Research Article/Review Article

Journal of Computing, Science &Technology. Vol. 2, 2024

Journal of Computing, Science &Technology

https://focjournal.unidel.edu.ng/

editorjcst@unidel.edu.ng

ISSN: 3043-6818 Print

https://focjournal.unidel.edu.ng

Sentiment Analysis in Election Domain: A Systematic Review

Uduak A. Umoh[1, 5], Victor E. Ekong[2, 5], Temitope Joel Fakiyesi[3, 5]and Samuel Robinson[4, 5]

[1]Department of Information System, Faculty of Computing, University of Uyo, Uyo, Akwa Ibom State, Nigeria

[2]Department of Software Engineering, Faculty of Computing, University of Uyo, Uyo, Akwa Ibom State, Nigeria

[3]Department of Computer Science, Faculty of Computing, University of Uyo, Uyo, Akwa Ibom State, Nigeria

[4]Department of Cyber Security, Faculty of Computing, University of Uyo, Uyo, Akwa Ibom State, Nigeria

[5] TETFund Centre of Excellence in Computational Intelligence Research, University of Uyo, Nigeria.

uduakumoh@uniuyo.edu.ng[1], victoreekong@uniuyo.edu.ng[2], temitopefakiyesi@uniuyo.edu.ng[3], samuelrobinson@uniuyo.edu.ng[4]

Corresponding Author's Email*:* victoreekong@uniuyo.edu.ng

Article Info

Date Received: 03-09-2024

Date Accepted: 12-11-2024

Keywords:

Election, Machine Learning, Sentiment, Sentiment analysis, Twitter

ABSTRACT

This systematic literature review (SLR) aims to comprehensively survey existing research on election results prediction, focusing on various approaches, methods, and challenges encountered in the field. It seeks to identify data-driven intelligent models and techniques used in sentiment analysis for predicting election outcomes, comparing their strengths and weaknesses to traditional polling methods. The review examines diverse data sources, including social media, news, and surveys, with a particular emphasis on Twitter as the most utilized source for sentiment analysis in the election domain as established in the studies of a hybrid method of sentiment analysis and machine learning algorithm for the US presidential election forecasting by Feng et al.,2023. Twitter-derived datasets are deemed reliable and sufficient for election outcome prediction. Employing a systematic approach, the review excludes duplicate articles, non-English works, and emphasizes relevant sources. Findings indicate prevalent use of machine learning tools and techniques such as text preprocessing, feature extraction, and supervised learning algorithms like Support Vector Machines (SVM) and Naive Bayes for sentiment classification tasks in election sentiment analysis with Twitter emerging as the most effective data source for sentiment analysis in election prediction. Sentiment analysis has notably enhanced election-related functions such as issue prioritization, real-time monitoring, crisis management, citizen predictive power, governance transparency, data-driven campaign strategy, and election integrity, thereby facilitating effective governance.

1.0 INTRODUCTION

1.1 Overview of Sentiment Analysis

Sentiment refers to the emotional or subjective tone expressed in a piece of text, speech, or communication [1, 2]. It reflects the attitude, opinion, or feeling of the author or speaker toward a particular subject, topic, or situation. Sentiment analysis (SA) also known as opinion mining has been applied in entertainment and media production industry to assess audience reactions to movies, TV shows, music, and other forms of entertainment so as to guide content creation and production based on viewer sentiment. In business and marketing, SA has been applied to monitor customer feedback and sentiment on social media and reviews, to assess brand perception and customer satisfaction as well as to inform marketing strategies and product development based on consumer opinions [3, 4]. In customer service, SA has the advantage of automating the analysis of customer support interactions by identifying and addressing customer issues and concerns more efficiently. In financial markets, SA has been used to analyze news articles and social media sentiment to gauge market sentiment by taking appropriate trading decisions and risk management in financial trading. Sentiment analysis as also been applied in hospitality and tourism sector to assess guest reviews and suggestions to improve hotel and travel services aiming at tailoring marketing campaigns and guest experiences based on feedback. The systematic literature review investigated in this study focuses on the use of sentiment analysis in the field of election. In spite of all the studies that have been carried out both past and present in SA on election results prediction, there has not been a research proposal that introduce a systematic literature review of SA in election domain.

1.2 Applications of Sentiment Analysis Across Industries

Sentiment analysis holds immense significance in electoral contexts, where it serves as a powerful tool for understanding and harnessing public sentiment and opinions. First and foremost, sentiment analysis provides valuable insights into the minds of voters. By examining the sentiments expressed in various media channels, such as social media, news outlets, and online forums, political campaigns and candidates gain a deeper understanding of how their messages and policies are perceived by the electorate [5]. Secondly, sentiment analysis plays a pivotal role in real-time monitoring during election campaigns as it offers a dynamic feedback mechanism that allows political actors to respond swiftly to emerging issues and concerns [6]. Also, sentiment analysis contributes to the predictive aspect of elections by analyzing sentiment trends over time and across different demographics. Promotes transparency by making public sentiment visible, holding candidates and elected officials accountable for addressing constituents' concerns [7]. Sentiment analysis facilitates effective communication between elected officials and constituents by addressing their expressed sentiments and concerns. In media strategy, SA helps in evaluating how campaign messages are portrayed in the media and shaping media strategies accordingly as well as in research and analysis where it provides valuable data for academic research on political behavior, public opinion, and sentiment trends. Sentiment analysis also improves civic engagement by giving voters a platform to express their opinions and concerns [8, 9]. This predictive power supplements traditional polling methods and provides a more holistic view of voter sentiment [10, 11]. Analysis of sentiments in elections can provide an indispensable tool that will enhance the democratic process by fostering transparency, promoting civic engagement, and ultimately helping political stakeholders better serve the electorate's needs and aspirations [12].

1.3 Techniques and Tools for Sentiment Analysis

Sentiment analysis involves using natural language processing (NLP) and machine learning (ML) techniques to determine the emotional tone behind a piece of text [13]. In the context of a country's presidential elections, SA can be used to gauge public opinion, identify trends, and understand the sentiment of voters, campaigners and media coverage. Many researchers have employed different types of techniques over time in sentiment analysis, advancing from lexicon-based methods, which rely on dictionaries of words and their associated semantic orientation, to machine learning techniques that automatically extract word relationships where machine learning operates without explicit programming but instead rely on the learning experience [14].

Over the years, the conventional methods and tools that have been utilized for SA include Valence Aware Dictionary and sEntiment Reasoning (VADER) for social media text, Linguistic Inquiry and Word Count (LIWC) for extracting emotional and cognitive dimensions, SenticNet for nuanced SA with multiple dimensions, emoji analysis for context-rich emotion interpretation also hybrid approaches that combine lexicon-based and machine learning methods using pretrained models, crowd-sourced or expert-annotated datasets for specialized SA and SA APIs such as Google Cloud Natural Language API and Microsoft Azure Text Analytics, all contributing to a diverse set of strategies to capture sentiment nuances in contexts like presidential elections [15].

Machine learning tools for SA are designed to automatically determine the sentiment or emotional tone expressed in a piece of text such as a review, comment or social media post. These tools use various machine learning algorithms and techniques to classify text into categories like positive, negative or neutral sentiment [16, 17, 18, 19]. Common machine learning tools and techniques for SA involve preprocessing steps such as text tokenization, stopword removal, stemming and lemmatization to prepare text data. Feature extraction methods, such as Bag of Words (BoW) or Term Frequency-Inverse Document Frequency (TF-IDF) are used to convert text into numerical representations. Supervised learning algorithms like Support Vector Machines (SVM), Naive Bayes and deep learning models, particularly recurrent neural networks (RNNs) and transformer-based architectures like Bidirectional Encoder Representations from Transformers (BERT) which was developed by Google and released in 2018 are widely employed for sentiment classification tasks.

Ensembling methods such as Random Forests or Gradient Boosting and fine-tuning techniques to adapt pre-trained models to specific SA tasks are commonly used to enhance performance. Evaluation metrics like accuracy, precision, recall, F1-score, and AUC-ROC are used to assess model performance for comparison while techniques such as cross-validation assist in robust model validation and hyperparameter tuning. ML for SA are known for its efficiency, scalability and adaptability which propel in processing vast datasets by capturing complex language patterns [20, 21]. It as well achieved high accuracy, particularly with context-aware deep learning models [22] and can adapt to diverse domains through fine-tuning, also in using probabilistic and statistical methods [23]. However, machine learning strongly relies on data quality, struggles with ambiguity and sarcastic language which may delay in highly specialized domains.

1.4 Sentiment Analysis in Elections

The concept of electing leaders was dated back to ancient times, but the modern idea of presidential elections gained prominence with the development of democratic systems in the late 17th and early 18th centuries. The United States is often credited with holding the world's first modern presidential election in 1788-1789 when George Washington was elected as the first President of the United States. Over time, various countries around the world adopted the idea of electing their leaders through direct or indirect elections. The methods and processes of presidential elections can vary significantly, ranging from popular votes to electoral college systems. [24] considers SA as an accurate predictive tool when the author rightly predicted the outcome of Colombia's presidential election. Several other studies had reflected the importance of SA for the prediction of an election results as identified from the works of [25, 26, 27, 28, 29, 30, 31, 32, 33].

2. Systematic Review Design

The systematic review attempts to accomplish the following objectives; (i) identify and categorize the various data-driven models, algorithms, and techniques employed in SA within the context of election results prediction, (ii) explores the role of diverse data sources in SA and assess their impact on the accuracy and reliability of election result predictions, (iii) evaluation of the advantages and limitations of data-driven intelligent models in comparison to traditional polling techniques for predicting election results; and (iv) identify the most used technique and data source by the SA in electoral contexts.

2.1 Research Questions

In the context of this systematic review, four guidelines have been developed in the form of research questions to align with the objectives:

RQ1. What are the key data-driven intelligent models and techniques used in SA for predicting election results?

RQ2. How do different data sources contribute to the effectiveness of SA in predicting election outcomes?

RQ3. What are the strengths and weaknesses of data-driven SA models in the electoral domain, and how do they compare to traditional polling methods?

RQ4. What advantages will SA carried-out on election domain offer?

2.2 Digital Libraries

The digital libraries that were used to execute the systematic literature review of this work is shown in Table 1. This table also presents the type of bibliographic source, the period of publication, language, and search strategy used in this study. As it can be noticed, a keyword-based search strategy was used to search for research works targeting on SA in election domain. This proposed guideline is described in detail in the adjoining section.

2.3 Search Strategy

In other to answer the formulated research questions, we use a keyword-based search strategy. By this approach, we identified a set of keywords related to SA in election domain as well as synonyms for the set of keywords identified. Once these terms were defined, we combined these terms with the connectors "AND" and "OR", resulting in the subsequent search chain below:

(sentiment analysis) AND (sentiment classification OR sentiment analysis techniques OR opinion mining OR election domain) AND/OR (election results prediction) AND/OR (country) AND/OR (presidential).

Table 1. Digital library sources

Language	Approach	Duration	Type	Digital library source
English	Keywords	Jan. 2014 to Jun. 2023	scholar articles, scientific journals, e-books, conferences, online workshops	ACM Digital Library
				Elsevier
				Google Scholar
				IEEE Xplore
				Springer Direct
				Springer Link Wiley
				Springer Nature

2.4 Exclusion criteria

Any research article that does not correspond with the search criteria and (or) not directly related to SA and election domain were excluded. Additional exclusion criteria include;

i. Duplicated research articles derived from digital libraries.
ii. Existing masters and doctoral thesis.
iii. Any research articles written in a language other than English language.

2.5 Inclusion criteria

Any research article that correspond with the search criteria and directly related an election domain are:

i. Any presidential election held within a country between 2014 and 2023.
ii. SA analysis tools and techniques for classification
iv. the Positive, Negative and Neutral predictions
v. Articles published in peer reviewed journal in English language

3. Systematic Review Implementation

In this section, the systematic review implementation is presented. It begins with a search for works which are related with SA and election domain from the selected digital library sources by taking into consideration the exclusion and inclusion criteria. The review responses to the research questions are discussed in the subsequent sections.

3.1 RQ1: What are the key data-driven intelligent models and techniques used in SA for predicting election results?

Data-driven intelligent models and techniques applied in SA for predicting election results encompass a range of advanced methodologies. These include natural language processing (NLP) models such as recurrent neural networks (RNNs) and transformer-based models like BERT, have excelled at capturing contextual information and understanding nuances in text. Machine learning algorithms such as Support Vector Machines (SVM), Random Forest, and Gradient Descent are leveraged for their ability to classify sentiments effectively. Hybrid approaches combine lexicon-based analysis with machine learning techniques.

These models are typically trained on extensive labeled datasets to discern sentiments with high accuracy, making them pivotal tools in forecasting election results outcomes based on public sentiment expressed across various media sources.

In Figure 1, a data-driven intelligent model and technique commonly used in SA for election results prediction is presented in five phases namely: data-driven intelligent model systematic review from digital library sources, natural language preprocessing tools, machine learning algorithms, lexicon-based analysis, transformer-based models, hybrid approach and the system model.

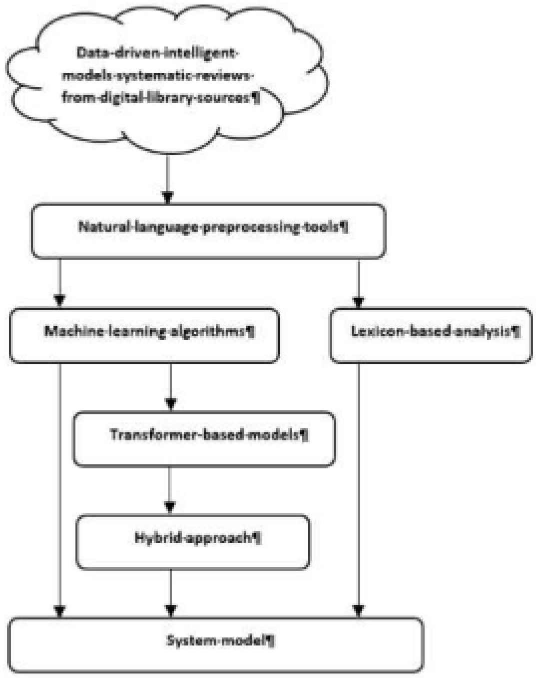

Fig. 1 Data-driven intelligent model systematic review

Table 2 indicate the set of articles analyzed in the systemic review. It shows the authors, year of publication, SA approach, classifier, techniques/tools, data source and performance evaluation metrics used.

Table 2. Sentiment analysis approach, classifier, techniques/tools, data source and performance evaluation metrics.

Author(s)	**Year**	**Sentiment approach**	**Classifier**	**Techniques/Tools**	**Data source**	**Performance evaluation metrics**
Oyewola, et al. (2023)	2023	Supervised machine learning	Probabilistic and linear classifiers	LSTMRNN, BERT and LSVC	Twitter	BERT accuracy = 0.94
Olabanjo, et al. (2023)	2023	Lexicon based with WEKA	Dictionary based approach	BERT	Twitter	precision = 0.83, recall = 0.92 and F1score = 0.91
Barreiro (2020)	2020	Supervised machine learning	Probabilistic and linear classifiers	MAKERRIMER, SVM, Naïve Bayes, Decision Trees, and Logistic Regression	Twitter	MAKERRIMER accuracy = 0.76
Yaqub et al., (2017)	2017	Lexicon based	Dictionary based approach	Statistical analysis	Twitter	Result corelated positively
CerónGuzmán (2016)	2016	Supervised machine leaning	Lexicon based classifiers	Logistic regression and sentiment score	Twitter	Positive Precision = 0.65, Recall = 0.43, F1score = 0.52
Prasetyo, (2014)	2014	Supervised machine learning	Probabilistic and linear classifiers	WEKA	Twitter	Precision = 0.80

From Table 2, we deduce that twitter house provided the most utilized data resource.

3.2 RQ2. How do different data sources, including social media, news, and surveys contribute to the effectiveness of SA in predicting election outcomes?

Various data sources, including social media, news, and surveys, significantly enhance the effectiveness of SA in predicting election outcomes. Social media platforms like Twitter, Facebook, and Instagram provide real-time, unfiltered expressions of public sentiment, thereby enabling timely insights into evolving voter opinions. News articles and reports offer a more structured and comprehensive view to reflect the sentiments expressed by journalists and editorial outlets. Surveys, on the other hand provide structured and targeted sentiment data collected from specific demographic groups by offering valuable insights into voters' preferences in the election domain for SA.

Through carefully crafted questionnaires, surveys collect detailed information on voters' opinions, issues that matter most to them, and their overall attitudes toward candidates and policies. These insights enable sentiment analysts to uncover specific sentiments related to key election issues, discern shifts in voter sentiment over time, and gauge the intensity of voter preferences. Surveys also provide poll statistical information, allowing for a deeper understanding of how sentiment varies among different voter groups. This comprehensive data enriches SA by offering a more nuanced and contextually informed perspective on voter sentiment, aiding in the prediction of election outcomes and the development of effective campaign strategies.

Combining these diverse sources allows SA models to capture a holistic perspective of public sentiment by making it a powerful tool for predicting election results. In summary, social media provides proximity and breadth, news articles offer context, and surveys provide structured depth, all of which contribute to a more accurate understanding of voter sentiment and its potential impact on election outcomes.

3.3 RQ3. Strengths and weaknesses of data-driven SA models in presidential election domain and the comparison with traditional methods?

Data-driven SA models in the electoral domain offer the strength of real-time and largescale data analysis by providing timely insights into evolving voter sentiment and allowing for nuanced sentiment categorization. It can capture public sentiment expressed on social media, news, and various online platforms. However, these models may have weaknesses related to potential bias in the data, limited representativeness of the online population, and challenges in handling sarcasm and context.

Traditional polling methods, on the other hand, offer the strength of random sampling and structured questionnaires, providing statistically valid and representative samples of the electorate. They are well-established and can assess sentiment across diverse demographics. However, they have limitations related to cost, time consuming during data collection, and potential nonresponse bias. Combining both data-driven SA and traditional polling methods can offer a more comprehensive and accurate understanding of public sentiment in the electoral context.

Table 3. Strengths and weaknesses of data-driven SA models in presidential election domain

Author(s)	Year	Country	Sentiment approach	Techniques/Tools
Oyewola, et al. (2023)	2023	Nigeria	Supervised machine learning	LSTMRNN, BERT and LSVC
Kermani and Rasouli (2022)	2022	Iran	Probabilistic classifiers	SNA and ECA
Mohapatra and Mohapatra (2022)	2022	United States of America	Supervised machine learning	BERT
RodríguezIbáñez, et al. (2021)	2021	Spain	Lexicon based	Dictionary based
Perera (2019)	2019	Sri Lanka	Supervised machine learning	Fuzzy rough set classifier, Naïve Bayes and SVM
CerónGuzmán (2016)	2016	Colombia	Supervised machine leaning	Logistic regression and sentiment score

RQ4. What advantages can SA conducted on election domain provide?

SA conducted in the election domain offers several advantages, including the ability to measure public sentiment in real-time, assisting political campaigns in formulating more vibrant messages and strategies, providing early detection of potential crises or negative sentiment, enhancing the accuracy of election outcome predictions, promoting transparency in governance, enabling more responsive representation of voter concerns, and facilitating academic research into political behavior and public opinion, ultimately it strengthening the democratic process and fostering informed decision-making by candidates, electorates and policymakers.

Sentiment analysis in election domain and the advantages it can provide is presented in Table 4.

Table 4. Advantages of Sentiment analysis on election domain

Sentiment analysis advantages	Articles
Crisis management	[16, 26]
Data-driven campaigns	[11, 21]
International perspective	[13, 19]
Issue prioritization	[5, 6]
Monitoring election integrity	[8]
Predictive power	[14, 15, 18, 19, 22]
Transparent governance	[25]

Brief outline for RQ4 findings is as follows:

a. Introduction to SA in Elections

Brief overview of SA's relevance and its potential impact on elections.

b. Presentation of Key Advantages

i. Crisis Management: Early detection of negative trends.
ii. Data-Driven Campaigns: Insights for campaign message refinement.
iii. International Perspective: Comparative analysis across regions.
iv. Issue Prioritization: Identification of key voter concerns.
v. Monitoring Election Integrity: Tracking sentiment on fairness and transparency.
vi. Predictive Power: Enhancing election outcome predictions.
vii. Transparent Governance: Fostering accountability and responsiveness.

Provided below is a brief overview of the articles referenced in the Table 4, grouped by sentiment analysis advantages:

a. Crisis Management [16]: This article examines how sentiment analysis (SA) can be utilized to detect early signs of public discontent and emerging crises, enabling swift responses by political campaigns to mitigate potential damage. [26]: Focuses on SA in social media contexts, showing how timely sentiment data helps manage public relations crises by identifying and addressing issues before they escalate.
b. Data-Driven Campaigns [11]: Discusses the role of SA in designing data-driven political campaigns, where real-time public opinion insights allow for tailored messaging and targeted strategies to engage different voter demographics. [21]: Highlights the effectiveness of using SA to optimize campaign resources and personalize outreach efforts based on evolving voter sentiment.
c. International Perspective [13]: Provides a comparative analysis of SA applications in different countries, emphasizing the unique political and cultural factors that influence public sentiment and election outcomes. [19]: Investigates

SA in a cross-national context, illustrating how political campaigns and governments can learn from sentiment trends observed globally to apply relevant strategies locally.

d. Issue Prioritization [5]: Analyzes how SA helps in identifying voter priorities by tracking sentiment around key issues, enabling campaigns to focus on topics that resonate most with the electorate. [6]: Explores how issue-based sentiment tracking allows policymakers to respond to shifting public concerns and allocate resources accordingly.
e. Monitoring Election Integrity [8]: Examines the use of SA in monitoring social media for signs of election related disinformation, fraud, and manipulation, contributing to a fairer electoral process.
f. Predictive Power [14]: Demonstrates SA's ability to predict election outcomes by analyzing trends in public sentiment across platforms, with historical data validating its accuracy. [15]: Discusses machine learning applications in SA that enhance predictive accuracy, showing a high correlation between public sentiment and election results. [18] and [22]: Both articles show SA's role in forecasting election results based on real-time data, emphasizing the predictive potential of sentiment trends. [19]: Also noted under "International Perspective," this article highlights predictive insights derived from cross-country sentiment data.
g. Transparent Governance [25]: This article discusses how SA fosters transparent governance by making public opinion on policy decisions accessible, allowing for more responsive and democratic decision making.

These research works collectively provide insights into how sentiment analysis across different domains can enhance campaign strategies, governance, and the democratic process.

4. Results

The breakdown of how each phase in Figure 1 contributes to the overall model's effectiveness in predicting election results is outlined below:

i. Data-driven intelligent models systematic reviews from digital library sources: Provides a foundation of structured and high-quality data for model training by gathering relevant, pre-evaluated literature.
ii. Natural language preprocessing tools: Prepares raw text data for analysis, ensuring cleaner, more relevant inputs that enhance model accuracy.
iii. Machine learning algorithms: Enables pattern recognition within the data, identifying trends and relationships that contribute to prediction insights.
iv. Lexicon based analysis: Uses predefined sentiment dictionaries to assign sentiment values, improving sentiment accuracy for election related text.
v. Transformer-based models: Leverages advanced neural networks for better context understanding and nuanced sentiment analysis, particularly useful in election discourse.
vi. Hybrid approach: Combines machine learning, transformer models, and lexicon methods to maximize predictive accuracy by balancing strengths and minimizing weaknesses of each approach.
vii. System model: Integrates outputs from all prior stages to generate a comprehensive prediction of election outcomes, leveraging both sentiment and trend analysis.

In Table 5, it is evident that the most frequently used data source for SA in election domain is Twitter house. This symbolized that the datasets derived from twitter can be generalized to serve as reliable and adequate input to SA in other to classify and generate a future prediction of an election.

Table 5. Data source for SA in election domain

Author(s)	Year	Data source	Sentiment approach	Techniques/ Tools	Performance evaluation metrics
Oyewola, et al. (2023)	2023	Twitter	Supervised machine learning	LSTMRNN, BERT and LSVC	BERT accuracy = 0.94
Kermani and Rasouli (2022)	2022	Twitter	Probabilistic classifiers	SNA and ECA	Result corelated positively
Mohapatra and Mohapatra (2022)	2022	Campaign speech and online news transcripts	Supervised machine learning	BERT	Result corelated positively
Rodríguez Ibáñez, et al. (2021)	2021	Twitter	Lexicon based	Dictionary based	N/A
Barreiro (2020)	2020	Twitter	Probabilistic and linear classifiers	MAKERRIMER, SVM, Naïve Bayes, Decision Trees, and Logistic Regression	MAKERRIMER accuracy = 0.76
Ekanem and Forsberg (2017)	2017	Electoral Management Body	Probabilistic classifiers	Binomial Regression and logit function	Pvalues = 0.1420, and 0.0346 respectively

Digital election measures utilized in SA within the election domain consists of various digital tools and platforms, such as social media, online campaign systems, and electoral management bodies. These measures collect and analyze data from digital sources to gauge public opinion, political discourse, and voter sentiment. It tracks candidate popularity, sentiment trends, and the virality of election related content on social media. Online campaign systems Online campaign systems are digital platforms and tools that facilitate the planning, execution, and management of political campaigns, enabling candidates and organizations to reach and engage with voters through various online channels, while election management bodies are responsible for overseeing and managing the electoral process, ensuring the fairness, transparency, and integrity of elections, and safeguarding the democratic principles of the electoral system. These measures offer a timely, data driven perspective on election related sentiment, assisting political campaigns, researchers, and policymakers in making informed decisions, understanding voter preferences, and predicting election outcomes.

Table 6. Digital election measures utilized in SA on election domain

Digital election measures	Articles
Twitter	[5, 8, 9, 10, 16, 17, 18, 19, 20, 21, 24, 26, 27, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56]
Electoral management body	[23]
Campaign speech	[25]
Online news transcripts	[25]

5. Discussion

From the review, it was discovered that integrating diverse data sources for election prediction can enhance predictive accuracy by combining insights from multiple perspectives, capturing a fuller picture of public sentiment and trends. Studies shown that combining data from sources like Twitter, campaign speeches, and news articles, alongside different sentiment analysis approaches, can yields robust predictions due to complementary information from varied platforms. This can be ascertained from these works:

i. Oyewola et al. (2023) achieved high predictive accuracy (BERT accuracy = 0.91) by using Twitter data with machine learning techniques like LSTM-RNN, BERT, and LSVC, showing the benefit of applying advanced deep learning on social media data for sentiment-rich, real-time insights.
ii. Kermani and Rasouli (2022) utilized probabilistic classifiers on Twitter data, finding positive correlations in results, which underscores the value of social network analysis (SNA) and entity-centric analysis (ECA) in identifying sentiment patterns relevant to election predictions.
iii. Mohapatra and Mohapatra (2022) combined campaign speech transcripts with online news articles, using BERT for sentiment analysis, and reported positive correlations. This integration highlights how combining real-time public discourse with media narratives can provide a broader and contextually nuanced prediction base.
iv. Rodríguez-Ibáñez et al. (2021) employed a lexicon-based approach on Twitter, showing that dictionary-based sentiment analysis can be effective for detecting the sentiment polarity in large-scale datasets, though it may lack the nuance of machine learning models.
v. Barreiro (2020) integrated various probabilistic and linear classifiers on Twitter data, achieving an accuracy of 0.76 with MAKERRIMER, showing that different classifiers can contribute uniquely to modeling public sentiment and election outcomes.
vi. Ekanem and Forsberg (2017) used data from electoral management bodies with binomial regression, which provided statistical significance (p-values of 0.1420 and 0.0346), highlighting the predictive value of electoral data itself as a baseline for modeling.

Also, these studies underscore that integrating data from different sources (e.g., social media, electoral bodies, speeches) and combining multiple sentiment approaches (e.g., lexicon-based, supervised machine learning) could allow a richer, more accurate predictive model by leveraging the strengths of each data type and analytical method.

6. Conclusion

A systematic literature review of SA in election domain has been carried out in detail. Many authors have disclosed the improvements SA has offered in the area of election domain which includes: issue prioritization, real-time election monitoring, crisis management, predictive power of the citizen, transparency in governance, data-driven campaign strategy and election integrity for an effective governance by policy makers. This study found out that Twitter and online news transcripts are the primary digital resources for collecting data for SA. The study also discovered that Twitter house has a strong predictive power over the outcome of an election. Lastly, the study observed that most researchers utilized supervised learning approach for SA in election domain due to its ability to leverage labeled training data to learn and generalize from known sentiments by enabling accurate classification of text based on predefined sentiment categories. A future research is intended by extending the study to include current digital libraries that will have real time views to enlighten the policy makers and the political stakeholders in the viability outcome of their activities and the impact on the masses.

REFERENCES

[1] M. E., Ekpenyong, Ananga, A. J., Udoh, E. O., & Umoh, N. M. (2019). Speech Prosody Extraction for Ibibio Emotions Analysis and Classification. *In Language and Technology Conference Proc., Cham: Springer International Publishing*, pp. 197213.

[2] G., Feng, Chen, K., Cai, H., & Li, Z. (2023). A Hybrid Method of Sentiment Analysis and Machine Learning Algorithm for the US Presidential Election Forecasting. *Proc of IEEE International Conference on Big Data*, pp. 14951500.

[3] F. Jemai, Hayouni, M., & Baccar, S. (2021). Sentiment analysis using machine learning algorithms. *IEEE Proc. of the International Wireless Communications and Mobile Computing (IWCMC)*, pp. 775779.

[4] N. Afroz, Boral, M., Sharma, V., & Gupta, M. (2021, March). Sentiment analysis of COVID19 nationwide lockdown effect in India. *IEEE International Conference on Artificial Intelligence*

and Smart Systems (ICAIS), pp. 561567.
[5] L. B. R, Barreiro (2020). Decision Modelling Driven by Twitter Data: A Case Study of the 2017 Presidential Election in Ecuador, A Ph. D dissertation, University of Manchester, UK.
[6] H. Wang, Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012). A system for realtime twitter sentiment analysis of 2012 us presidential election cycle. *In Proc. of the ACL 2012 system demonstrations*, pp. 115120.
[7] H. Kermani, & Rasouli, F. (2022). Protesting is Not Everything: Analyzing Twitter Use during Electoral Events in Nondemocratic Contexts. *Journal of Digital Social Research*, 4(4): 2251.
[8] R. W. Miller Jr. (2020). Political Spectacle and Twitter Usage by 2016 US Presidential Candidates: A Content Analysis, Ph. D dissertation, Walden University.
[9] A. Tsakalidis, Papadopoulos, S., Cristea, A. I., & Kompatsiaris, Y. (2015). Predicting elections for multiple countries using Twitter and polls. *IEEE Intelligent Systems*, 30(2):1017.
[10] D. Shaw, Li, Y., & Hadaegh, A. (2021). Twitter based sentiment analysis of each presidential candidate using long shortterm memory. *International Journal of Computer Science and Security*, 15:8796.
[11] M. RodríguezIbáñez, GimenoBlanes, F. J., CuencaJiménez, P. M., SogueroRuiz, C., & RojoÁlvarez, J. L. (2021). Sentiment analysis of political tweets from the 2019 Spanish elections. *IEEE Access*, 9:101847101862.
[12] A. VignaGómez, Murillo, J., Ramirez, M., Borbolla, A., Márquez, I., & Ray, P. K. (2023). Design and analysis of tweetbased election models for the 2021 Mexican legislative election. *EPJ Data Science*, 12(1): 23.
[13] O. Olabanjo, Wusu, A., Padonu, R., Afisi, O., & Mazzara, M. (2023). Twitter Sentiment Analysis of Lagos State 2023 Gubernatorial Election Using BERT. *Journal of Advanced Research in Social Sciences*, 6(2): 5975.
[14] R. Liu, Yao, X., Guo, C., & Wei, X. (2021). Can we forecast presidential election using twitter data? an integrative modelling approach. *Annals of GIS*, 27(1): 4356.
[15] R. Sorathiya (2021). Political analytics on election candidates and their parties in context of the US Presidential elections 2020, Ph. D dissertation, Laurentian University of Sudbury.
[16] L. Oikonomou (2017). Predicting the USA presidential elections using Twitter Data. M.Sc. Thesis, International Hellenic University.
[17] P. Singh, Sawhney, R. S., & Kahlon, K. S. (2017). Forecasting the 2016 US presidential elections using sentiment analysis. *In Digital Nations–Smart Cities, Innovation, and Sustainability: 16th IFIP WG 6.11 Conference on eBusiness, eServices, and eSociety*, 412423.
[18] L. L Tavabi (2015). Evaluation of political sentiment on twitter. University of California, Los Angeles, USA.
[19] U. Yaqub, Chun, S. A., Atluri, V., & Vaidya, J. (2017). Analysis of political discourse on twitter in the context of the 2016 US presidential elections. *Government Information Quarterly*, 34(4):613626.
[20] D. O. Oyewola, 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).
[21] N. D. Prasetyo (2014). Tweet based election prediction. M.Sc. Thesis, Delft University of Technology.
[22] S. Malla, Kumar, L. K., & Alphonse, P. J. A. (2023). Novel fuzzy deep learning approach for automated detection of useful COVID19 tweets. *Artificial Intelligence in Medicine*, 143:102627.
[23] U. O. Ekanem, & Forsberg, O. J. (2017). An Analysis of the 2015 Nigerian Presidential Election. *Purdue: Undergraduate Research Journal*, 1(2):1.
[24] J. A. CerónGuzmán (2016). A Sentiment Analysis Model of Spanish Tweets. Case Study: Colombia 2014 Presidential Election, Ph.D Thesis, Universidad, Nacional De Colombia.
[25] S. Mohapatra & Mohapatra, S. (2022). Sentiment is all you need to win US Presidential elections. arXiv preprint arXiv:2209.13487.
[26] W. A. S. N. Perera (2019). Sentiment based analysis of social media data using fuzzy rough set classifier for the prediction of the presidential election, M.Sc. Thesis from School of Computing, University of Colombo, Sri Lanka.
[27] U. Yaqub, Chun, S. A., Atluri, V., & Vaidya, J. (2017). Sentiment based analysis of tweets during the US presidential elections, *Proc. of the 18th Annual International Conference on Digital Government Research*, pp. 110.
[28] A. Ligthart, Catal, C., & Tekinerdogan, B. (2021). Systematic reviews in sentiment analysis: a tertiary study. *Artificial Intelligence Review*, pp. 157.
[29] N. A. S. Abdullah, & Rusli, N. I. A. (2021). Multilingual Sentiment Analysis: A Systematic Literature Review, *Pertanika Journal of Science & Technology*, 29(1).
[30] Q. A. Xu, Chang, V., & Jayne, C. (2022). A systematic review of social mediabased sentiment analysis: Emerging trends and challenges, *Decision Analytics Journal*, 3:100073.
[31] E. Aburaghiega, Farrag, M. E., Hepburn, D., & Haggag, A. (2018). Enhancement of power transformer state of health diagnostics based on fuzzy logic system of DGA. *Proc. of 20th IEEE International Middle East Power Systems Conference (MEPCON*, pp. 400405.
[32] H. A. AlJamimi, & Saleh, T. A. (2019). Transparent predictive modelling of catalytic hydrodesulfurization using an interval type2 fuzzy logic. *Journal of cleaner production*, 231:

10791088.
[33] L. AmadorAngulo, Castillo, O., Melin, P., & Castro, J. R. (2022). Interval type3 fuzzy adaptation of the bee colony optimization algorithm for optimal fuzzy control of an autonomous mobile robot. *Micromachines*, 13(9): 1490.
[34] D. MartínGutiérrez, HernándezPeñaloza, G., Hernández, A. B., LozanoDiez, A., & Álvarez, F. (2021). A deep learning approach for robust detection of bots in twitter using transformers. *IEEE Access*, 9:5459154601.
[35] J. Pfeiffer, Rücklé, A., Poth, C., Kamath, A., Vulić, I., Ruder, S., & Gurevych, I. (2020). Adapterhub: A framework for adapting transformers. arXiv preprint arXiv:2007.07779.
[36] C. Wang, Chen, Y., Zhang, S., & Zhang, Q. (2022). Stock market index prediction using deep Transformer model. *Expert Systems with Applications*, 208: 118128.
[37] A. Gruzd, Doiron, S., & Mai, P. (2011). Is Happiness Contagious Online? A Case of Twitter and the 2010 Winter Olympics. *Proc. of the 44th Hawaii International Conference on System Sciences*.
[38] S. Baccianella, Esuli, A., & Sebastiani, F. (2010). SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. *Proc. of the 7th International Conference on Language Resources and Evaluation (LREC)*.
[39] J. Mutinda, Mwangi, W., & Okeyo, G. (2023). Sentiment analysis of text reviews using lexicon enhanced Bert embedding (LeBERT) model with convolutional neural network. *Applied Sciences*, 13(3): 1445.
[40] G. R. Narayanaswamy (2021). Exploiting BERT and RoBERTa to improve performance for aspectbased sentiment analysis, Ph. D Dissertation, University of Dublin, UK.
[41] R. M. Nesse (1990). Evolutionary explanations of emotions. *Human Nature*, 1(3):261289.
[42] F. Å. Nielsen (2011). A New ANEW: Evaluation of a Word List for Sentiment Analysis in Microblogs. *Proc. of the ESWC2011 Workshop on 'Making Sense of Microposts*'.
[43] S. L. Althaus & Coe, K. (2011). Predicting political participation: The influence of novel and traditional predictors. *The Journal of Politics*, 73(1): 278288.
[44] R. M. Alvarez, Atkeson, L. R., & Hall, T. E. (2001). Economic and noneconomic factors influencing the 2000 presidential vote: The first application of the preelection/postratification election exit poll. *Political Research Quarterly*, 54(4):805–826.
[45] S. Malla, Kumar, L. K., & Alphonse, P. J. A. (2023). Novel fuzzy deep learning approach for automated detection of useful COVID19 tweets. *Artificial Intelligence in Medicine*, 143: 102627.
[46] S. Malviya, Tiwari, A. K., Srivastava, R., & Tiwari, V. (2020). Machine learning techniques for sentiment analysis: A review. SAMRIDDHI: *A Journal of Physical Sciences, Engineering and Technology*, 12(2): 7278.
[47] S. Moghaddam, Ester, M., & Jung, J. (2019). DeepWordBug: A Textual Adversarial Attack Against Deep Learning Based Sentiment Analysis Models. *Proc. of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*.
[48] B. Pang & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2(1–2): 1–135.
[49] S. Poria, Cambria, E., Hazarika, D., Vij, P., & Hussain, A. (2017). Multilevel Multiple Attentions for Contextual Multimodal Sentiment Analysis. *Proc. of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.
[50] G. Roccabruna, Azzolin, S., & Riccardi, G. (2022). Multisource multidomain sentiment analysis with BERTbased models. *In Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pp. 581589.
[51] D. Shah, Li, Y., & Hadaegh, A. (2021). Twitter based sentiment analysis of each presidential candidate using long short term memory. International Journal of Computer Science and Security, 15: 8796.
[52] L. L. Tavabi (2015). Evaluation of political sentiment on twitter. University of California, Los Angeles.
[53] Tsakalidis, A., Papadopoulos, S., Cristea, A. I., & Kompatsiaris, Y. (2015). Predicting elections for multiple countries using Twitter and polls. *IEEE Intelligent Systems*, 30(2): 1017.
[54] U. A. Umoh, Udoh, S., Isong, E., Asuquo, R. & Nyoho, E (2019), PSO Optimized Interval Type-2 Fuzzy Design for Elections Results Prediction. International Journal of Fuzzy Logic Systems (IJFLS) Vol.9, No.1, 1-19, January 2019
[55] U. A. Umoh, Udoh, S. & Nyoho E. (2018). An Interval Type-2 Fuzzy Framework for Elections Results Prediction. Proceedings of NIWIT 4th International Conference, 9th – 10th May 2018, Precious Palm Royal Hotel, Isiohor Village, Benin –City, Edo State, Nigeria, 59 - 74.
[56] V. E. Ekong, Asuquo, D. E. & Ejodamen, P. U. (2021). Towards an Analytic Framework for Twitter influencers in Boko Haram Discourse. *In Proceedings of the International Conference on Innovative Systems for Digital Economy* (ISDE). Nigeria. 204 –212.