International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Multifaceted Method for Election Prediction on Multiple Social Media Platforms Using Machine Learning

[1]Sanjay Kumar Gupta, [2]Prof. K. P. Yadav

Submitted: 03/05/2024 **Revised**: 16/06/2024 **Accepted**: 23/06/2024

Abstract: The paper presents a Multifaceted Method for Election Prediction on Multiple Social Media Platforms Using Machine Learning" . It is a novel approach to forecasting election outcomes by leveraging data from various social media platforms like Facebook, twitter and YouTube. . By employing machine learning algorithms, the method integrates diverse social media signals, such as user sentiment, engagement metrics, and topic trends, to create a comprehensive prediction model. The model has optimize sentiment analysis as well as use of proposed ensemble approach. This multifaceted approach aims to improve the accuracy of election predictions by capturing a wider array of public opinions and behaviors across different social media ecosystems. The study demonstrates the potential of this method through case studies and empirical analysis, highlighting its effectiveness. Evaluation of proposed method with existing models have been conducted and found that proposed method got good accuracy over others. It is approximate 94.27 % for combined data and proposed ensemble approach also have good accuracy over state of art methods.

Keywords*: Election Prediction, Social media platforms, machine learning, ensembled learning, Natural language processing*

1. Introduction

For politicians, the upcoming LokSabha elections are not about elections; This is about political choice. Social media has changed the way people think, write and feel, but political experts believe social media still influences the way people make choices. Most politicians understand the heartbeat of media and public support over time, so more and more politicians are realizing the importance of social media in election. Predicting election results based on major political events can save large amounts of money and effort from campaigns and advertising. Predicting people's political preferences through social media can add to or change polls. But data on social media surveys can be complex and confusing. Voting allows politicians to use Twitter, Facebook, YouTube, Google plus etc. requires communication via social media. It is clear that online tools play an important role in shaping public opinion and setting political agendas[1]. Methods used to predict election results; Quantitative analysis is divided into two groups: sentiment analysis and social analysis [2].

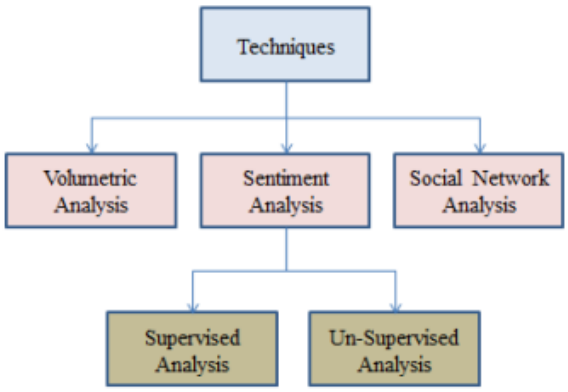

Fig 1: Technique used to predict electoral result [2]

In literature mostly work has been done on twitter and with sentiment analysis. But it is important to include multiple social media indicators into the account and also use an efficient method for sentiment analysis. It is also required to develop an optimized solution for sentiment analysis of user generated contents. Lastly combining data extraction, social media indicators & sentiment analysis on various social media platform is also a critical task. Proposed research work has developed a framework to provide solution to all these problems.

2. Related Work

P. Juneja and U. Ojha [10] classified Twitter data based on sentiment (positive or negative) using different machine learning to predict Delhi Corporation election results and identify the latest trends in machine learning. This study shows the performance of different classifiers on information from party-related Twitter profiles.

[1]*Research Scholar MATS University Raipur*
sanjiit@rediffmail.com
[2]*Vice Chancellor MATS University, Raipur*
dryadavkp@gmail.com

Experiments show that the polynomial naive Bayes classifier is the most accurate hypothesis estimator with 78% accuracy. D. A. Kristiyant [9] compared two types of classification such as Support Vector Machine Algorithm (SVM) and Naive Bayes Classifier (NBC). The results show that Naive Bayes Classifier (NBC) algorithm has higher accuracy than Support Vector Machine (SVM) and the accuracy of is as high as 94%. P. KhuranaBatra et al. [7] proposed a model for election prediction using the Twitter platform dataset. K.d.S. Brito and P. J. L. [8] proposed an election prediction system based on news media using machine learning. Social networks like Facebook, Twitter, and Instagram have revolutionized the way voters and campaigns communicate. Irmalasari and L. Dwiyanti [3] generally use decision trees, support vector machines, linear regression, etc. for classification. It uses many simple models including: The most effective model for prediction using the given data is the gradient-assisted decision tree, which achieves an F1 score of 0.7445 by evaluating the parameters and using the selection results. Dr. Ulandari et al. [4] Use artificial intelligence and optimization techniques to improve sentiment analysis for Indonesian election predictions. In this study, emotional analysis is to use a large number of keywords to extract information from Twitter and use the Valence Aware Dictionary for Emotional Reasoning (VADER) method to plot it, and use three algorithms such as Naive Bayes and K- to do its job well. . I went out to check. Research results show that the VADER method combined with SVM and Grid Search CV for parameter optimization achieves the best results when the index is 0.1, the accuracy reaches 91.93%, the return value is 91.76%, and the actual value is 91.76%. It shows that it reached 97.94. . and f1 score reaches 91.84%. S. Perera and K. Karunanayaka [5] use social media for sentiment analysis. This study focuses on machine learning technology to analyze Twitter data. Special attention is paid to classifiers based on fuzzy sets and coarse methods. The accuracy of classifiers based on fuzzy rough sets is higher than other classifiers. The actual results of the 2019 presidential election defied the candidates' predictions. M. Fachrie and F. Ardiani [6] developed an approach based on text mining and machine learning techniques to create an emotional evaluation model as a basis for prediction. Therefore, the accuracy of the prediction model reached 72.96% after the post-pruning and pre-pruning processes were completed. Brito [24] stated that Twitter alone cannot make good decisions [25].

Major contribution in the field is as follows:

References/Year	Algorithm Used	Dataset	Results	State/Place
[23]/2024	Hybrid Machine Learning	Contextual & synthetic data	Not fixed	Election 2019,2021
[11]/2023	LSTM, RNN, BERT and LSVC.	Twitter Dataset	Accuracy=94 Precision=88.5 Recall= 92.5 AUC= 94.7 F-measure =91.7. BERT has the highest of all three.	Nigeria Presidential Election 2023.
[12]/2023	BERT	Twitter Dataset	NA	Indian General Election 2019.
[14]/2023	Review of Various base methods (LR, SVM,and NN) and Advanced methods (BERT, T5 etc.).	Common Used dataset.	Comparative Analysis.	Mixed Domains.
[20]/2023	Detecting impact of social media on elections.	All Media TV and Others.	the disruption in TV coverage of politics had an impact on voters' source of	Italian 2013 electoral race

			political information	
[13]/2022	Latent Dirichlet Allocation (LDA).	Twitter Dataset	Statistical analysis.	U.S. Presidential Election 2020
[21]/2022	BERT based sentiment analysis.	Different domain dataset.	Precision = 0.98% Recall= 0.98% Accuracy= 0.97%	General Dataset.
[16]/2022	long short-term memory (LSTM), peephole long short term memory (PLSTM), and two-stage residual long short-term memory (TSRLSTM) models.	Twitter Dataset	Accuracy= 71. 12 by TSELSTM.	Presidential election in Nigeria, Atiku, Tinubu, and Obi, by conducting a sentiment analysis on tweets.
[17]/2022	Political satire in the 2019 election. Mojok.co has provided political education to the community through the pictures/images.	Signs are represented through the elements of colour, illustration, and typography in the info graphic content design of Mojok.co.	Only discussion.	Indonesian Election 2019.
[15]/2021	Review on social media responsibility in political manipulation.	European Commission database.	Review Analysis.	Political and Social Domains.
[22]/2020	LSTM	Twitter Dataset	Precisions= 0.81% Recall=0.92%	General Elections of India in 2019
[19]/2019	News-sharing on Twitter during the UK General Election in 2015. Analysis of the network structure of users and the news articles they share identifies multiple distinct user communities.	Twitter Dataset	Suggesting that social media users may be exposed to a more diverse information flow than suggested by some other studies of political social media.	Hybrid
[18]/2017	Tweets counting based analysis.	Twitter Dataset	Relationship between social media Buzz of political parties on seats won.	Indian General Election 2014.

3. Proposed Methodology

The proposed research work has developed the multifaceted method for election result prediction on social media. The system takes data from multiple social media like twitter, facebook and youtube. The system uses Natural Language Processing and Machine Learning method for generating optimized results. The proposed framework is shown below:

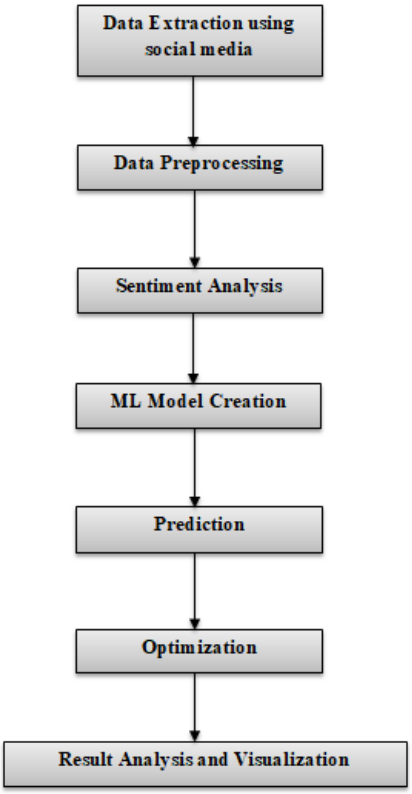

Fig 2: Proposed System

At the data collection stage there is a process like first, data retrieval is done from Social Media Platform on Youtube, Facebook and Twitter through Social Media API. Preprocessing reduces processing overhead up to many extents. Facebook and twitter posts contain noisy data which is useless. Before social media contents are analyzed, they are pre-processed, transform and categorize. Prediction includes optimized sentiment analysis & score generation phase Proposed prediction algorithm is an ensemble machine learning method that include ensemble of LR, NB & SVC. Steps are listed below:

Algorithm: EnsemblePrediction

Step 1: Load Data

LoadData()

Step 2: Split Data into Training and Testing Sets

SplitData()

Step 3: Train Individual Models

TrainLogisticRegression()

TrainNaiveBayes()

TrainSupportVectorClassifier()

Step 4: Make Predictions with Individual Models

LR_Predictions = PredictWithLogisticRegression()

NB_Predictions = PredictWithNaiveBayes()

SVC_Predictions = PredictWithSupportVectorClassifier()

Step 5: Combine Predictions using Ensemble

EnsemblePredictions = CombinePredictions(LR_Predictions, NB_Predictions, SVC_Predictions)

Step 6: Evaluate Ensemble Model

EvaluateModel(EnsemblePredictions, TrueLabels)

End Algorithm

Sentiment score is calculated for text comments in social media. But this has to be optimizing as the roles of other indicators for example like, dislikes are also important. The proposed system will utilize following

1. First take the dataset with likes and comment
2. Calculate sentiment score for tweet.
3. Calculate sentiment score for comment
4. Calculate entropy for comment and like using the formula
   $H=-(pl\log2(pl)+pc\log2(pc))$

   Where H is the entropy

   Pl is the probability of like>500

   Pc is the probability of comment sentiment>1
5. If like>500 and comment sentiment >1 update the cumulative sentiment by

4. Result

The proposed system is implemented with taking data from twitter,you tube and facebook platforms for Delhi Assembly Election 2020. After data extraction prediction has been performed by implementing proposed ensemble method as well as other existing methods like naive bayes, logistic regression, support vector, etc. The accuracy found is represented in table below:

Table 1: Comparison of accuracy achieved in different methods

Model	Youtube	Facebook	Twitter	Combined (Proposed)
Naïve Bayes	88.65	94.20	88.46	**92.8**
Logistics Regression	86.90	93.97	89.60	**91.56**
Support Vector Machine	84.15	91.25	77.21	**88.65**
Ensemble (Proposed)	**96.47**	**97.33**	**95.66**	**94.27**

Graphical chart for comparison is shown below:

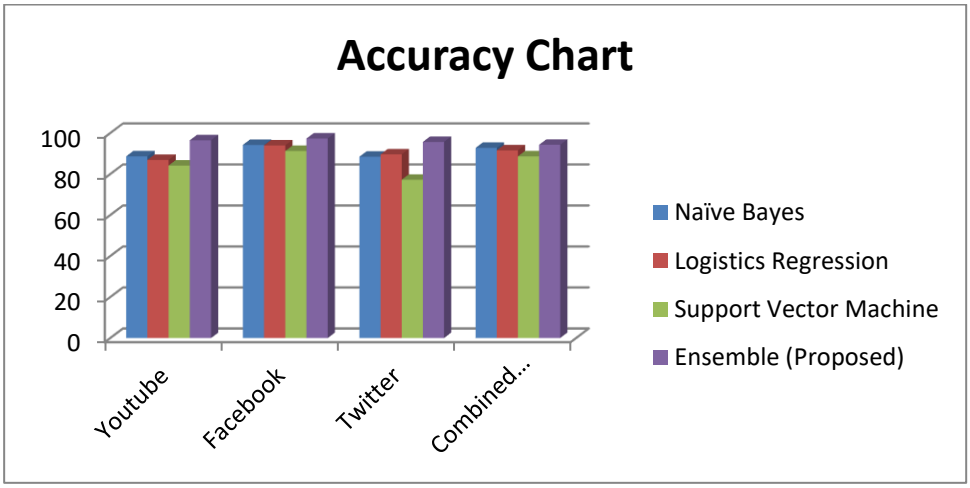

Fig 3: Performance analysis

It is found that the proposed method has provided best result as compare to all other methods.

5. Conclusion

The use and influenceof social media in politics have been recognized by manyresearchers and political parties in making electoralpredictions and devising future campaign strategies. In literature mostly work has been done on twitter and with sentiment analysis. But it is important to include multiple social media platform and indicators into the account and also use an efficient method for sentiment analysis. The proposed ensemble method produces more accurate result as shown in fig 3. It includes you tube, facebook& twitter platforms. Only sentiment analysis of comments cannot produce correct sentiment. Social media indicators like no of likes can add up accurate sentiment prediction. In future more indicators and real time analysis can be introduced. Best performing result is of proposed ensemble method. It gives 96.47, 96.33, 95.66 percent accuracy for youtube, facebook and twitter data but when we combine all data along with other social media KPI than accuracy achieve is 94.27%. In future system can be extended to other social platform as well as improved algorithmic development.

References

[1] Woolley, A.W., Chabris, C.F., Pentland, A., Hashmi, N., & Malone, T.W., “Evidence for a Collective Intelligence Factor in the Performance of Human Groups.” Science, 330, 686, 2010.

[2] Muhammad Bilal, Abdullah Gani, Mohsen Marjani, Nadia Malik, “predicting elections: social media data and techniques”, IEEE, 2019.

[3] Irmalasari and L. Dwiyanti, "Algorithm Analysis of Decision Tree, Gradient Boosting Decision Tree, and Random Forest for Classification (Case Study: West Java House of Representatives Election 2019)," 2023 International Conference on Electrical Engineering and Informatics (ICEEI), Bandung,

Indonesia, 2023, pp. 1-5, doi: 10.1109/ICEEI59426.2023.10346727.

[4] D. R. Wulandari, M. A. Murti and H. F. T.S.P, "Sentiment Analysis Based on Text About President Candidate 2024 in Indonesia Using Artificial Intelligence with Parameter Optimization Algorithm," 2023 IEEE International Conference on Internet of Things and Intelligence Systems (IoTaIS), Bali, Indonesia, 2023, pp. 216-222, doi: 10.1109/IoTaIS60147.2023.10346038.

[5] S. Perera and K. Karunanayaka, "Sentiment Analysis of Social Media Data using Fuzzy-Rough Set Classifier for the Prediction of the Presidential Election," 2022 2nd International Conference on Advanced Research in Computing (ICARC), Belihuloya, Sri Lanka, 2022, pp. 188-193, doi: 10.1109/ICARC54489.2022.9754173.

[6] M. Fachrie and F. Ardiani, "Predictive Model for Regional Elections Results based on Candidate Profiles," 2021 8th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), Semarang, Indonesia, 2021, pp. 247-252, doi: 10.23919/EECSI53397.2021.9624256.

[7] P. KhuranaBatra, A. Saxena, Shruti and C. Goel, "Election Result Prediction Using Twitter Sentiments Analysis," 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC), Waknaghat, India, 2020, pp. 182-185, doi: 10.1109/PDGC50313.2020.9315789.

[8] K. d. S. Brito and P. J. L. Adeodato, "Predicting Brazilian and U.S. Elections with Machine Learning and Social Media Data," 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 2020, pp. 1-8, doi: 10.1109/IJCNN48605.2020.9207147.

[9] D. A. Kristiyanti, A. H. Umam, M. Wahyudi, R. Amin and L. Marlinda, "Comparison of SVM & Naïve Bayes Algorithm for Sentiment Analysis Toward West Java Governor Candidate Period 2018-2023 Based on Public Opinion on Twitter," 2018 6th International Conference on Cyber and IT Service Management (CITSM), Parapat, Indonesia, 2018, pp. 1-6, doi: 10.1109/CITSM.2018.8674352.

[10] P. Juneja and U. Ojha, "Casting online votes: To predict offline results using sentiment analysis by machine learning classifiers," 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Delhi, India, 2017, pp. 1-6, doi: 10.1109/ICCCNT.2017.8203996.

[11] OlusolaOlabanjo, AshiriboWusu, OseniAfisi, MautonAsokere, Rebecca Padonu, OlufemiOlabanjo, OluwafolakeOjo, OlusegunFolorunso, Benjamin Aribisala, Manuel Mazzara, "From Twitter to Aso-Rock: A sentiment analysis framework for understanding Nigeria 2023 presidential election", Heliyon, Volume 9, Issue 5, 2023, e16085, ISSN 2405-8440, https://doi.org/10.1016/j.heliyon.2023.e16085.

[12] MudasirMohd, SaheebaJaveed, MohsinAltafWani, Hilal Ahmad Khanday, AbidHussainWani, Umar Bashir Mir, Sheikh Nasrullah, poliWeet — Election prediction tool using tweets,

[13] Software Impacts, Volume 17, 2023, 100542, ISSN 2665-9638, https://doi.org/10.1016/j.simpa.2023.100542.

[14] Amir Karami, Spring B. Clark, Anderson Mackenzie, Dorathea Lee, Michael Zhu, Hannah R. Boyajieff, Bailey Goldschmidt, 2020 U.S. presidential election in swing states: Gender differences in Twitter conversations, International Journal of Information Management Data Insights, Volume 2, Issue 2, 2022,100097, ISSN 2667-0968, https://doi.org/10.1016/j.jjimei.2022.100097.

[15] Margarita Rodríguez-Ibanez, Antonio Casanez-Ventura, Felix Castejon-Mateos, Pedro-Manuel Cuenca-Jimenez, "A review on sentiment analysis from social media platforms", Expert Systems with Applications, Volume 223, 2023, 119862, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2023.119862.

[16] Ulrike Reisach, "The responsibility of social media in times of societal and political manipulation", European Journal of Operational Research, Volume 291, Issue 3, 2021, Pages 906-917, ISSN 0377-2217, https://doi.org/10.1016/j.ejor.2020.09.020.

[17] David OpeoluwaOyewola, LawalAbdullahiOladimeji, SoworeOlatunji Julius, LummoBalaKachalla, Emmanuel Gbenga Dada, "Optimizing sentiment analysis of Nigerian 2023 presidential election using two-stage residual long short term memory", Heliyon, Volume 9, Issue 4, 2023, e14836, ISSN 2405-8440, https://doi.org/10.1016/j.heliyon.2023.e14836.

[18] Sunarso, BenniSetiawan, Ni PutuPandeSatyaAnjani, "The political satire of Mojok.co in the 2019 Indonesian election", Heliyon, Volume 8, Issue 7, 2022, e10018,

[19] ISSN 2405-8440, https://doi.org/10.1016/j.heliyon.2022.e10018.

[20] MdSafiullah, PramodPathak, Saumya Singh, AnkitaAnshul, "Social media as an upcoming tool for political marketing effectiveness", Asia Pacific Management Review, Volume 22, Issue 1, 2017, Pages 10-15, ISSN 1029-3132, https://doi.org/10.1016/j.apmrv.2016.10.007.

[21] Iain S. Weaver, Hywel Williams, Iulia Cioroianu, LorienJasney, Travis Coan, Susan Banducci, "Communities of online news exposure during the

UK General Election 2015", Online Social Networks and Media, Volumes 10–11, 2019, Pages 18-30, ISSN 2468-6964, https://doi.org/10.1016/j.osnem.2019.05.001.

[22] Giulia Caprini, "Does candidates' media exposure affect vote shares? Evidence from Pope breaking news", Journal of Public Economics, Volume 220, 2023, 104847, ISSN 0047-2727, https://doi.org/10.1016/j.jpubeco.2023.104847.

[23] SayyidaTabindaKokab, SohailAsghar, ShehneelaNaz, "Transformer-based deep learning models for the sentiment analysis of social media data", Array, Volume 14, 2022,

[24] 100157, ISSN 2590-0056, https://doi.org/10.1016/j.array.2022.100157.

[25] MohdZeeshan Ansari, M.B. Aziz, M.O. Siddiqui, H. Mehra, K.P. Singh, "Analysis of Political Sentiment Orientations on Twitter", Procedia Computer Science, Volume 167, 2020, Pages 1821-1828, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2020.03.201.

[26] Zuloaga-Rotta L, Borja-Rosales R, Rodríguez Mallma MJ, Mauricio D, Maculan N. "Method to Forecast the Presidential Election Results Based on Simulation and Machine Learning". *Computation*. 2024; 12(3):38. https://doi.org/10.3390/computation12030038

[27] KellytonBrito, RogérioLuiz Cardoso Silva Filho, Paulo Jorge LeitãoAdeodato, "Stop trying to predict elections only with twitter – There are other data sources and technical issues to be improved",Government Information Quarterly,Volume 41, Issue 1,2024,101899,ISSN 0740-624X,https://doi.org/10.1016/j.giq.2023.101899.

[28] Kirkizh, N., Ulloa, R.,Stier, S., &Pfeffer, J. "Predicting political attitudes from web tracking data: a machine learning approach". Journal of Information Technology & Politics, 1–14. 2024, https://doi.org/10.1080/19331681.2024.2316679