analysis for promoting governance and participatory democracy in multilingual and resource-constrained contexts.

Keywords: *Deep learning, sentiment analysis, electoral forecasting, social media, Nigeria's presidential election.*

1.0 Introduction

The digital transformation of Nigeria has profoundly reshaped political participation, with broadband subscriptions surpassing 100 million. However, as of May 2025, broadband penetration stood at approximately 47.73%, while the Nigerian National Broadband Plan (NBP) targets 70% penetration by the end of 2025 (Nigerian Communications Commission (NCC), 2020). Social media platforms, particularly X (formerly Twitter), have become central arenas for political expression and civic engagement, especially among youths. During the 2023 presidential elections, over 1.9 million election-related posts reflected public sentiment on pressing issues such as unemployment, governance reforms, and education (Attai et al., 2024). This unprecedented surge in online political discourse presents a fertile ground for computational political analysis, offering real-time insights into voter attitudes and electoral dynamics (Muñoz et al., 2024). However, extracting actionable knowledge from such unstructured, multilingual, and context-dependent data demands sophisticated Natural Language Processing (NLP) techniques that go beyond conventional polling approaches, which remain costly, slow, and prone to sampling bias (Zhang et al., 2018), (Myilvahanan et al., 2023), (Jose Gonzalez-Gomez et al., 2024).

Sentiment analysis, enabled by NLP, has emerged as a transformative approach for electoral forecasting. Recent studies have explored a range of techniques, including lexicon-based approaches, classical machine learning models, and deep learning architectures. For instance, Oyebode and Orji (2019) applied VADER-based models to Nigerian election data, achieving moderate predictive performance but lacking contextual depth. Similarly, Alade and Nwankpa (2022) utilised Naïve Bayes classifiers on student-related tweets, reporting limited accuracy due to small dataset sizes. More recent work has explored deep learning frameworks such as LSTM-based architectures and ensemble models, demonstrating improved classification performance but often constrained by data quality and linguistic diversity (Barik et al., 2023; Alsayat, 2022). Transformer-based architectures such as BERT, RoBERTa, and their variants have also achieved remarkable

Received: 20-08-2025 / Accepted: 12-09-2025 / Published: 27-04-2026

https://doi.org/10.70118/ujet.2026.0302.06

accuracy in detecting nuanced expressions of opinion, including sarcasm and code-switching (Islam et al., 2024).

Despite these advances, existing studies exhibit several critical limitations. First, many studies present results without sufficient contextual discussion, often relying heavily on tabular summaries rather than synthesizing findings to establish clear research gaps. Second, prior works frequently overlook the unique linguistic characteristics of the Nigerian digital space, including code-mixing, regional dialects, and Nigerian Pidgin, which significantly affect sentiment classification performance (Inuwa-Dutse, 2025). Third, while some models report high accuracy levels, they often fail to capture the dynamic and context-specific evolution of political discourse, limiting their applicability in real-time electoral forecasting (Paul Anule & Jennifer Ifeoma, 2025).

Even as academic attention increases, three critical limitations remain unresolved in the literature. First, most studies rely on static datasets collected prior to or after election periods, overlooking the real-time volatility of voter sentiment during campaigns and polling (Rita et al., 2023). Second, deep learning models have yet to be systematically adapted to Nigeria's linguistic ecosystem, which is characterized by Pidgin, slang, and rich code-switching practices (Garba et al., 2024). Third, few attempts have empirically validated whether sentiment-derived predictions correspond to actual electoral outcomes, raising questions about the external validity of computational forecasting (Oyewola et al., 2023). Addressing these gaps is crucial for advancing both the methodological rigor and practical relevance of computational electoral studies in emerging democracies.

This study responds to these gaps by proposing a novel deep learning framework for sentiment-driven election forecasting in Nigeria's 2023 presidential elections using X data. The framework integrates a hybrid LSTM-RNN optimised for local linguistic features such as sentiment-carrying emojis, code-mixing, and informal dialects. In doing so, it benchmarks performance against traditional classifiers such as Naïve Bayes and SVM, using accuracy, precision, recall, and F1-score as evaluation metrics. Furthermore, it extends the scope of prior work by quantifying the degree of alignment between sentiment-derived forecasts and official election results, thereby providing an empirical assessment of social media's predictive capability in the Nigerian context**.**

The contributions of this study are threefold. First, it develops a replicable methodological framework for real-time election sentiment analysis in multilingual settings. Second, it offers empirical evidence on the comparative effectiveness of deep learning and traditional machine learning approaches for electoral forecasting, highlighting their strengths and limitations. Third, it provides actionable policy insights into how digital engagement can be leveraged as a participatory governance tool, informing electoral transparency and democratic consolidation.

Table 1: Summary of Empirical Review

Authors/Year	Focus	Dataset & Size	Method	Strength (Accuracy)	Weakness (Research Gap)
(Aliyu et al., 2024)	Sentiment Analysis in Low-Resource Settings	Social media	Transfer learning	Effective for low-resource contexts	Scarcity of annotated datasets
(Ciuverca, 2024)	Sentiment Analysis of AI Ethics	Web of Science Clarivate database	Latent Dirichlet Allocation (LDA)	Increasing scholarly focus on AI ethics	No quantitative accuracy metrics provided
(Gong et al., 2024)	Multimodal sentiment Analysis Method	MVSA-datasets	Multimodal joint learning	76.98% MVSA-Single 75.3% MVSA-Multiple	Required improve accuracy

Received: 20-08-2025 / Accepted: 12-09-2025 / Published: 27-04-2026

https://doi.org/10.70118/ujet.2026.0302.06

(Barik et al., 2023)	LSTM-DGWO-Based Sentiment Analysis	10,000 app reviews	LSTM-DGWO	98.89%	Inability to manage optimisation issues
(Oyebode & Orji, 2019)	Social Media Sentiment Analysis	118,421 posts	VADER-EXT, VADER, TextBlob	Overall precision (81.6%)	No evaluation of advanced deep learning models
(Alade & Nwankpa, 2022)	Sentiment Analysis of Nigerian Students	4,016 tweets	UML and Naive Bayes	63%	Small dataset limitations
(Onyenwe et al., 2022)	Location-based Sentiment Analysis	583,816 tweets	TextBlob, SentiWordNet and VADER	N/A	Missing insights into how real-time events affect sentiment polarity.
(Chaudhry et al., 2021)	Sentiment Analysis of before and after Elections	18,432,811 tweets	Naïve Bayes	94.58%	Lacks model diversity
(Alsayat, 2022)	Improving Sentiment Analysis for Social Media Applications	Twitter (18K), Amazon (34.6M), Yelp (299K)	Ensemble Deep Learning Language Model	90.25% 95.70% 96.66%	Narrow DL technique scope
(Caballero, 2021)	Predicting the 2020 US Presidential Election	3,200 tweets	VADER	achieves an MAE of 1.5%	Insufficient data volume

2.0 Methodology

This study adopts a mixed-methods research design, integrating quantitative computational analysis with qualitative expert insights to assess the predictive potential of social media sentiment in electoral contexts. The methodological framework, depicted in Figure 1, comprises five sequential computational phases: data collection, pre-processing, feature extraction, and model development and evaluation, followed by a qualitative validation phase through stakeholder interviews. This design enables triangulation of results, combining large-scale quantitative evidence with context-rich qualitative perspectives to ensure both methodological rigor and practical relevance.

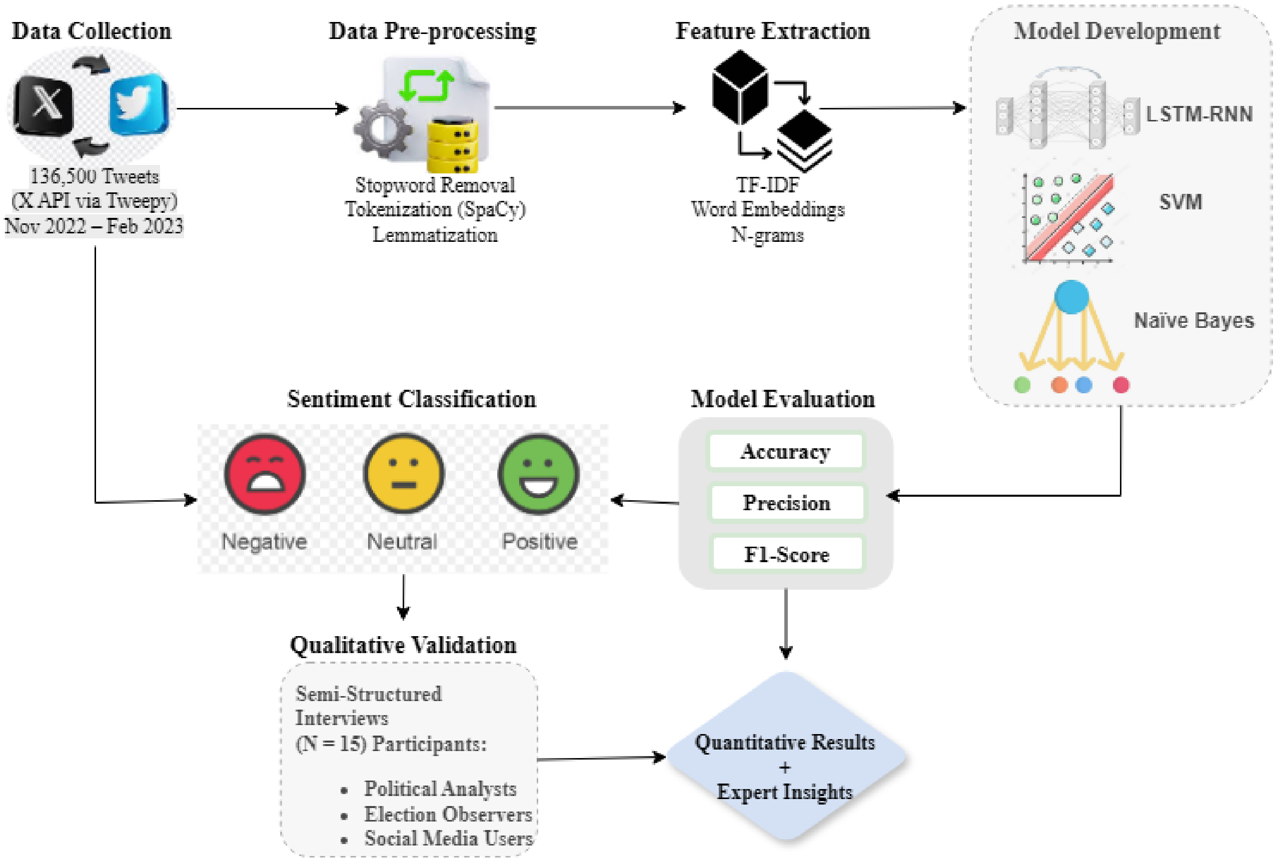


Figure 1: Methodological Framework

2.1 Data Collection

A corpus of 136,500 tweets was compiled via the X API using Tweepy, an open-source Python library, over the three-month campaign period preceding Nigeria's 2023 presidential election (24 November 2022 to 24 February 2023). The dataset targeted political discourse relating to the three leading candidates and their parties, employing both hashtags (e.g., #Atiku, #PDP, #Tinubu, #APC, #PeterObi, #LabourParty) and candidate-specific keywords to maximize coverage. The collection strategy ensured linguistic homogeneity by retaining only English-language tweets and was conducted in compliance with X's rate limits and ethical guidelines, excluding tweets from deleted or private accounts to mitigate potential selection bias.

2.2 Data Pre-processing

The raw dataset underwent a rigorous five-stage NLP pipeline designed to minimize noise, standardize linguistic structure, and enhance feature discriminability for sentiment classification. First, data cleaning involved removing duplicates, retweet markers (RT), @-mentions, non-alphanumeric characters, and other metadata irrelevant to sentiment polarity. Second, normalization included lowercasing (case folding) and removal of stop words using the Natural Language Toolkit (NLTK) English stop word list to focus analysis on semantically meaningful terms. Third, tokenization was performed using SpaCy's tokenizer, segmenting tweets into unigrams and bigrams for granular sentiment pattern detection. Fourth, lemmatization standardized inflected forms to their base lemmas (e.g., "elections" → "election") using WordNet's morphy algorithm. Finally, spelling correction addressed orthographic variations, particularly Nigerian Pidgin and political slang, via a custom lexicon curated from prior political discourse datasets.

2.3 Feature Extraction

Two complementary feature representation strategies were implemented to transform the pre-processed tweets into machine-interpretable vectors suitable for model training. The first approach, term frequency–inverse document frequency (TF-IDF) vectorization, quantified the relative importance of terms by balancing their intra-document frequency with their inverse frequency across the corpus, thereby deprioritizing common but low-information words such as "Nigeria" or "election". The second approach employed randomly initialized word embeddings with 300-dimensional dense vectors trained in-domain to capture semantic and syntactic relationships specific to Nigerian political discourse. This embedding approach enabled the representation of context-dependent associations, such as linking "PVC" (Permanent Voter Card) to electoral participation, and clustered semantically related terms within the embedding space.

Received: 20-08-2025 / Accepted: 12-09-2025 / Published: 27-04-2026

https://doi.org/10.70118/ujet.2026.0302.06

2.4 Model Development

The quantitative modelling phase involved training three distinct classifiers: LSTM-RNN, SVM, and Naïve Bayes, on the extracted feature sets. The LSTM-RNN model was selected for its proven ability to capture sequential dependencies and contextual information in text, making it well-suited to political discourse characterised by evolving narratives and contextual sentiment shifts. The SVM and Naïve Bayes models served as baseline comparators, enabling evaluation of deep learning's marginal predictive advantage over traditional machine learning approaches.

2.5 Qualitative Validation

To complement the computational analysis, semi-structured interviews were conducted with a purposive sample (N = 15) of social media users, political analysts, and election observers. These interviews explored participants' perceptions of political discourse on X, their interpretations of sentiment dynamics, and their views on the credibility of social media as a predictive indicator for election results. Qualitative data were thematically coded and compared with quantitative findings to identify convergences and divergences, enabling triangulation and increasing the ecological validity of the study's conclusions. This integration of qualitative perspectives ensured that model-driven inferences were grounded in the lived realities of political communication in Nigeria's digital public sphere.

2.6 Proposed Models Architecture

LSTM-RNNs are a special kind of RNN, capable of learning long-term dependencies in data. They are particularly well-suited for processing and making predictions based on time series data or sequential data, such as text for sentiment analysis (Waqas & Humphries, 2024). LSTM addresses the vanishing gradient issue in RNNs by incorporating gates that regulate information flow. The LSTM-RNN architecture is composed of a single unit, the memory unit (also referred to as the LSTM unit). This unit consists of four feed-forward neural networks, each comprising an input layer and an output layer, with full connectivity between input and output neurons. This results in four fully connected layers within the LSTM unit. Three of these neural networks are responsible for information selection, functioning as gates: the forget gate, input gate, and output gate. Figure 2 illustrates the data flow through the storage unit, controlled by these gates.

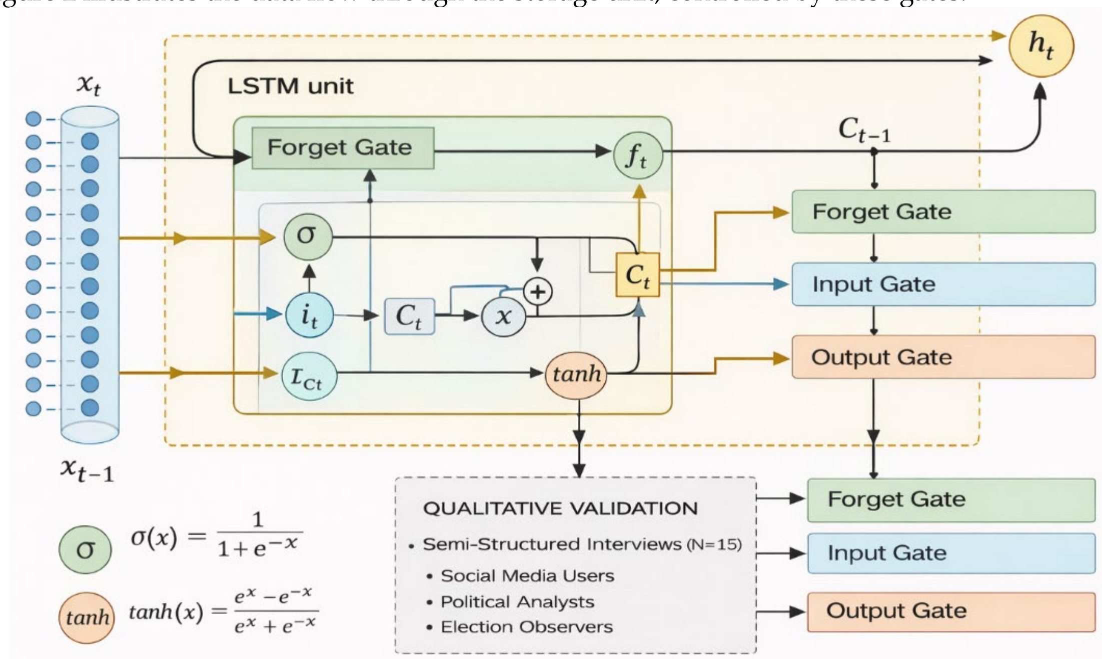


Figure 2: Architecture of the LSTM-RNN

The LSTM unit utilizes both short-term memory (h) and external input (x) to update long-term memory (cell state, c). Subsequently, it employs the long-term memory (c) to update the short-term memory (hidden state, h). The hidden state (h) determined at time t serves as the output of the LSTM unit at that instant, representing the behaviour evaluated for performance assessment. The three gates (forget gate, input gate, and output gate) function as information selectors, generating selector vectors through neural networks with sigmoid activation functions in the output layer. The forget gate, as shown in 'equation 1' is the first activity in the LSTM unit, decides what information to discard from the previous cell state vector based on X_t and H_{t-1} (Venna et al., 2019). This decision yields a selector vector, calculated using the following equation:

$$f_t = \sigma\left(X_t * U_f + H_{t-1} * W_f\right) \quad (1)$$

https://doi.org/10.70118/ujet.2026.0302.06

WhereX_t: Input to the current time. U_f: Weight associated with the input. H_{t-1}: The hidden state of the previous time. W_f: It is the weight matrix associated with the hidden state. The input gate produces a selector vector, which is element-wise multiplied by the candidate vector. The input gate's equation is as follows:

$$i_t = \sigma(X_t * U_i + H_{t-1} * W_i) \quad (2)$$

The output gate determines the value of the hidden state outputted by the LSTM at time t and received as input at time t+1. The output gate utilizes the sigmoid function as the activation function for its output neurons. This gate's equation is similar to those of forget and input gate:

$$o_t = \sigma(X_t * U_o + H_{t-1} * W_o) \quad (3)$$

Where o_t represents the output gate's output, σ denotes the sigmoid function, W_o is the learnable weight matrix, and b_o is the bias term.

SVMs are supervised learning models that learn a decision boundary to separate data points belonging to different classes (Du et al., 2024). The algorithm is applied for classifying sentiment labels from X data and measures the accuracy of the classified data. In which, the labelled tweets from the data set are trained using the classifier. The classification obtained shows the accuracy of the data set. It aim to find the optimal dividing line (hyper-plane) to separate data points (e.g., positive/negative tweets)(Ghaddar & Naoum-Sawaya, 2018). Figure 3 illustrates the SVM architectural framework, showcasing the internal structure and mechanics of the algorithm as it performs classification tasks.

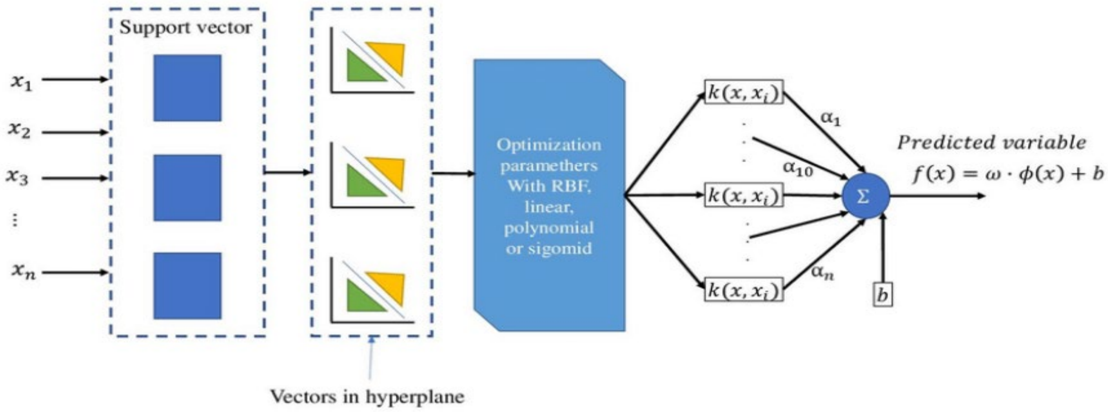


Figure 3: Architecture of SVM model (Deo et al., 2016).

Naive Bayes, according to (Berrar, 2019) is a probabilistic classifier based on Bayes' theorem. It uses Bayes statistics while assuming that features are statistically independent of each other. Due to this assumption, Naive Bayes can learn high-dimensional data with minimal training. It was adopted for classification since the X data is not labelled; therefore, obtaining training data is not forthright. Moreover, Naive Bayes is scalable and is very lightweight. Since tweet data grows steadily with time, it is one of the most suitable classifiers with stable and predictable results. Figure 4 displays the Naïve Bayes and Text Classification Architectural Pipeline, highlighting the sequence of steps involved in the Naive Bayes algorithm for classifying text documents based on various features. Mathematically, it calculates the probability of a class given a set of features as the product of individual feature probabilities conditioned on the class. Despite its "naive" assumption, Naive Bayes often performs surprisingly well, especially for text classification tasks. Naive Bayes equations can also be expressed as follows:

$$\mathrm{P(y)} = \frac{p(x) * p\left(\frac{y}{x}\right)}{p(y)} \quad (4)$$

$$\mathrm{P\ (xb)} = \frac{d(w,e)(xb))}{sl(d(w,e(x))y)} \quad (5)$$

Received: 20-08-2025 / Accepted: 12-09-2025 / Published: 27-04-2026
https://doi.org/10.70118/ujet.2026.0302.06

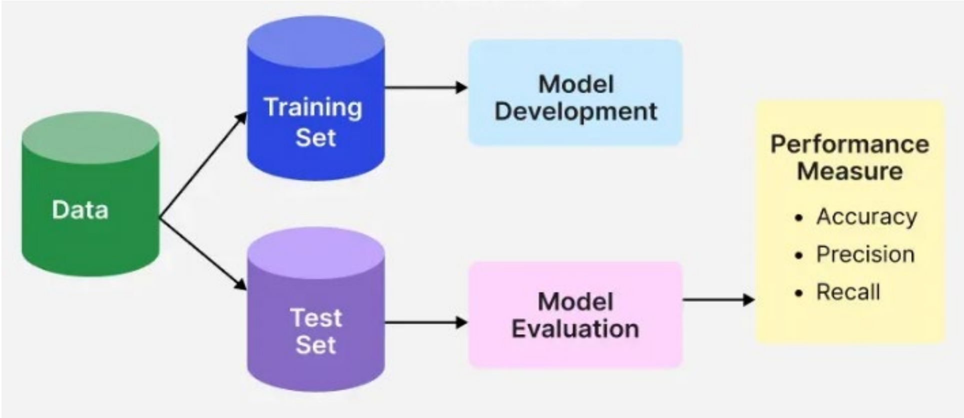


Figure 4: Naïve Bayes Architectural Pipeline. (Keerthana et al., 2024)

2.7 Model Evaluation Metrics

To assess the performance of the models, accuracy, precision, recall, and F1-score were calculated. Accuracy is the percentage of correct predictions. It does not always provide the whole picture of a model's performance, this is where precision, recall, and F1-score offer a more nuanced perspective:

i. Precision: Measured how many of the cases the model predicted as positive were actually positive.

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \tag{6}$$

ii. Recall: Measures how many of the actual positive cases the model correctly identifies (true positives).

$$\text{Recall} = \frac{TP}{(\text{TP} + \text{FN})} \tag{7}$$

iii. F1 Score: The harmonic mean of precision and recall.

$$\text{F1 Score} = \frac{2 * (\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{8}$$

The parameters used in Equations (6), (7), and (8) are defined as follows: TP (True Positives), TN (True Negatives), FP (False Positives), and FN (False Negatives).

3.0 Results and Discussion

3.1 Results

LSTM-RNN as shown in table 2 has the highest performance with an accuracy of 93.8%, precision of 96.6%, recall of 93.6%, and an F1-score of 95.0%. The SVM model achieves moderate performance metrics, with accuracy of 81.7% and F1-score of 86.0%. However, its precision (82.9%) and recall (89.3%) are relatively high, they fall short compared to the LSTM-RNN model. The Naïve Bayes model performs the poorest among the three models, with an accuracy of 74.5%, precision of (83.3%), F1-score of 78.8%, and recall (74.7%).

Table 2: Comparative Performance of Sentiment Analysis Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
LSTM-RNN	93.8	96.6	93.6	95.0
SVM	81.7	82.9	89.3	86.0
Naïve Bayes	74.5	83.3	74.7	78.8

Received: 20-08-2025 / Accepted: 12-09-2025 / Published: 27-04-2026
https://doi.org/10.70118/ujet.2026.0302.06

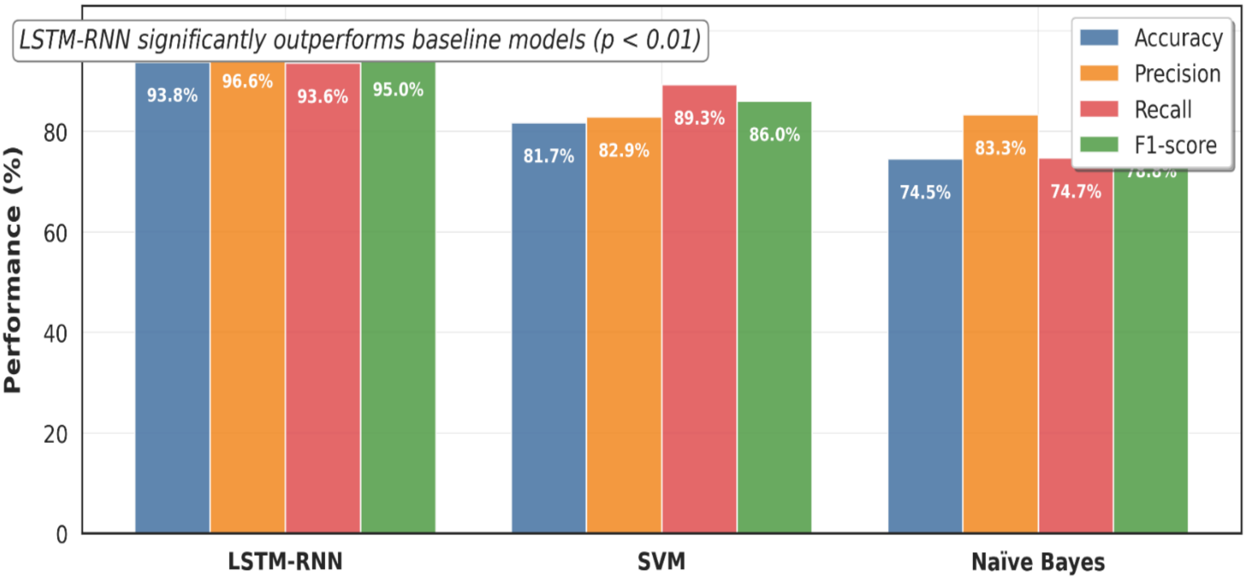


Figure 5: Model Performance Comparison across Metrics

Table 3 compares the sentiment analysis model's predicted percentage of votes for each presidential candidate with the actual election results. The analysis revealed the degree to which sentiment expressed on X aligns with the Nigerian 2023 Presidential election outcomes. The Pearson correlation coefficient (r) revealed a moderately strong negative association ($r = -0.804$). This suggests that candidates with higher predicted positive sentiment tended to receive lower vote percentages in the actual election. However, this correlation lacked statistical significance ($p = 0.405$), indicating that we cannot confidently conclude a definitive relationship between sentiment and voting patterns in this electoral context.

Table 3: Sentiment Analysis Predictions vs. Actual Election Outcomes

Candidate	Political Party	Predicted Result	Actual Outcome
Atiku Abubakar	PDP	46.4%	29%
Bola Ahmed Tinubu	APC	46.1%	37%
Peter Obi	LP	49.6%	25%

Received: 20-08-2025 / Accepted: 12-09-2025 / Published: 27-04-2026
https://doi.org/10.70118/ujet.2026.0302.06

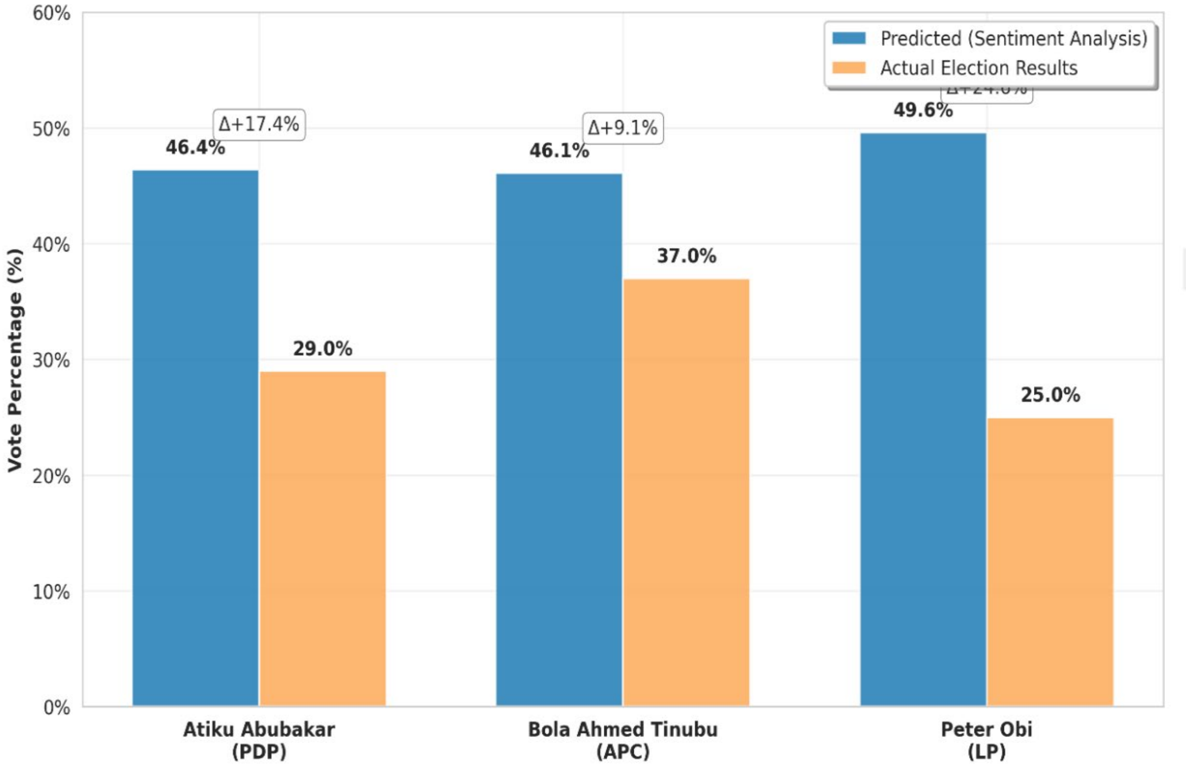
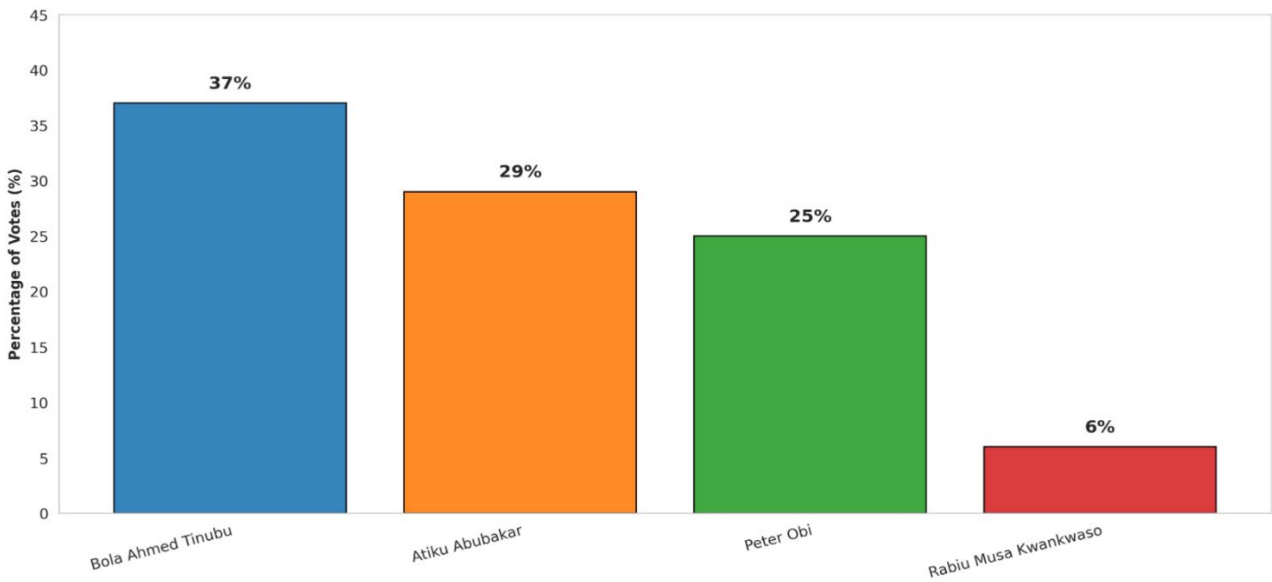


Figure 6: Discrepancy Analysis between model predictions and official INEC results



Candidate	Party	Vote %	Status
Bola Ahmed Tinubu	APC	37%	Winner
Atiku Abubakar	PDP	29%	Runner-up
Peter Obi	LP	25%	Third
Rabiu Musa Kwankwaso	NNPP	6%	Fourth

Figure 7: 2023 Nigerian Presidential Election Results (INEC)

Table 4. Summary of Thematic Analysis of Stakeholder Interviews

Theme	Summary of Stakeholder Insights	Implications for Predictive Potential
Influence on Political Discourse	Social media, particularly X, amplifies youth voices and drives national agenda-setting through viral hashtags.	Confirms that online sentiment reflects political salience and can strengthen electoral prediction validity.
Sentiment Volatility	Public mood on X changes quickly in response to campaign events and controversies.	Supports the development of real-time or streaming sentiment analysis models.
Credibility and Representation	Online discourse is shaped by misinformation, bot activity, and an urban bias, leaving rural voices underrepresented.	Limits social media's role as a standalone predictor; calls for demographic weighting and multimodal integration.
Multilingualism and Code-Mixing	Nigerian Pidgin and regional code-switching influence political talk and challenge standard NLP tools.	Highlights the need for locally adapted deep learning models to improve classification accuracy.
Trust in Predictive Utility	Stakeholders see online sentiment as reflective of electoral mood, but caution that "online noise" does not always translate into votes.	Reinforces the necessity of triangulating sentiment with ground-level political data and polling.
Governance and Transparency	Social media sentiment is seen as a tool for accountability and early detection of voter concerns.	Suggests wider applications of predictive models in governance beyond elections.

As summarised in Table 5, to validate and contextualise the computational findings, semi-structured interviews were conducted with key stakeholders, including social media users, political analysts, and election observers, thereby leveraging methodological triangulation to strengthen the robustness of the study. These qualitative findings not only corroborate the computational analysis but also reveal important limitations that challenge the generalizability of sentiment-derived predictions.

Table 5: Impact of Optimisation on LSTM-RNN Sentiment Analysis Performance

Optimisation Technique	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Baseline Model	90.5	95.9	88.9	92.2
Optimised Model	93.8	96.6	93.6	95.0
Change in Performance	3.3	0.7	4.7	2.8

The optimised LSTM-RNN in Table 5 achieves an accuracy of 93.8%, representing a 3.3% improvement over the baseline model. The precision increases by 0.7% compared to the baseline model, reaching 96.6%. Similarly, the recall improves by 4.7% compared to the baseline model, reaching 93.6%. The F1-score also demonstrates a 2.8% improvement over the baseline model, reaching 95.0%.

Received: 20-08-2025 / Accepted: 12-09-2025 / Published: 27-04-2026

https://doi.org/10.70118/ujet.2026.0302.06

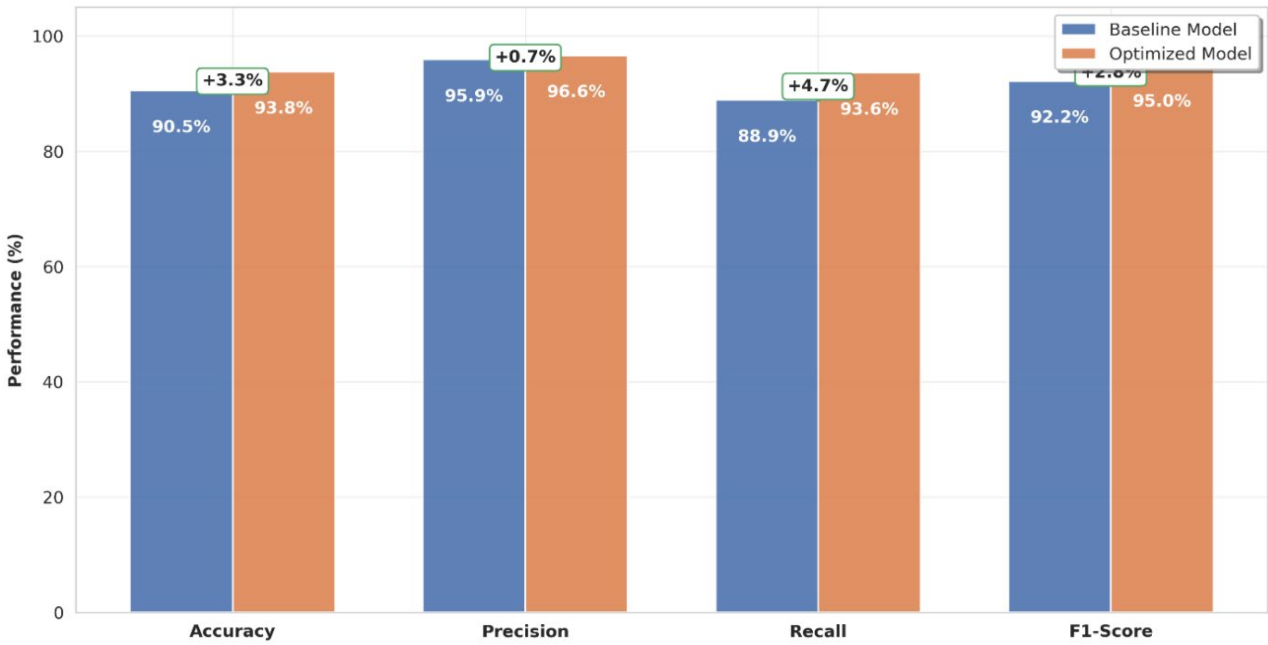


Figure 8: Performance Gains achieved through Hyper-parameter Optimisation

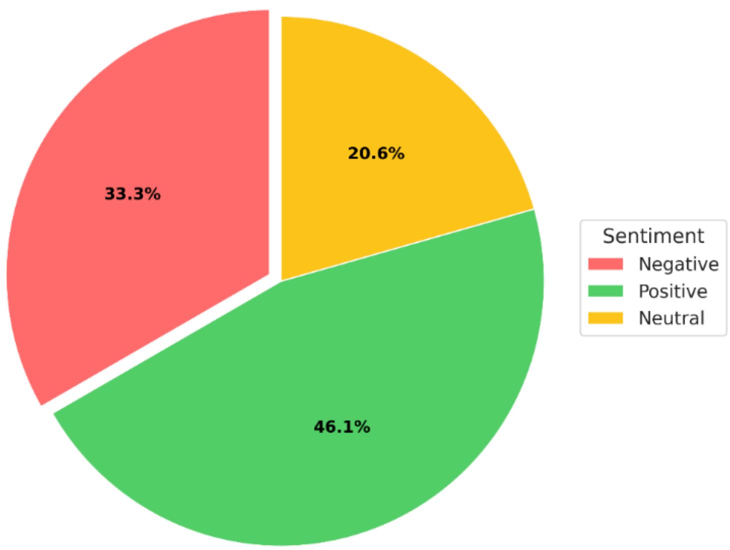


Figure 9: Sentiment Distribution of Bola Ahmed Tinubu

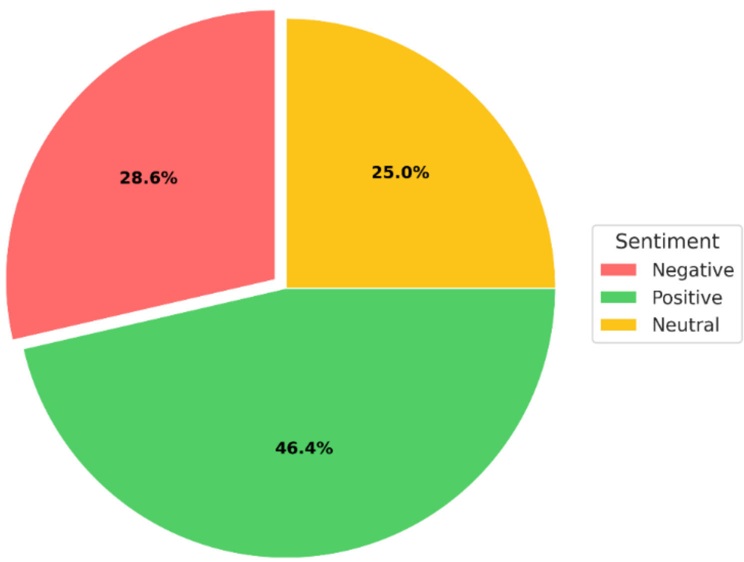


Figure 10: Sentiment Distribution of Atiku Abubakar

Received: 20-08-2025 / Accepted: 12-09-2025 / Published: 27-04-2026
https://doi.org/10.70118/ujet.2026.0302.06

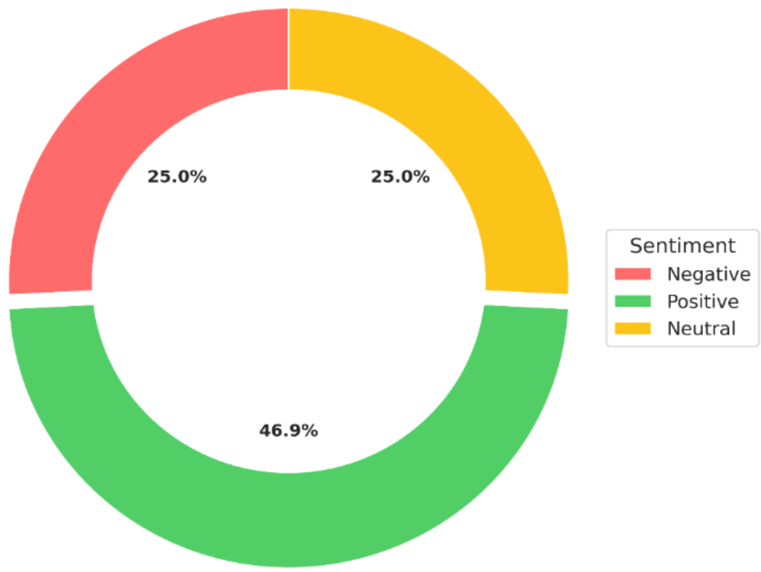


Figure 11: Sentiment Distribution of Peter Obi

3.2 Discussion

The performance assessment of LSTM-RNN, SVM, and Naïve Bayes models revealed distinct variations in their effectiveness for sentiment analysis of X data. The LSTM-RNN model exhibited superior performance across all metrics: accuracy, precision, recall, and F1-score. This superior performance is primarily attributed to its ability to capture sequential dependencies and contextual relationships in text data, which are essential for understanding evolving sentiments on social media (Waqas & Humphries, 2024). In contrast, SVM and Naïve Bayes, while effective for classification tasks, demonstrated limited capability in modelling contextual nuances, resulting in comparatively lower performance. This highlights the importance of aligning model selection with the intrinsic characteristics of textual data (Arévalo-Cordovilla & Peña, 2024).

The moderately strong negative correlation between predicted sentiment scores and actual election results, though not statistically significant, raises concerns about the reliability of sentiment analysis as a standalone predictive tool. Notable discrepancies were observed across candidates: Peter Obi's predicted outcome exceeded the actual result by 24.6 percentage points, while Bola Ahmed Tinubu showed a smaller deviation of 9.1 percentage points. These disparities suggest underlying biases in the data and limitations in capturing the true distribution of voter preferences, particularly in digitally mediated environments. This finding reinforces the need for hybrid predictive frameworks that integrate sentiment data with complementary socio-demographic and electoral variables (Feng et al., 2023).

Model optimisation further contributed to performance improvements, with adjustments such as RMSprop optimisation, increased batch size, and dropout regularisation enhancing training stability and generalisation. These results underscore the importance of hyperparameter tuning in improving deep learning performance, particularly for high-dimensional textual data. Beyond model performance, the findings highlight structural limitations in sentiment-based electoral prediction. The presence of silent voters and disparities in voter turnout introduce disconnect between expressed online sentiment and actual voting behaviour. Additionally, social media data is inherently skewed toward more active and urban populations, limiting its representativeness. These factors collectively constrain the predictive validity of sentiment analysis in electoral contexts.

Thematic analysis of interview data both corroborated and extended the quantitative findings. Participants emphasized the central role of social media in amplifying political discourse, particularly among youths, while also identifying critical challenges. These include the volatility of online sentiment, linguistic complexity arising from Pidgin and code-mixing, and concerns regarding misinformation and bot activity. Such factors introduce noise and bias, further complicating sentiment interpretation.

Overall, the findings demonstrate that while deep learning models, particularly LSTM-RNN, offer strong capabilities for sentiment classification, their predictive application in electoral forecasting remains limited without contextual augmentation. Future research should therefore focus on integrating multimodal data sources and developing more context-aware models to improve robustness and generalisability.

4.0 Conclusion

By addressing the research questions, this study provides a comprehensive evaluation of sentiment analysis approaches for electoral prediction, highlighting the comparative advantage of the LSTM-RNN over traditional models such as SVM and Naïve Bayes. The optimised LSTM-RNN achieved high predictive

Received: 20-08-2025 / Accepted: 12-09-2025 / Published: 27-04-2026
https://doi.org/10.70118/ujet.2026.0302.06

performance (93.8% accuracy and 95.0% F1-score), demonstrating its effectiveness in capturing contextual and sequential patterns in social media data. However, despite this strong classification performance, the observed correlation between sentiment-derived predictions and actual election outcomes was not statistically significant, indicating limitations in the predictive validity of sentiment analysis when used in isolation. This finding underscores the need for more robust frameworks that incorporate contextual, demographic, and behavioural factors to improve predictive reliability. Furthermore, this study highlights critical challenges associated with social media data, including representativeness bias, linguistic complexity, and the influence of non-participatory ("silent") voters. These limitations emphasise that sentiment analysis should be interpreted as a complementary analytical tool rather than a standalone predictor of electoral outcomes. Finally, the study underscores the importance of ethical and responsible deployment of sentiment analysis, ensuring transparency, fairness, and awareness of potential biases in computational models. Future research should focus on integrating multimodal data sources and developing context-aware models to enhance robustness and applicability in real-world electoral settings.

References

Alade, M. S., & Nwankpa, J. M. (2022). Sentiment analysis of Nigerian students' tweets on education: A data mining approach. *International Journal of Computer*, 45(1), 1–27.

Aliyu, Y., Sarlan, A., Usman Danyaro, K., Rahman, A. S. B. A., & Abdullahi, M. (2024). Sentiment analysis in low-resource settings: A comprehensive review of approaches, languages, and data sources. *IEEE Access*, 12, 66883–66909. https://doi.org/10.1109/ACCESS.2024.3398635

Alsayat, A. (2022). Improving sentiment analysis for social media applications using an ensemble deep learning language model. *Arabian Journal for Science and Engineering*, 47(2), 2499–2511. https://doi.org/10.1007/s13369-021-06227-w

Arévalo-Cordovilla, F. E., & Peña, M. (2024). Comparative analysis of machine learning models for predicting student success in online programming courses: A study based on LMS data and external factors. *Mathematics*, 12(20), 3272. https://doi.org/10.3390/math12203272

Attai, K., Asuquo, D., Okonny, K. E., Johnson, E. A., Bassey, A., John, A., Bardi, I., Iroanwusi, C., & Michael, O. (2024). Sentiment analysis of Twitter discourse on the 2023 Nigerian general elections. *European Journal of Computer Science and Information Technology*, 12(4), 18–35. https://doi.org/10.37745/ejcsit.2013/vol12n41835

Barik, K., Misra, S., Ray, A. K., & Bokolo, A. (2023). LSTM-DGWO-based sentiment analysis framework for analyzing online customer reviews. *Computational Intelligence and Neuroscience*, 2023, 6348831. https://doi.org/10.1155/2023/6348831

Berrar, D. (2019). Bayes' theorem and Naïve Bayes classifier. In *Encyclopedia of Bioinformatics and Computational Biology* (pp. 403–412). Elsevier. https://doi.org/10.1016/B978-0-12-809633-8.20473-1

Caballero, M. (2021). Predicting the 2020 US presidential election with Twitter. In *Computer Science & Information Technology (CS & IT)*, 53–65. https://doi.org/10.5121/csit.2021.111006

Chaudhry, H. N., Javed, Y., Kulsoom, F., Mehmood, Z., Khan, Z. I., Shoaib, U., & Janjua, S. H. (2021). Sentiment analysis of before and after elections: Twitter data of U.S. election 2020. *Electronics*, 10(17), 2082. https://doi.org/10.3390/electronics10172082

Ciuverca, A.-C.-D. (2024). Sentiment analysis of research on AI ethics: A web-based study. In *Proceedings of the International Conference on Economics and Social Sciences*, **Bucharest, Romania**. https://doi.org/10.24818/ICESS/2024/060

Deo, R. C., Wen, X., & Qi, F. (2016). A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset. *Applied Energy*, 168, 568–593. https://doi.org/10.1016/j.apenergy.2016.01.130

Du, K.-L., Jiang, B., Lu, J., Hua, J., & Swamy, M. N. S. (2024). Exploring kernel machines and support vector machines: Principles, techniques, and future directions. *Mathematics*, 12(24), 3935. https://doi.org/10.3390/math12243935

Feng, G., Cai, H., Chen, K., & Li, Z. (2023). A hybrid method of sentiment analysis and machine learning algorithm for the U.S. presidential election forecasting. *arXiv preprint arXiv:2312.05584*. https://doi.org/10.48550/arXiv.2312.05584

Garba, K., Kolajo, T., & Agbogun, J. B. (2024). A transformer-based approach to Nigerian Pidgin text generation. *International Journal of Speech Technology*, 27(4), 1027–1037. https://doi.org/10.1007/s10772-024-10136-2

Ghaddar, B., & Naoum-Sawaya, J. (2018). High dimensional data classification and feature selection using support vector machines. *European Journal of Operational Research*, 265(3), 993–1004. https://doi.org/10.1016/j.ejor.2017.08.040

Gong, L., He, X., & Yang, J. (2024). An image-text sentiment analysis method using multi-channel multi-modal joint learning. *Applied Artificial Intelligence*, 38(1), 2371712. https://doi.org/10.1080/08839514.2024.2371712

Harem Fatima, & Skander Gasmi. (2024). Optimization strategies in machine learning for improved big data analysis and natural language processing. *Preprint*. https://doi.org/10.13140/RG.2.2.28130.75208

Inuwa-Dutse, I. (2025). NaijaNLP: A survey of Nigerian low-resource languages (Version 2). *arXiv preprint arXiv:2502.19784*. https://doi.org/10.48550/arXiv.2502.19784

Islam, Md. S., Kabir, M. N., Ghani, N. A., Zamli, K. Z., Zulkifli, N. S. A., Rahman, Md. M., & Moni, M. A. (2024). Challenges and future in deep learning for sentiment analysis: A comprehensive review and a proposed novel hybrid approach. *Artificial Intelligence Review*, 57(3), 62. https://doi.org/10.1007/s10462-023-10651-9

Jose Gonzalez-Gomez, L., Margarita Hernandez-Munoz, S., Borja, A., Daniel Azofeifa, J., Noguez, J., & Caratozzolo, P. (2024). Analyzing natural language processing techniques to extract meaningful information on skills acquisition from textual content. *IEEE Access*, 12, 139742–139757. https://doi.org/10.1109/ACCESS.2024.3465409

Keerthana, B., Vamsinath, J., Kumari, C. S., Appaji, S. V. S., Rani, P. P., & Chilukuri, S. (2024). Machine learning techniques for medicinal leaf prediction and disease identification. *International Journal of Experimental Research and Review*, 42, 320–327. https://doi.org/10.52756/ijerr.2024.v42.028

Muñoz, P., Bellogín, A., Barba-Rojas, R., & Díez, F. (2024). Quantifying polarization in online political discourse. *EPJ Data Science*, 13(1), 39. https://doi.org/10.1140/epjds/s13688-024-00480-3

Myilvahanan, K., P, Y., Pasha, S., Ismail, M., & Tharun, V. (2023). A study on election prediction using machine learning techniques. In *Proceedings of the 3rd International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, **India**, 1518–1520. https://doi.org/10.1109/ICAIS56108.2023.10073693

Nigerian Communications Commission (NCC). (2020). *Nigerian National Broadband Plan 2020–2025*. Abuja, Nigeria.

Onyenwe, I. E., Nwagbo, S. N. C., Orji, R., Onyedinma, E., Ikechukwu-Onyenwe, O., Nwafor, C. A., & Agbata, O. (2022). Location-based sentiment analysis of 2019 Nigeria presidential election using a voting ensemble approach. *Preprint*. https://doi.org/10.21203/rs.3.rs-2104488/v1

Oyebode, O., & Orji, R. (2019). Social media and sentiment analysis: The Nigeria presidential election 2019. In *Proceedings of the IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, **Canada**, 0140–0146. https://doi.org/10.1109/IEMCON.2019.8936139

Oyewola, D. O., Oladimeji, L. A., Julius, S. O., Kachalla, L. B., & Dada, E. G. (2023). Optimizing sentiment analysis of Nigerian 2023 presidential election using two-stage residual long short term memory. *Heliyon*, 9(4), e14836. https://doi.org/10.1016/j.heliyon.2023.e14836

Paul Anule, A., & Jennifer Ifeoma, C. (2025). Application of support vector machine for effective prediction of election for sentiment analysis. *Middle East Journal of Applied Science & Technology*, 8(1), 115–131. https://doi.org/10.46431/MEJAST.2025.8111

Rita, P., António, N., & Afonso, A. P. (2023). Social media discourse and voting decisions influence: Sentiment analysis in tweets during an electoral period. *Social Network Analysis and Mining*, 13(1), 46. https://doi.org/10.1007/s13278-023-01048-1

Venna, S. R., Tavanaei, A., Gottumukkala, R. N., Raghavan, V. V., Maida, A. S., & Nichols, S. (2019). A novel data-driven model for real-time influenza forecasting. *IEEE Access*, 7, 7691–7701. https://doi.org/10.1109/ACCESS.2018.2888585

Waqas, M., & Humphries, U. W. (2024). A critical review of RNN and LSTM variants in hydrological time series predictions. *MethodsX*, 13, 102946. https://doi.org/10.1016/j.mex.2024.102946

Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *WIREs Data Mining and Knowledge Discovery*, 8(4), e1253. https://doi.org/10.1002/widm.1253

Received: 20-08-2025 / Accepted: 12-09-2025 / Published: 27-04-2026
https://doi.org/10.70118/ujet.2026.0302.06