Informing Science: the International Journal of an Emerging Transdiscipline

An Official Publication of the Informing Science Institute
InformingScience.org

Inform.nu

Volume 27, 2024

CRITICAL REVIEW OF STACK ENSEMBLE CLASSIFIER FOR THE PREDICTION OF YOUNG ADULTS' VOTING PATTERNS BASED ON PARENTS' POLITICAL AFFILIATIONS

Godwin Elo *	University of Education, Winneba, Central Region, Ghana	mychrist.eg@gmail.com
Benjamin Ghansah	University of Education, Winneba, Central Region, Ghana	bghansah@uew.edu.gh
Ephrem Kwaa-Aidoo	University of Education, Winneba, Central Region, Ghana	ekkaidoo@uew.edu.gh

*Corresponding Author

ABSTRACT

Aim/Purpose	This review paper aims to unveil some underlying machine-learning classification algorithms used for political election predictions and how stack ensembles have been explored. Additionally, it examines the types of datasets available to researchers and presents the results they have achieved.
Background	Predicting the outcomes of presidential elections has always been a significant aspect of political systems in numerous countries. Analysts and researchers examining political elections rely on existing datasets from various sources, including tweets, Facebook posts, and so forth to forecast future elections. However, these data sources often struggle to establish a direct correlation between voters and their voting patterns, primarily due to the manual nature of the voting process. Numerous factors influence election outcomes, including ethnicity, voter incentives, and campaign messages. The voting patterns of successors in regions of countries remain uncertain, and the reasons behind such patterns remain ambiguous.
Methodology	The study examined a collection of articles obtained from Google Scholar, through search, focusing on the use of ensemble classifiers and machine learning classifiers and their application in predicting political elections

Accepting Editor Eli Cohen │Received: December 21, 2023│ Revised: January 22, February 5, February 13, 2024 │ Accepted: February 14, 2024.

Cite as: Elo, G., Ghansah, B., & Kwaa-Aidoo, E. (2024). Critical review of stack ensemble classifier for the prediction of young adults' voting patterns based on parents' political affiliations. *Informing Science: The International Journal of an Emerging Transdiscipline, 27,* Article 2. https://doi.org/10.28945/5254

	through machine learning algorithms. Some specific keywords for the search include "ensemble classifier," "political election prediction," and "machine learning", "stack ensemble".
Contribution	The study provides a broad and deep review of political election predictions through the use of machine learning algorithms and summarizes the major source of the dataset in the said analysis.
Findings	Single classifiers have featured greatly in political election predictions, though ensemble classifiers have been used and have proven potent use in the said field is rather low.
Recommendations for Researchers	The efficacy of stack classification algorithms can play a significant role in machine learning classification when modelled tactfully and is efficient in handling labelled datasets. however, runtime becomes a hindrance when the dataset grows larger with the increased number of base classifiers forming the stack.
Impact on Society	The implications of the study on society at large are the need not downplay the impact of young adults' political contribution as a culmination of the experiences they have had with immediate political parents before voting age and their early political party preferences.
Future Research	There is the need to ensure a more comprehensive analysis, alternative data sources rather than depending largely on tweets, and explore ensemble machine learning classifiers in predicting political elections. Also, ensemble classification algorithms have indeed demonstrated superior performance when carefully chosen and combined.

Keywords	politics, elections, prediction, machine learning, classification, ensemble algorithm

INTRODUCTION

This review paper sought to review the nuances in the prediction algorithms of the political elections landscape and the role of machine learning classification, specifically ensemble algorithms; for example, how parents' political attributes impact the voting choices of young adults in Ghana.

While there has been some research on the influence of parents' political attributes on voting behavior, most of this research has been conducted in developed countries. Little is known about the influence of parents' political attributes on voting behavior in developing countries like Ghana.

The findings of this study could have important implications for political parties, policymakers, and researchers. For example, political parties could use the findings to tailor their campaigns to young voters. Policymakers could use the findings to design policies that encourage young people to vote. Researchers could use the findings to develop new theories about political socialization.

Voting behavior across nations is shaped by diverse factors encompassing socioeconomic and cultural influences, political activities, etc. Identifying the most influential factor in political voting patterns has proven to be challenging for researchers, as it is difficult to isolate a single factor that determines how voters cast their votes. The transmission of political traits and tendencies from parents to children has been observed in the United States, where children tend to adopt their parents' political activities and affiliations passively. This transmission is influenced by factors such as high literacy rates and education, as children are well-informed about political issues and learn from their parents' choices. However, the literacy level in Africa, particularly in countries like Ghana, is relatively lower, and the research conducted in Palestine by Habashi and Worley (2013) showed that children's political preferences were influenced by extraterritorial factors rather than specific variables. Kim and Lim's (2019) research confirmed a positive link between parental education and offspring's political engagement, examining family processes in 30 countries. The study identified crucial factors, highlighting adolescents' expectations of their education and exposure to politics at home as significant contributors to their intention to vote.

Consequently, the study will contribute to a broader understanding of intergenerational political transmission and the formation of political ideologies. It can shed light on how family socialization processes impact individuals' political beliefs and behaviors. This knowledge can inform discussions on political socialization, democratic processes, and civic education.

The potential, understanding, and ability to predict the influence of parents' political activities on the voting patterns of young adults can help political actors tailor their strategies and messages to engage with this demographic effectively. It will also aid in the development of targeted outreach programs and campaigns to mobilize young voters based on their family backgrounds. Additionally, policymakers can use the findings to design policies and initiatives that foster political awareness and participation among young adults.

Machine learning prediction plays a crucial role in supervised learning, encompassing labeled and continuous datasets. In supervised learning, prediction can take the form of classification, where the predicted variable is categorical, or regression, where it involves continuous datasets. Classification algorithms are designed to handle datasets with categorical labels, such as "Yes" or "No," "1" or "0," and so on (Yaqoob et al., 2023). Some classification algorithms can accommodate nominal or categorical datasets and numerical datasets, such as the Naïve Bayes classifier. Consequently, researchers and pollsters have shown an increasing interest in predicting political election results ahead of time to assess candidates' potential success.

Ensemble classifiers, e.g., stacking ensembles, will serve as a valuable tool for predicting voting patterns in other contexts and countries. By adapting the algorithm to different datasets, researchers can explore the generalizability of findings and identify commonalities or differences in the factors influencing voting behavior across diverse populations.

Moreover, studies have shown that young first-time voters with politically active parents exhibit higher voter turnout rates, indicating the influence of family on voting behavior. Political ideology and positional attitudes are also developed in early adolescence, suggesting that lowering the voting age to 15 could be meaningful. Family upbringing was found to have a greater influence on first-time voters in Norway (Kristensen & Solhaug, 2017). However, the specific ways in which family influences respondents' political lives were not explored in the research. Once the base classifiers are trained, the stacking ensemble algorithm will combine their predictions using a meta-classifier. The meta-classifier takes the outputs of the base classifiers as input features and learns to make a final prediction based on this combined information. The stacking ensemble leverages the strengths of different base classifiers, improving the overall prediction accuracy compared to using a single classifier alone.

Despite the extensive research conducted on various aspects of voting patterns, little work has been done to investigate the influence of parents' political attributes on the voting choices of young adults simply because the dataset does not exist. In recent years, there has been a gradual shift in machine learning towards the use of ensemble classification instead of relying solely on individual classifiers when dealing with large amounts of data. However, it is important to recognize the continued significance of single classification algorithms as they form the fundamental building blocks for creating composite classifiers. This is a consequence of ensemble classifiers ameliorating the defect of classifier bias and variance level (Khan et al., 2024). Additionally, an emergent trend involves the review of the applicability of machine learning algorithms in prognosticating political elections. Furthermore, the stacking ensemble's potential as a versatile predictive tool for diverse voting contexts and nations is worth exploring to benefit data mining.

The performance of the stacking ensemble algorithm is evaluated using appropriate evaluation metrics, such as accuracy, precision, recall, and F-score ROC and AUC, depending on the intent of the user. Cross-validation techniques may also assess the algorithm's robustness and generalization ability. The comparison and experiments can be conducted to evaluate the performance of the stacking ensemble algorithms against other classification algorithms or approaches to ascertain their potency in classifying a given dataset.

REVIEW METHODOLOGY

This study examined a collection of articles obtained from Google Scholar, focusing on the use of ensemble classifiers and machine learning classifiers and their application in predicting political elections through machine learning algorithms. The cited articles reviewed in this work were those published between 2019 and 2023, using keywords like "ensemble classifier," "political election prediction," and "machine learning." The primary focus of the analysis was on the stacking ensemble classifier. Reputable publishers such as IEEE, Springer, MDPI, ResearchGate, and ACM were utilized as the main sources for the articles included in this review. A systematic approach was followed to ensure a thorough and well-organized analysis of the literature.

The review is structured into several sections. The initial section provides an overview of conventional ensemble classifiers, discussing their principles and techniques. The subsequent section delves into articles specifically addressing ensemble classifiers in the context of predicting political elections. Various machine learning classification algorithms applied to political election prediction are also reviewed. The final section of the review presents the conclusions derived from the analyzed literature.

CONVENTIONAL ENSEMBLES

Over the years, the utilization of a solitary classifier in machine learning has greatly contributed to the classification and prediction capabilities of the data mining ecosystem. However, due to certain limitations of these individual classifiers, data scientists have consistently aimed to enhance them by proposing a combined approach known as ensemble classifiers. These ensemble classifiers are designed to overcome or reduce the weaknesses of single classifiers. Classifiers function by making decisions based on predetermined conditions learned from training and testing data. However, confidence in these decisions increases significantly when they are supported by other classifiers that possess similar characteristics. Combining these classifiers into an ensemble enhances the overall decision-making process, resulting in increased confidence levels (Narassiguin, 2018). The employment of a single-classifier predictive model inherently suffers from drawbacks such as high variation, poor accuracy, and susceptibility to feature noise and bias (Alhamid, 2022). To address these issues, an ensemble classifier model is proposed for data analysis and machine learning, aiming to mitigate the severity of these drawbacks associated with using a single classifier. Prominent contributors to the development of conventional ensemble classifier models include Breiman, Schapire, Wolpert, and others. (Yang, 2017) The framework of the ensemble classifier is illustrated in Figures 1, 2, and 3.

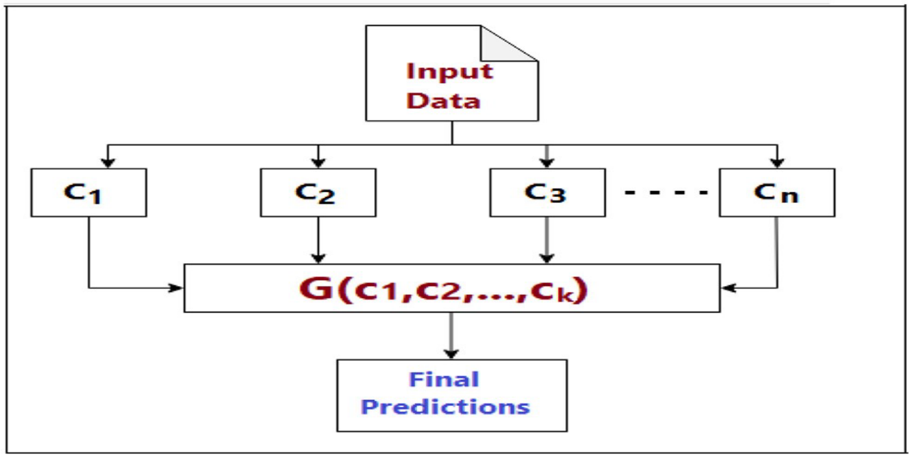

Figure 1. General framework of the ensemble classifier

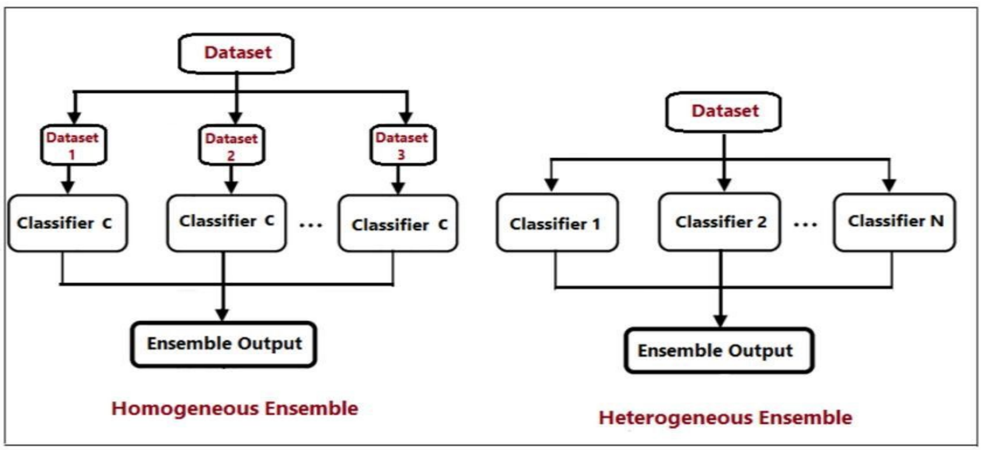

Figure 2. Framework of ensemble groups

In 1996, Leo Breiman introduced the Bagging algorithm, which is an ensemble technique designed to improve the accuracy of classification (Yang, 2017). The concept involves combining the judgments of multiple base classifiers that are trained using bootstrapped training sets. This ensemble classifier, known as bagging, aggregates the predictions of individual weak learners to make more accurate predictions and reduce classification errors (Rocca, 2019). Breiman's work was an advancement over T. K. Ho's random forest algorithm (Polikar, 2009). The bagging algorithm offers several advantages. It can train numerous weak learners, which helps to mitigate overfitting issues that arise when using a single learner on the entire dataset. Additionally, by combining the predictions of multiple learners, the algorithm can reduce bias and eliminate high-variance errors associated with weak learners. Furthermore, the individual bootstraps can be processed independently and in parallel, allowing for efficient computation (Rocca, 2019). One popular example of the bagging algorithm is the random forest. In 1995, T. K. Ho introduced the general method of random decision forests. The random forest algorithm leverages the predictions of decision trees by averaging or aggregating their outputs to determine the outcome. Increasing the number of trees in the random forest improves the accuracy of the results (Mbaabu, 2020). The training process of random trees within the random forest can be performed concurrently, as illustrated in Figure 3.

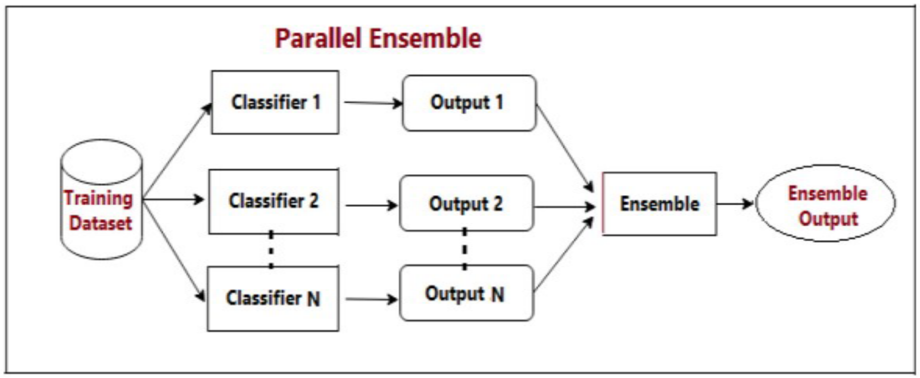

Figure 3. Parallel ensemble nature (Geeksforgeeks, 2023)

The mathematical representation of bagging is given as:

$$\boldsymbol{f(x) = \frac{1}{B}\sum_{B=1}^{B} f_{b(x)}} \quad (1)$$

where $\boldsymbol{fb(x)}$weak learners, $\frac{1}{B}$ generate bootstrapping set
Therefore, for samples D1, D2, D3, from D with replacement.

For $\mathbf{D}\boldsymbol{j}$ train a full decision tree $\boldsymbol{fb}$ (x).

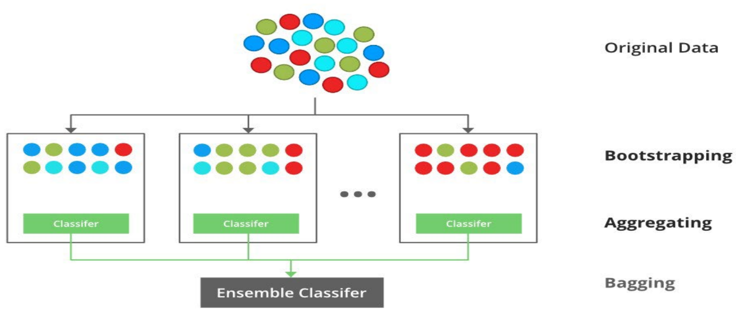

Figure 4. Stages of bagging ensemble (Anusha, 2022)

Figure 4 shows the various stages of the bagging ensemble classifier in a given dataset**.**

BOOSTING

The boosting algorithm introduced by Schapire in 1990 aimed to address the high error rate associated with using a single learner by combining multiple weak learners (Friedman et al., 2000). This approach significantly reduces the error rate compared to using a single learner alone (Walia, 2021). Learning is a challenging problem, and there is no one-size-fits-all algorithm that performs optimally on validation datasets. However, it is important to note that combining classifier algorithms does not automatically enhance accuracy if the base learners are not carefully selected and properly tuned. Additionally, combined classifier algorithms often encounter issues related to runtime and space, which need to be fine-tuned for optimal performance. Boosting, a category of machine learning technique operates on the principle that a combination of basic classifiers generated by weak learners can outperform individual simple classifiers. A weak learner, a type of learning method, is capable of creating classifiers with a marginally lower error probability than random guessing (0.5 in the binary case). Conversely, a strong learner can produce classifiers with arbitrarily low error probabilities when provided with an adequate training dataset.

The weak binary classifier (for m = 1 …, M), and x; $z \in X$ … (1)

Letting E_m: X⟶ {-1, +1} be the m, some input pattern to be classified. There are many ways to combine the outputs E_1 (x)…Em(x) into a single class prediction as an example, we might train several simpler N.N.s and combine their outputs to produce the final output. The function of boosting can be given as:

$$f(x) = \sum_t \alpha_t h_t(x) \tag{2}$$

where the strong classifier $f(x)$ is from several weak classifiers $h_t(x)$. This is, however, done by building a model from the training dataset and then creating a second model that attempts to correct the errors from the earlier model α_t.

Adaptive Boosting, Stochastic Gradient Boosting (SGB), and Extreme Gradient Boosting (XGB), also known as XGBoost, are three types of boost algorithms. Boosting simplifies the understanding of a model and reduces bias and variance in an ensemble of machine-learning models. In a boosting ensemble, the individual models learn one after another, with each subsequent model attempting to correct the mistakes made by the previous models. However, boosting has a drawback: each classifier needs to address the weaknesses of the earlier versions. Implementing sequential training in boosting poses several challenges. As the number of iterations increases, it becomes computationally expensive and more susceptible to overfitting. It is important to note that boosting algorithms may require more time to train compared to bagging because the model's behavior can be influenced by numerous parameters.

Other techniques for enhancing ensemble algorithms in boosting include LightGBM and categorical boosting, known as CatBoost (Zhang et al., 2022). AdaBoost, originally introduced by Freund and Schapire in 1997 (Liu, 2021), was based on the principle of using weighted versions of the same training data rather than randomly selected subsets. This approach differs from earlier boosting methods as it relies heavily on the training set, allowing for smaller datasets to be used effectively. Initially designed to improve the efficiency of binary classifiers, AdaBoost aims to overcome the limitations of the first boosting algorithm. In Gradient Boosting Machines (GBM), the weak learners are decision trees. XGBoost, an enhanced version of GBM, implements parallel preprocessing at the node level, resulting in faster processing compared to GBM. Additionally, XGBoost incorporates various regularization techniques to address overfitting. Many of the advantages of XGBoost, such as parallel training, regularization, and sparse optimization, are also present in LightGBM. However, the tree construction process differs significantly between the two. LightGBM utilizes a leaf-wise splitting strategy, where the node with the highest delta loss is selected for the subsequent split after the initial split. This approach enables LightGBM to handle large datasets efficiently. Moreover, LightGBM often employs a histogram-based technique to select optimal splits and reduce training time. However, LightGBM may not perform well with limited sample sizes.

CatBoost, on the other hand, is an improved version of the XGBoost gradient boosting decision tree (GBRT) algorithm. As the name suggests, CatBoost is particularly suitable for dealing with categorical variables in the data, as it has built-in mechanisms to handle them effectively.

STACK GENERALIZATION

In 1992, Wolpert introduced a technique called stacked generalization, which differs from voting methods by incorporating a combiner system or meta-classifier (Polikar, 2009). This system merges the outputs of base learners in a non-linear manner rather than a linear combination, as in voting methods. The combiner system learns how to appropriately merge the base learners' outputs, even if they are not linearly related, and this combined output is used for making predictions.

To prevent base learners from memorizing the training set, combiner systems must learn from their mistakes. Stacking is employed to estimate and correct the biases of the base learners. Therefore, the combiner system is trained using data that was not used to train the base learners. The tough sum rule, simple majority voting, and weighted majority voting are commonly used procedures for combining ensemble classifiers due to their specific guarantees (Ziweritin, 2022).

Ho (1995) demonstrated that when random forests are restricted to being sensitive to specific feature dimensions, they can improve their accuracy over time without overtraining. In classification problems, the output of a random forest is determined by the majority class chosen by the trees, while in regression tasks, the average prediction of each tree is returned.

However, diversity is a crucial factor in the stacking classification algorithm. Diversity refers to the independence of base classifiers in terms of their operation and predictions, which allows them to complement the weaknesses of the ensemble classifiers (Zhang et al., 2022). Metrics such as Q-statistics in classification and the variance of ensemble predictions around their weighted mean in regression are used to measure diversity. Furthermore, the number of base classifiers in the stack significantly influences the model's performance. Increasing the number of base classifiers traditionally results in a longer runtime for the model.

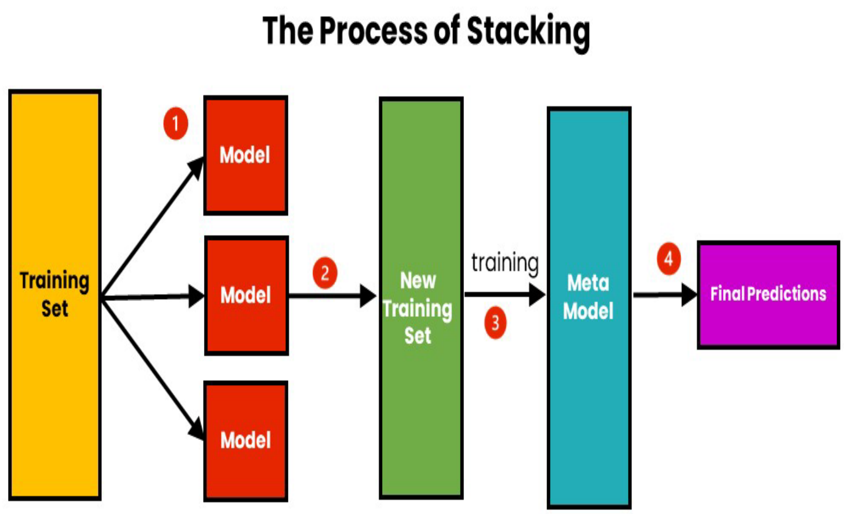

Figure 5. Stacking diagram (Kalirane, 2023)

Figure 5 shows the stages of a stack classifier and its sequential training processes to the final prediction.

The function of stacking is represented as:

$$fs(x) = \sum_{i=1}^{n} a_i fi(x) \qquad (3)$$

Prediction is made from some models (m1, m2, m3, …, mn) to build the desıred model. Consequently, the new model uses a test dataset to make a prediction using the stack of the combined weight of prediction of (mn) a_j, …, (i = 1, 2, 3, … n).

Classification Model Evaluation Matrices

Assessing the effectiveness of a machine learning algorithm is a crucial aspect for experts in the field. Evaluating the performance of a developed model is considered the foremost measure in this regard. It is essential for machine learning models to demonstrate their ability to mimic human decision-making accurately. Therefore, these models must provide evidence of their efficacy in the assigned tasks. Parameters used include Accuracy, Recall, Precision, F-score or F-measure, Mean Absolute Error, and Root Mean Square Error. The latter two, the closer their values are to zero, the better. On the contrary, with the former four, the closer their values are to one, the stronger their efficacy. This is most times expressed in percentages.

CONFUSION MATRIX

The assessment of a classification model's performance can be done in various ways, but one method that has proven its reliability over time is the confusion matrix (Bhandari, 2023). By definition, a confusion matrix is a square matrix with dimensions "N x N" where N represents the total number of classes in the target variable. Its purpose is to evaluate the effectiveness of a classification model by comparing the predicted values with the actual target values. This comprehensive analysis provides valuable insight into the accuracy of the classification model and the types of errors it may be generating. Figure 6 shows the parameters of this matrix.

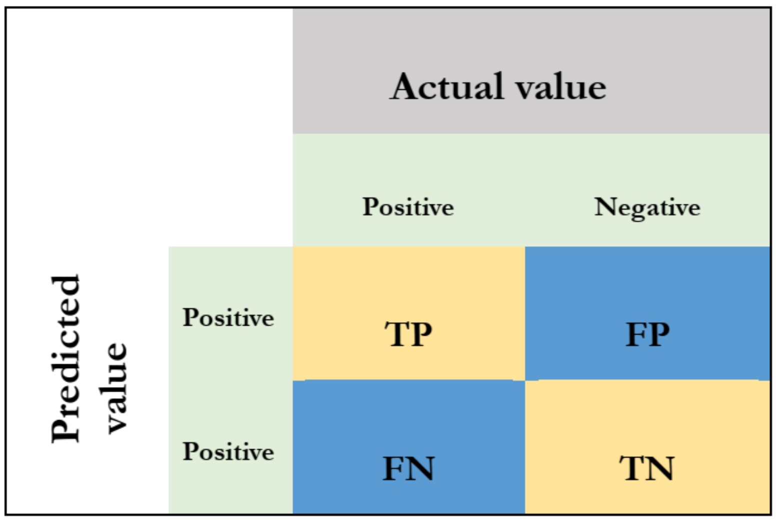

Figure 6. Confusion matrix

TP: True Positive
FP: False Positive
TN: True Negative
FN: False Negative

Accuracy

Accuracy relates to the classification's accuracy for a specific program. It can be described as the proportion of correctly categorized patterns in all samples. The following is a description of the accuracy measurement formula:

$$\boldsymbol{ACCURACY} = \frac{\boldsymbol{TP + TN}}{\boldsymbol{TP + TN + FP + FN}}$$

Precision

Precision is the number of positive samples subtracted from the total number of samples recognized as positive by the classification model. The relation is:

$$\boldsymbol{PRECISION} = \frac{\boldsymbol{TP}}{\boldsymbol{TP} + \boldsymbol{FP}}$$

Recall/sensitivity

The recall is the proportion of genuine positive patterns to the total positive proclaimed patterns. The equation is given as:

$$\boldsymbol{RECALL} = \frac{\boldsymbol{TP}}{\boldsymbol{TP} + \boldsymbol{FN}}$$

F-Measure/F-score

F-score is the harmonic mean of a classifier system's precision and recall values. This matrix measures the predictive performance of a classification algorithm, especially those of binary dataset outputs. The following equation is the mathematical expression.

$$\boldsymbol{F\text{–}SCORE} = 2\,\boldsymbol{X}\,\frac{\boldsymbol{PRECISION\ XRECALL}}{\boldsymbol{PRECISION} + \boldsymbol{RECALL}}$$

Frequently Modeled Conventional Ensemble Classifier

Data scientists and analysts who utilize machine learning classification methods often lean towards adopting a specific traditional ensemble model. This section examines the commonly used conventional ensemble models that have been explored by researchers and data scientists.

One of the challenges faced by researchers is the lack of standardization in techniques for the early detection of mental illnesses based on various everyday human behaviors, such as speech analysis, analysis of social media behavior patterns, and visual activity pattern analysis (Sushma Koushik & Deepu, 2021). In the study, datasets related to depression were utilized to predict mental diseases using both single classifiers and ensemble models. The study statistically compared the performance of these approaches and found that the ensemble model outperformed the majority of single classifiers.

Furthermore, Hema and Kannan (2021) proposed a classification algorithm that aimed to outperform existing methods in recognizing human facial emotions. The proposed ensemble model, called PATCH-SIFT, achieved an accuracy of 98% on a dataset comprising 300 facial images for recognition, surpassing existing classifiers such as L.R., R.F., LDD, KNN, CART, and N.B. The authors emphasized that their proposed approach does not require additional GPU resources or a large dataset to function effectively. The impact of dataset size on classifier models is an aspect that warrants further investigation by researchers in the field. Table 1 shows a summary of the ensemble modeled by some researchers.

Table 1. A summary of composition of ensemble and their models

Year	Author	Ensemble name	Classifiers combined	Conventional ensemble modeled
2023	Wu et al.	SKNN	KNN, LR	Stacking
2023	Thockchom et al.	Not mentioned	GaussianNB, LR, D.T.	Stacking
2021	Hema & Kannan	PATCH SIFT	Discriminant Analysis (LDA) algorithms and R.F.	Stacking
2022	Janardhan & Kumaresh	Not mentioned	(GNB), (SVM), (KNN), (LR), (RF)	Stacking
2021	Yousuf & Khan	Not mentioned	KNN, R.F.	Bagging
2020	Pan et al.	Not mentioned	XGBOOST, R.F., L.R.	Stacking
2019	Fudenberg & Liang	Hybrid decision tree	R.F.	Bagging

Identifying enhancers from DNA sequence data is challenging because they are distributed throughout non-coding regions without specific sequence features. Wu et al. (2023) conducted a study using a stacking classification model called SKNN to predict the presence of enhancers in DNA sequences. The model predicted two classes – strong and weak enhancers. The proposed SKNN model utilized a 5-fold K-nearest neighbors (KNN) approach with a meta-classifier, which was Logistic Regression. The study reported an accuracy of 75%. The dataset used in the study was obtained from a DNA sequence and consisted of nine different cell lines. The dataset was divided into five-folds, and each fold was used to train the KNN classifier. The results from each fold were aggregated into a stack, and the stack was then tested using the Logistic Regression model. It is worth noting that the dataset used in the study was derived from four different databases.

In a separate study by Immanuel Jeo Sherin and Radhika (2022), a stacked classification model was employed to train and test the NSL-KDD dataset using Random Forest as the meta-classifier. The authors compared their results with other known classifiers and suggested that a stacked classifier can perform better when features are reduced. In their methodology, the dataset was split into train and test sets, encoded, and feature extraction was performed before modeling. Similarly, when Thockchom et al. (2023) employed stochastic gradient descent as the meta-classifier and logistic regression, decision trees, and Gaussian Naive Bayes as the base classifier, they demonstrated the effectiveness of ensemble classifiers through stacking. Their model was tested using the chi-square test as the feature selection approach on three datasets: KDD Cup 1999, UNSW-NB15, and CIC-IDS 2017. Their empirical findings demonstrate that the suggested approach model performs better in binary class and multiclass prediction when classifying network intrusion detection. The approach aimed to predict student academic performance at a higher education level. Alwarthan et al.'s (2022) study identified an ensemble classification model that employed a Random Forest Ensemble with a dataset sourced from educational institutes' repositories. The study highlighted the scarcity of research predicting student academic performance in arts and humanities majors using students' data. Olteanu et al. (2022) used the Naive Bayesian suit to compare individual classifiers leveraging machine learning on semi-structured data. The study concluded that different classifiers exhibited varying performances based on different parameters but did not combine these classifiers.

In another study conducted by Tarimo et al. (2022), the researchers developed an ensemble classification algorithm consisting of Adaboost, Gradient Boosting, and Extreme Gradient Boosting (Adaboost, Gboost, and XGBoost). Data processing and prediction were performed using the Python programming language. The study found that the prevalence of low Apgar scores (<7) was 9.5%.

The proposed ensemble models performed similarly to their baseline models. Through the application of resampling methods, borderline-SMOTE significantly improved the predictive performance of all the boosting-based ensemble methods, enhancing sensitivity, F1-score, AUROC, and PPV.

Furthermore, Bakhshipour (2021) employed the Boosting algorithm and Logit Boost for feature filtering and boosting in plant classification based on images. The dataset consisted of 150 colored images of plants from a peanut farm. The combination of boosting algorithms, Linear Discriminant Analysis (LDA), and Random Forest achieved an accuracy of 95%, outperforming similar algorithms. In summary, these studies explored various approaches and models in different domains, ranging from predicting enhancer presence in DNA sequences to student academic performance and neonatal Apgar scores. They utilized stacking classification models, ensemble methods and boosted algorithms to enhance predictive performance and accuracy in their respective fields.

Yousuf and Khan (2021) introduced a novel approach to detecting plant diseases and enabling prompt action on crop farms. Their method involved combining Random Forest and K-nearest neighbors (KNN) algorithms in an ensemble. By employing a dataset containing 1000 images of healthy leaves and three types of plant diseases, they utilized R.F. with K-means clustering for feature selection. Evaluating the proposed model against SVM, they measured parameters like accuracy, precision, and recall. Their findings indicated that the ensemble model they proposed surpassed SVM in terms of performance.

In the study conducted by Gu (2021), a Convolutional Neural Network (CNN) was utilized for the classification of handwritten characters ranging from 0 to 9. The accuracy of the proposed model, using the K-MNIST dataset, reached an impressive 98.77%. The methodology involved the exploration of various CNN layers and the optimization of kernel parameters. To validate the model's performance, it was compared against other classifiers, such as Support Vector Machines (SVM), logistic regression, and baseline CNN models.

Another study by Pan et al. (2020) focused on utilizing XGBoosting, Random Forest, and Logistic Regression in a stacking approach for a classification task. The study employed a dataset of students' performance in entrance examinations, exercises, daily training scores, and other related factors. Feature ranking using XGBoosting was applied to prepare the dataset, and S-fold validation was utilized for dataset operations. The accuracy of the model achieved was 74.80%. However, it was noted that the proposed algorithm had a relatively poor computational time compared to other algorithms. In addition to these approaches, researchers have also explored bagging, a conventional ensemble method. Fudenberg and Liang (2019) proposed an algorithm that combines bagging with a hybrid decision tree for predicting gameplay in a matrix game. This algorithm aimed to uncover irregularities in the initial play of the game.

Overall, these studies demonstrate the application of various machine learning techniques in different domains, showcasing their effectiveness and highlighting the importance of considering factors such as accuracy, computational time, and performance compared to other algorithms.

CLASSIFICATION ALGORITHMS FOR POLITICAL ELECTION PREDICTION

Election prediction has become an essential tool for researchers and politicians, providing valuable insights into the performance of political parties in elections. Machine learning algorithms have gained interest in recent times as a means of predicting election outcomes using datasets.

The ability to correctly predict the outcome of a given event based on past trained classification algorithms is important to the vitality of analyzing data for a given purpose. Consequently, it becomes a challenge when the data needed to train the algorithms is not readily available, for example, the dataset on parents' political activities. In Ghana, for instance, some regions have been tagged as strongholds of some political parties since the Fourth Republic. However, as to which factors contribute to this persistent phenomenon remains unclear. Therefore, it will be a significant leap in the study of politics in Ghana or elsewhere in the world when it is possible to predict young adults' voting patterns based on their parents' political affiliations.

Another study by Polat and Korpe (2022) focused on evaluating the predictive power of demographic traits using natural language processing (NLP) and machine learning techniques. The researchers analyzed Turkish parliamentary debates to estimate the demographic characteristics of deputies, including gender, age, education, occupation, election region, party, and party status. Sub-datasets were created with extracted features, and classification accuracy percentages were determined. The experimental findings demonstrated the effectiveness of using NLP, traditional machine learning, and deep learning approaches in estimating the demographics of deputies based on the debates held between 2012 and 2020.

Examining the influence of macroeconomic factors on elections, Guan and Mani (2022) conducted a study to explain and predict incumbent party losses in federal elections using machine learning models. Various models were compared, and the multilayer perceptron produced the most accurate classification results. However, social sentiments from social media platforms were not included in this particular study, suggesting that incorporating such data could enhance predictive accuracy. Additionally, Ali et al. (2022) presented a five-step process in their study, which aimed to analyze overall election results by quantifying various volumetric social media approaches. Their model achieved an average accuracy of 71% using different classifiers, including deep learning and support vector machines (SVM). Twitter's Tweepy API was utilized to gather data and feedback from the general public. Likewise, most election predictors also rely on Twitter as the source of datasets for sentiment analysis due to its wide use by people to express their views openly on social media (Batra et al., 2020).

Wordliczek (2022) discussed the benefits and limitations of using neural networks in political science research. Despite their shortcomings, neural networks were found to be superior to other networks, especially when dealing with non-stationary datasets over time. This study aimed to provide social scientists with a deeper understanding of the evolving techniques in the machine learning toolbox.

Overall, these studies highlight the increasing importance of machine-learning techniques in election prediction. They demonstrate the potential of using various algorithms and datasets, including social media data, to improve the accuracy of election predictions.

Many classification algorithms and machine learning have featured in the prediction of political elections and have covered dataset sources and their role in comparison with other statistical methods. Kiingati (2021) aims to provide a thorough analysis of these data using statistics. The purpose of this article is to formulate and apply a Bayesian model for the comparison of two front-runners in the race for president. Kenya's presidential election predictions were made using a Bayesian hierarchical model. The outcomes are based on pre-election surveys taken no later than four months before the 2007 elections. Incorporating polling data to account for the evolution of opinions during campaigns, the Bayesian estimation approach, and Kenya's presidential elections, as well as the 2013 general elections, were able to develop a robust methodological option for predicting the outcome of Kenya's presidential elections. According to the findings, if the observed pattern does not hold, the poll-top candidate will prevail. Table 2 is a summary of some authors and their use of machine learning algorithms in political election predictions.

Table 2. Summary of classification algorithms usage and the predictive environment-machine learning in political election prediction

Date	Author	Title of article	Prediction algorithm	Data source	Data size	Dataset date	Prediction environment
2023	Ali et al.	Deep learning-based election results prediction using Twitter activity	SVM, Naïve Bayes, and NN	Twitter	3000 5weets	2018	RapidMiner
2022	Polat & Korpe	Estimation of demographic traits of the deputies through parliamentary debates using machine learning	LR, SVM, FFNN	GNAT website Turkish parliamentary 65,570 stenographic transcripts of individual speeches of 1024 MPs	65,570 stenographic transcripts 1024 MPs	2012–2020	Not Stated
2022	Guan & Mani	Election forecasting using macroeconomic and social indicatorsvia machine learning	LG, SVM, KNN, Naïve Bayes, DT, RF, AdaBoost, XGBoost, and MLP	World Bank	Not stated	1961-2021	Python 3.7
2022	Kiingati	A Bayesian model for forecasting the choice of candidate in a presidential election	Bayesian theorem	Pre-election polls were collected at most four months prior to the 2007 and 2013	1,067 in the current sample	2007 and 2013	Not Stated
2022	Mahendrakar	Feature engineering for election results prediction using machine learning	LR, RF, KNN	Tweeter	1600 Records	2015	Python
2021	Beqiri et al.	Analyses of methods for prediction of elections using software systems	NB, SVM, DT	Tweeter	Not stated	2012	Python
2021	Redl and Hlatshwayo	Forecasting social unrest: A machine learning approach	NN, SVM, RF, DT	IMF's World Economic Outlook	Not stated	1996 to 2020	Not stated
2021	Awais et al.	Leveraging big data for politics: Predicting general election	Bayesian optimization	Twitter	640,000 tweets	2018	Python
2021	Onyenwe et al.	Location-based sentiment analysis of the 2019 Nigeria presidential election using a voting ensemble approach	Voting Ensemble Approach (VEA)	Twitter	583816 bytes	2019	Python

Date	Author	Title of article	Prediction algorithm	Data source	Data size	Dataset date	Prediction environment
2021	Bach et al.	Predicting voting behaviour using digital trace data	XGBoost	2,000 online users who were registered to vote in the 2017 German federal election	1,991	2017	Not stated
2021	Kulkarni et al.	Analyzing political opinions and prediction of election result of the Indian election using data mining approach	Naïve Bayes, SVM and DT	Twitter	60000	Loksabha Election 2019	Not stated
2021	Althnian et al.	Performance, impact of dataset size on classification: An empirical evaluation in the medical domain	SVM, N.N., C4.5 D.T., R.F.	Medical UCI	20		
2020	Immer et al.	Sub-matrix factorization for real-time vote prediction	Generalized Linear Models GLM	Switzerland Binary Munic. 2 196 330 – 1981–2020, US Binary State 50 11 – 1976–2016, Germany Categ. State 16 6 5 1990–2009, Germany Categ. District 538 5 5 1990–2005	Switzerland Binary Munic. 2 196 330 – 1981–2020 US Binary State 50 11 – 1976–2016, Germany Categ. State 16 6 5 1990-2009 Germany Categ. District 538 5 5 1990–2005		Not stated
2020	Skoric et al.	Electoral and public opinion forecasts with social media: A meta-analysis		(1) Twitter; (2) Facebook; (3) Forums (Usenet, PTT, Uwant.com, etc.); (4) Blogs; (5) YouTube; and (6) other platforms.	74 studies	2018	ANCOVA
2020	Şan et al.	Comparison of the accuracy of classification algorithms on three data sets in data mining: Example of 20 classes	KNN, Naive Bayes, SMO, J48, NBM, BAGGING, and JRIP	1975 and 2018, the Higher Education Council's database	6631	1975 and 2018	WEKA
2020	Brito and Adeodato	Predicting Brazilian and US elections with machine learning and social media data	LR, Multilayer Perceptron (MLP) (ANN)	Facebook, Twitter, and Instagram			

Date	Author	Title of article	Prediction algorithm	Data source	Data size	Dataset date	Prediction environment
2018	Houshmand et al.	Disentangling bias and variance in election polls	Bayesian statistical inference.	last three weeks of the campaigns for 608 state-level presidential, senatorial, and gubernatorial elections between 1998 and 2014	4221	1998 and 2014	Bayesian statistical inference.
2018	Mallavarapu et al.	Political profiling using feature engineering and NLP	Random Forest Regressor	Wikipedia pages, Google presence	150	2016	Not stated

In the project conducted by Mahendrakar (2022) in the field of technology, machine learning techniques were utilized to forecast the likelihood of a political party winning an election in India. Various algorithms incorporating machine learning were employed to process and analyze a given dataset using Python programming. The K-Nearest Neighbor Classifier, Random Forest Classification, and Logistic Regression methods were implemented to determine the accuracy of the predictions. The accuracy rates were 87% for logistic regression, 85% for KNN, and 90% for R.F. The dataset used in this study was collected from the social media platform Twitter and focused specifically on India's 2015 elections.

Similarly, the study by Beqiri et al. (2021) explored the potential of software systems in forecasting and predicting election results, with a specific emphasis on social media applications. The literature review examined methods of accessing and analyzing the opinions of potential voters through social media platforms. Python programming, along with powerful libraries such as Numpy for mathematical operations and array manipulation and Matplotlib for graphical representations, was employed to build a predictive model for election outcomes. The required data were obtained using the Tweepy library. Comparing Naive Bayes with two other algorithms, the research found that Naive Bayes exhibited high accuracy in anticipation and classification processes. The study also employed machine learning and lexicon-based methods to determine sentiment scores in tweets, although the classifier was built using a dictionary-based approach. Data mining techniques utilizing Naive Bayes, Support Vector Machine, and Decision Tree algorithms were applied to political and election-related tweets with positive, negative, or neutral labels.

Overall, both studies employed machine learning algorithms, Python programming, and social media data to forecast election results in India. While Mahendrakar (2022) focused on the 2015 elections and used algorithms such as Logistic Regression, K-Nearest Neighbors Classifier, and Random Forest Classification, Beqiri et al. (2021) explored the potential of Naive Bayes, Support Vector Machine, and Decision Tree algorithms in predicting election outcomes by analyzing sentiment scores and classifying tweets. Classification algorithms perform better when the features of the dataset match the capabilities of the chosen classifier. Kulkarni et al. (2021) employed machine learning and a lexicon-based method to detect emotions and predict sentiment scores in tweets. They conducted text mining on politically and election-related tweets using a dictionary-based approach. Utilizing Naïve Bayes, Support Vector Machine, and Decision Tree algorithms, they built a classifier for categorizing test data into positive, negative, and neutral sentiments. The research, focused on 60,000 tweets from three national parties in the 17th Lok Sabha Election of the Indian General Election, concluded that the Naïve Bayes classifier demonstrated superior accuracy to the other classifiers. Figure 6 shows the flow of processes used by Kulkarni et al. (2021) to collect and manipulate tweet datasets. Additionally, Figure 7 shows so frequently used words on Twitter to express sentiments on political elections. Table 3 also is a summary of data collected on the specified political actors in the Nigerian 2019 presidential pre-election tweets.

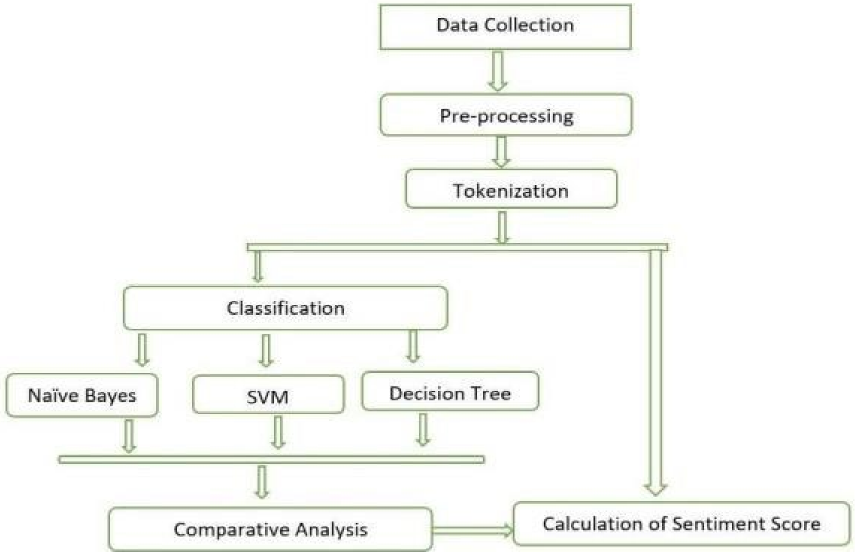

Figure 6. Tweeter data workflow (Kulkarni et al., 2021)

Figure 7. Some frequent words used in the tweet about political election

Table 3. Sample tweet dataset

SN	Political actors collection	Tweet total before preprocessing	Tweet total after preprocessing
1	atiku	52844	3166
2	buhari	94411	7165
3	pdp[1]	101520	4850
4	apc[2]	101894	5591
5	atiku—buhari[3]	23958	1309
6	buhari—apc[4]	15009	607
7	atiku—pdp[3]	8000	295
8	apc— pdp[6]	67464	2611
9	unclassified[7]	118716	27450

Note: adapted from Onyenwe et al. (2022)

[1]Tweet that mentions People's Democratic Party. Presidential candidate =Atiku
[2]Tweet that mentions All Progressives Congress party. Presidential candidate = Buhari.
[3]Tweets that mention Atiku and Buhari.
[4]Tweets that mention Buhari and APC.
[4]Tweets that mention Atiku and PIDP.
[6]Tweets that mention APC and PDP.
[7]Tweets do not fall within S.N. 1-6 mentions

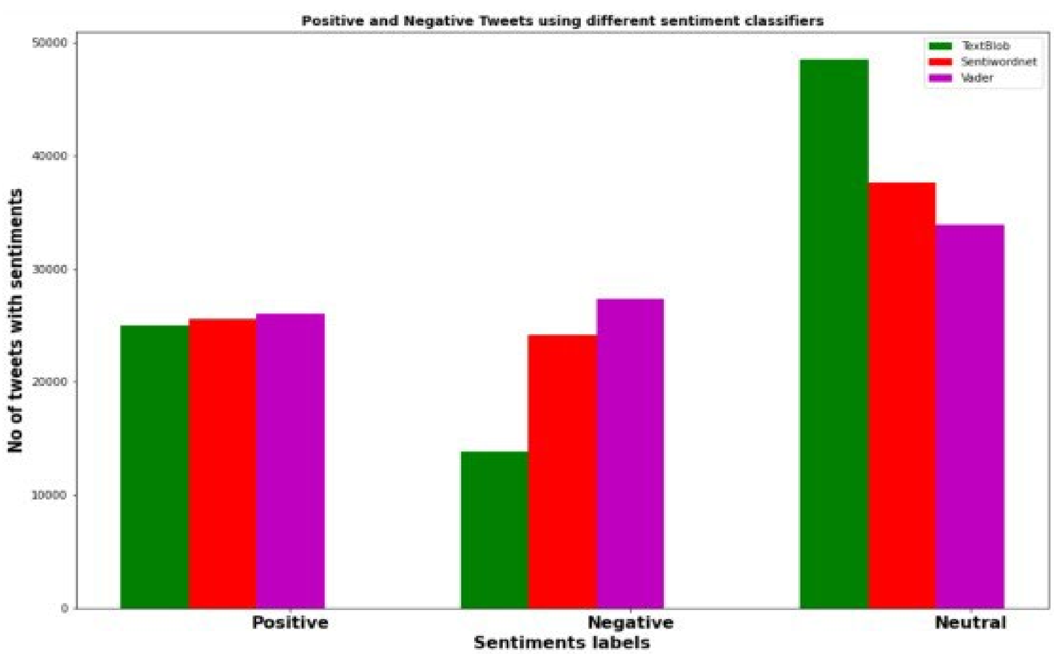

Figure 8. TextBlob, SentiWordNet, and VADER sentiment classifications (Onyenwe et al., 2022)

Figure 8 is a chart of sentiment data expressed by tweets regarding further data collected on sentiments classified as positive, negative, and neutral on the Nigerian elections but using TextBlob, SentiWordNet, and VADER.

Sometimes, it becomes evident that individual expressed opinions vary over time, which can impact the collected data. In a study by Vendeville et al. (2021), the researchers utilized the voter model as the foundation of their investigation. In this model, individuals hold binary opinions and frequently revise them based on new prevailing information. These opinions are shared through a network that incorporates connections with zealots, who are characterized as stubborn agents who do not waver from their beliefs. The study's model yielded an absolute error of 0.4% and a mean absolute error (MAE) of 4.74%. Gayo-Avello (2013) suggested that any model used to forecast election results should have an MAE of no more than 1% or 2%, based on the U.K. general elections' official database results. To obtain the dataset, the researchers utilized the House of Commons as their source. In many countries worldwide, social unrest is a prevalent characteristic of political elections, particularly when such elections are not conducted in a free, fair, and transparent manner. Redl and Hlatshwayo (2021) created a social unrest risk index covering 1996 for 125 countries. This index was established using over 340 indicators that encompassed macroeconomic, socioeconomic, political, and development factors. Using a machine learning model, the researchers predicted social unrest in upcoming elections with an approximate accuracy of two-thirds of the time. The standard likelihood function of the Ridge and Lasso models was modified by incorporating a quadratic penalty term (L2 norm) and a penalty for absolute value (L1 norm) in the input parameters of linear regression. The Ridge model assumes that all predictors are significant in the true model, while the Lasso model assumes that only a small number of predictors matter, making the model sparse. Additionally, Algorithm 1, known as Random Forest Classification, utilizes a neural network and a Support Vector Machine (SVM). The 340 indicators cover a wide range of macroeconomic, socioeconomic, political, and development variables. The Random Forest model achieved a balanced accuracy rate of 65.9%.

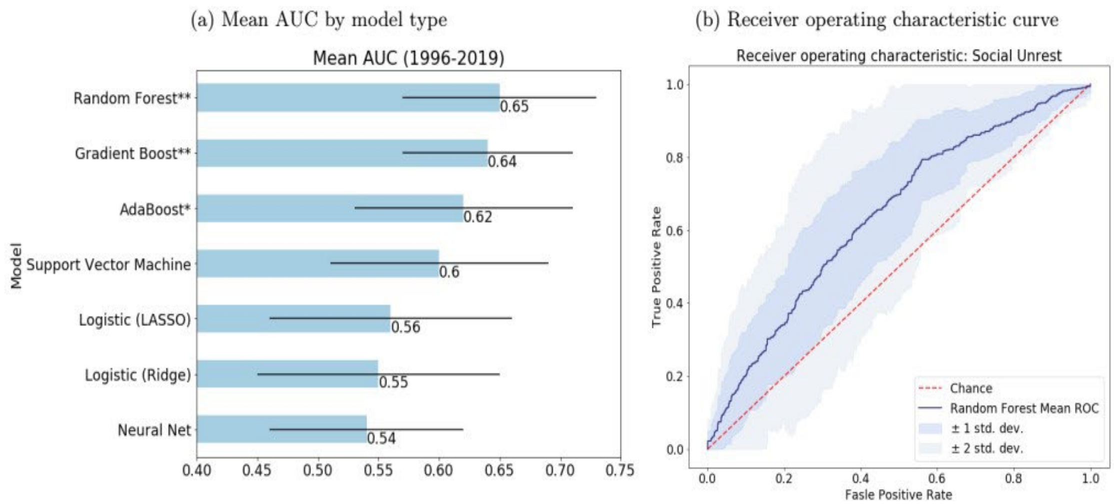

Figure 9. Summary results of models – social unrest index (Redl & Hlatshwayo, 2021)

Figure 9 is a graph of classifiers' performance on the dataset of social unrest index by Redl and Hlatshwayo.

In a study conducted by Awais et al. (2019), a novel machine learning-based model was introduced for election forecasting. The model achieved remarkable success by winning a national competition and demonstrating the highest accuracy in predicting the outcomes of Pakistan's 2018 general election. The key innovation of the model was its utilization of Bayesian optimization to combine probabilities. By analyzing past election data, the model was able to extract valuable insights regarding demographic trends associated with each political party across different districts. Additionally, the model incorporated real-time data from Twitter and approval polls to capture the current levels of popularity for candidates. As a result, it accurately identified the winning candidates for each national assembly seat and predicted the representation of political parties in the national assembly with an impressive 83% accuracy.

Similarly, Onyenwe et al. (2022) conducted a study in 2021 focusing on sentiment analysis of tweets related to the 2019 Nigerian presidential election. They employed a voting ensemble approach (VEA) that combined predictions from multiple techniques to determine the most appropriate polarity of each tweet. To gather data, they utilized the Twitter API to live-stream the voting process during the Nigerian election, resulting in a dataset comprising 583,816 bytes of tweet information. After preprocessing the data, the VEA was applied using three different sentiment classifiers, namely TextBlob, SentiWordNet, and VADER (Valence Aware Dictionary and Sentiment Reasoner). The findings revealed a correlation between election results and the sentiments expressed on Twitter in various locations. However, there were instances where the election results did not align with the sentiment analysis conducted by the model, indicating a lack of similarity.

Overall, both studies employed innovative approaches to analyzing election-related data. While the model developed by Awais et al. (2019) focused on accurately forecasting election outcomes in Pakistan, Onyenwe et al. (2022) concentrated on sentiment analysis of tweets during the Nigerian presidential election. These studies contribute to the growing field of machine learning techniques to gain insights into electoral processes and public opinion.

Grimmer et al. (2021) discussed the integration of machine learning into society and the need to reevaluate applications of machine learning techniques and social science best practices to advance scientific research. The authors argue that machine learning techniques offer more flexibility in manipulating datasets compared to traditional statistical methods, making them valuable for addressing a wide range of research questions in the field of social science. They emphasize the agnostic approach to machine learning methods that focus on social science tasks, particularly in areas such as voting synthesis and text-based analysis, including the utilization of sample splitting, V-fold cross-validation, model fitting with regularization, hyperparameter search, automatic feature engineering, and dimensionality reduction. The authors highlight the limitations of individually applying supervised machine learning to each label, as it ignores the current "best practice" in the field. They point out the lack of information on inter-label associations, which can result in poor model performance. To address this issue, they propose the use of multilabel prediction techniques. Even when the correlations between multiple labels are low, multilabel prediction outperforms conventional supervised learning methods. The authors demonstrate this through simulations, showing that multilabel classification performs better than traditional classification methods when the correlation between target variables increases or the training/test data ratio changes. They also explore the impact of misclassified labels and subset accuracy on text classification within a dataset using a simulated dataset from Mexico's ATI system. Another study by Bach et al. (2021) focused on predicting voting behavior in Germany, a sensitive personal information subject to strict privacy regulations. The researchers employed XGBoost models and 10-fold cross-validation to predict political outcomes using survey information and browsing history data from online users registered to vote in the 2017 German federal election. However, the predictions fall short of the performance standards set by sociodemographic benchmark models. The study concludes that digital trace data, despite some success, are not able to accurately identify undecided voters compared to self-reporting methods. In a different approach, Immer et al. (2020) explored the prediction of overall vote results based on fragmented regional results. The author proposes an algorithm that combines matrix factorization methods with generalized linear models (GLMs) to create an adaptable and precise algorithm. The algorithm learns representations of regions from historical data and uses these representations to fit a GLM and forecast unobserved results. Experimental results demonstrate the algorithm's success in predicting Swiss, U.S., and German legislative elections, as well as presidential elections.

Brito and Adeodato (2020) presented a new approach for training machine learning models to predict vote share, focusing on two scenarios: the 2018 Brazilian presidential election and the 2016 U.S.

Presidential Election. The authors experimented with logistic regression (LR) and multilayer perceptron (MLP) artificial neural network (ANN) algorithms using data from three different sources: Facebook, Twitter, and Instagram. Although the specific data sets used in the experiments are not provided due to erasure, the authors highlight the small size of the data sets. Overall, these studies emphasize the growing integration of machine learning techniques into the social sciences and their potential for advancing research in various areas, such as voting behavior prediction, text classification, and overall vote result forecasting. However, they also highlight the challenges and limitations associated with using machine learning in these contexts, including issues related to inter-label associations, privacy regulations, and performance standards. The obtained results provide statistical evidence ($p<0.05$) to support the claim being made. The errors obtained using both Artificial Neural Network (ANN) strategies were lower than those obtained from the Datafolha survey and the LR (Logistic Regression) method ($p=0.031$ for all four tests). Additionally, when testing the similarity of errors obtained by ANN using fixed and gridded search parameters, a p-value of 1 was found. Subsequent analysis suggests that the ANN with grid search parameters may have produced the best results when using a window of 1 day, with a Mean Absolute Error (MAE) of 0.49 at 28 days. On the other hand, the ANN's fixed parameter settings might have yielded the best results when considering the window size. Two methods were employed for parameterization: manual parameter selection and grid search. Due to the data's characteristics, particularly small samples, the study initially opted for manual parameter selection. The following parameters were chosen: a hidden layer with three neurons to prevent overfitting, the L-FBGS solver known for its effectiveness with small samples, an alpha value of 0.05, a constant learning rate for fast training, and logistic activation. Grid search was also utilized to explore additional parameter settings. Another study conducted by Sanlı et al. (2020) aimed to evaluate the performance of data mining algorithms on three datasets using the WEKA program. Through expert analysis, it was determined that the dataset containing keywords outperformed the other two datasets. Various algorithms, such as Naive Bayes, SMO, J48, NBM, BAGGING, and JRIP, were implemented, while KNN, Decision Tree, Naive Bayes, and Linear SVC were used as classification methods. The decision tree algorithm exhibited the highest performance, achieving an average accuracy of 86.3%. The dataset used in this study was sourced from GitHub and the Twitter library called "tweepy." One of the major challenges in this context is the presence of instances where the outcomes become unobservable after a certain time point or when some instances are missing. Machine learning and survival analysis techniques are considered the most effective to address this issue, a concept referred to as censoring (Ping et al., 2019). The study employed classification strategies using the R programming language. The advantages of machine learning over traditional statistical approaches were discussed, along with recommended classifiers such as Random Forest (RF), Naive Bayes (NB), and Support Vector Machine (SVM). Additionally, the study addressed methods for handling missing data and noise in datasets.

The research conducted by Hunt et al. (2019) examined the reasons why a trading strategy based on forecasts fails to produce abnormal results, while stepwise logit consistently offers reliable predictions about future earnings changes. The study investigates the effectiveness of the elastic net modification of stepwise logit in improving the trading strategy's performance, but it does not yield any significant improvement in return rates. To explore alternative approaches, the study employs Random Forest, a non-parametric machine learning method, and finds that it significantly enhances the accuracy of out-of-sample forecasts, leading to abnormal returns. Ultimately, the study combines the results from all three forecasting techniques to predict the direction of stock market returns. Each of the three models utilizes the same 60 financial ratios and selects a subset to forecast changes in one-year earnings, drawing on the insights of Ou and Penman from 1989 (Ou & Penman, 1989, as cited in Hunt et al., 2019). The findings suggest that recent non-parametric machine learning techniques can benefit various accounting contexts where predictions of binary outcomes are required. In the realm of election prediction, behavioral analysis plays a crucial role. It incorporates geographical remarks, sentiments, and participant psychology to understand and predict election results. Factors such as gender, place of birth, native tongue, and ancestry also influence electoral polls. Moreover, analyzing the impact of events, whether positive or negative, on public opinion is vital for understanding election dynamics. To enhance the accuracy and reliability of polls, the study proposes the SLEPS framework, which involves training dual-layer Neural Nets on a large and diverse dataset. The dataset utilized for this research included 2,100 participants, comprising university undergraduate students and workers from Amazon Mechanical Turk. By employing the SLEPS framework, the study aimed to improve the precision of election outcome predictions based on sentiment analysis. Furthermore, in a study by Zhou and Makse (2019), a machine learning model called Long Short-Term Memory (LSTM) was proposed. The study focused on fine-tuning the model's hyperparameters to improve its accuracy in predicting election results using user or supporter sentiment gathered from Twitter's APIs. The dataset used in this study consisted of textual data collected from posts about the election and political campaigns. The model achieved an accuracy rate of 85%, indicating its effectiveness in estimating the likelihood of winning an upcoming election based on social media sentiment analysis. The growing demand for authenticated olive oils has led to increased research on determining their geographic origins (Gumus et al., 2019). The study utilized Weka and the statistical program MINITAB 15 for analysis. After evaluating various classification schemes, the study identified the most reliable algorithm for authenticating Turkish olive oils. The algorithms assessed included Multilayer Perception, IBK, BayesNet, Naive Bayes, Kstar, SMO, Random Forest, J48, LWL, Logistic Regression, Simple Logistic, and LogitBoost. The study examined 49 olive oils using 61 different chemical analysis parameters and collected samples from six locations in Western Turkey. The BayesNet, Random Forest, and LogitBoost algorithms demonstrated the highest classification accuracies, with values of 93.88%, 91.84%, and 93.88%, respectively. The variety of sampling methods employed in election polls presents a significant challenge. Assessing the total survey error, which encompasses sampling, non-sampling, and other types of errors, is a common practice (Houshmand et al., 2018). However, reported margins of error often overlook non-sampling errors that occur when defining the target population, focusing primarily on sampling variability. To shed light on this issue, the study examined 4,221 polls conducted in the final three weeks of 608 state-level presidential, senatorial, and gubernatorial elections from 1998 to 2014, taking into account the actual election outcomes. The study found that the average survey error, as measured by root mean square error, amounted to approximately 35%, which is roughly twice the magnitude suggested by the majority of reported margins of error. By utilizing election-level bias and variance measures, the researchers decomposed the survey error. Furthermore, the study discovered that the average absolute election-level bias stood at around two percentage points, indicating a consistent source of error across polls conducted for a given election. During the aforementioned period, a total of 4,221 polls were conducted in the final three weeks of 608 state-level presidential, senatorial, and gubernatorial elections between 1998 and 2014. The model generated a posterior estimate bias of 0.2%.

Another study by Mallavarapu et al. (2018) focused on evaluating the effectiveness of utilizing dynamic public data for forecasting election results. While public surveys have been predominantly employed for this purpose, they have displayed limitations in terms of accuracy and reliability. The study employed a Random Forest Regressor with 150 estimators for its initial iteration, resulting in a Pearson correlation of 0.5493 and a mean squared error (MSE) of 0.0355. The dataset was divided into a 50/50 split between training and testing data. The study accessed politician information using Google's Knowledge Graph4 API, which allowed the downloading of relevant data organized according to standard schema specifications. The second iteration of the model yielded a Pearson correlation of 0.79 and an MSE of 0.02. In the study conducted by Wang and Kosinski (2018), deep neural networks were employed to extract features from 35,326 facial images to categorize sexual orientation. These extracted characteristics were then utilized in a logistic regression model. The results showed that the classifier accurately differentiated between gay and heterosexual men based on a single facial image with an 81% success rate while achieving a 71% accuracy for women. These percentages were higher than those achieved by human judgments, which stood at 61% for men and

54% for women. When the algorithm was provided with five facial images for each person, its accuracy increased to 91% for men and 83% for women.

The classifier utilized a combination of fixed features such as nose shape and temporary facial characteristics like grooming style. This study highlights the significant impact of machine learning algorithms across various aspects of human life, extending beyond politics. On a different note, Suhaid et al. (2021) aimed to review feature extraction techniques employed in diverse applications, specifically utilizing a dataset comprising social media messages. The dataset used in this research represented real-world data that exhibited severe imbalances. Through experimentation, it was determined that the TF-IDF approach, in conjunction with stop-word deletion and stemming, yielded the best overall results for this particular dataset. The review concluded that not all feature extraction techniques enhance classification performance, and no single technique outperforms others when applied to different classifiers. Therefore, it is crucial to consider the data structure when selecting an appropriate extraction method. Figure 10 shows a summary of frequently used algorithms in political election predictions.

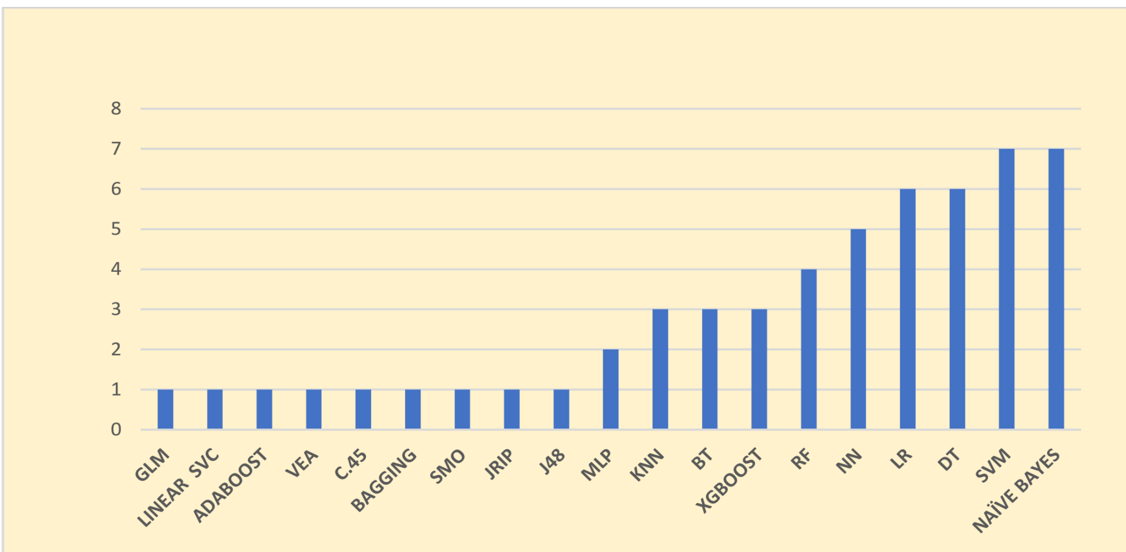

Figure 10. Summary of frequent algorithms used in political predictions

DATASET CLEANING AND USAGE METHODS

Parents' political activities and their influence on children's voting patterns as first voters or political engagements is pronounced in early adolescence but fluctuates in late adolescence in Britain (Janmaat & Hoskins, 2022). This goes to say that in whichever way parents engage in any form of political activity, it has some impact on their children's political choices at voting age. Though such data on parents' activities may not be intentionally gathered and stored, it could be useful for later prediction of future children's political engagements.

The study conducted by Skoric et al. (2020) presents the findings of a meta-analysis that focused on examining the predictive power of social media data using sentiment analysis and structural feature analysis. The analysis utilized various data sources and prediction techniques, employing the ANCOVA tests as the methodology. The results, derived from 74 published studies, revealed significant variation in the accuracy of predictions, with the average performance falling behind traditional survey research benchmarks. The study found that sentiment analysis and a combination of structural features yielded the most accurate predictions, with machine learning-based estimates generally outperforming pre-existing lexica.

However, determining the appropriate sample size for these datasets has been a challenge for researchers in this field. In a study by Rajput et al. (2023), the researchers aimed to establish standards for assessing sample size in machine learning. They evaluated the performance of five machine learning methods on real and simulated datasets and examined the effect sizes and classification accuracy. The findings revealed that as the sample size increased, the effect sizes and classification accuracy improved, particularly in datasets with strong class distinctions. However, for indeterminate datasets, increasing the sample size did not lead to improved effect sizes or classification accuracy.

In a study by Haralabopoulos et al. (2020), a stacked ensemble approach with a weighted average was proposed. It demonstrated good accuracies on two datasets: the Semeval 2018 Task 1 Dataset (SEM2018) and the Toxic Comments Dataset from Kaggle (TOXIC). The stacked ensemble, employing a layered architecture and weights derived from each dataset's mean and median, achieved 87% and 84% accuracy for the SEM2018 dataset and 97.76% and 97.82% for the TOXIC dataset.

Another research by Madhu et al. (2019) focused on addressing missing data in machine learning using a novel imputation method based on XGBoost ensemble algorithms. The proposed method was compared to the imputation techniques of KNN and Miss Forest using benchmark medical datasets with missing values ranging from 1.98% to 50.65%. The results demonstrated that the proposed method outperformed the other techniques in terms of RMSE, Accuracy, and Variance. Handling missing data during dataset preparation is crucial for ensuring data cleanliness and facilitating further processing. Althnian et al. (2021) conducted a study intending to investigate the impact of dataset size on the performance of six popular supervised machine-learning models in the medical industry. The research involved in-depth testing of six classification models, including support vector machines (SVM), neural networks (NN), C4.5 decision trees (DT), random forest (RF), Adaboost (AB), and Naive Bayes (NB), using 20 medical UCI datasets. The study evaluated how the models' performance changed with smaller datasets in terms of accuracy, precision, recall, f-score, specificity, and area under the ROC curve. Statistical tests were applied to compare the performance differences in various scenarios using three randomly selected small datasets obtained through sampling without replacement. The chosen datasets were obtained from the UCI data repository and represented medical specialties with limited availability.

Although the definition of small datasets is not explicitly specified in the literature, the study concluded that dataset size may not be a barrier to achieving high-performing models, depending on the problem domain. The research indicated that specific compact datasets yielded exceptional classifier performance, with an average accuracy of 99%. These results emphasize that the choice of model plays a significant role, even when working with large datasets. Table 4 shows the models' performances of the Althnian et al. (2021) study on the SEM2018 dataset and consequently, Figure 11 is a graph of the mean square error rates of the models used.

Table 4. Result of accuracy score of SEM2018 datasets

Model	Train	Dev.
NN	0. 8981	0-9013
Model1	0. 9759	0-9643
Model2	0. 9637	0-9623
Model3	0.9777	0-9672
Model4	0. 9776	0.968
Model5	0.9768	0-9623
Stacked en.	0.9776	0-9667
Weighted en.	0-9782	0-9698

Figure 11. Result of the proposed model with missing data - SEM2018 datasets (Althnian et al., 2021)

Feature Selection in Ensemble Algorithm

Dataset preprocessing plays a crucial role in data preparation for prediction, and a key aspect of this process is feature selection. The techniques and procedures involved in feature selection are quite extensive. The machine learning model, trained on a specific dataset, provides predictions for test data based on the knowledge gained during training. However, it is important to determine which input data is necessary for training the model by eliminating redundant and unnecessary information. This step is significant as it reduces the dimensionality of the data, leading to improved performance and decreased time complexity of the model.

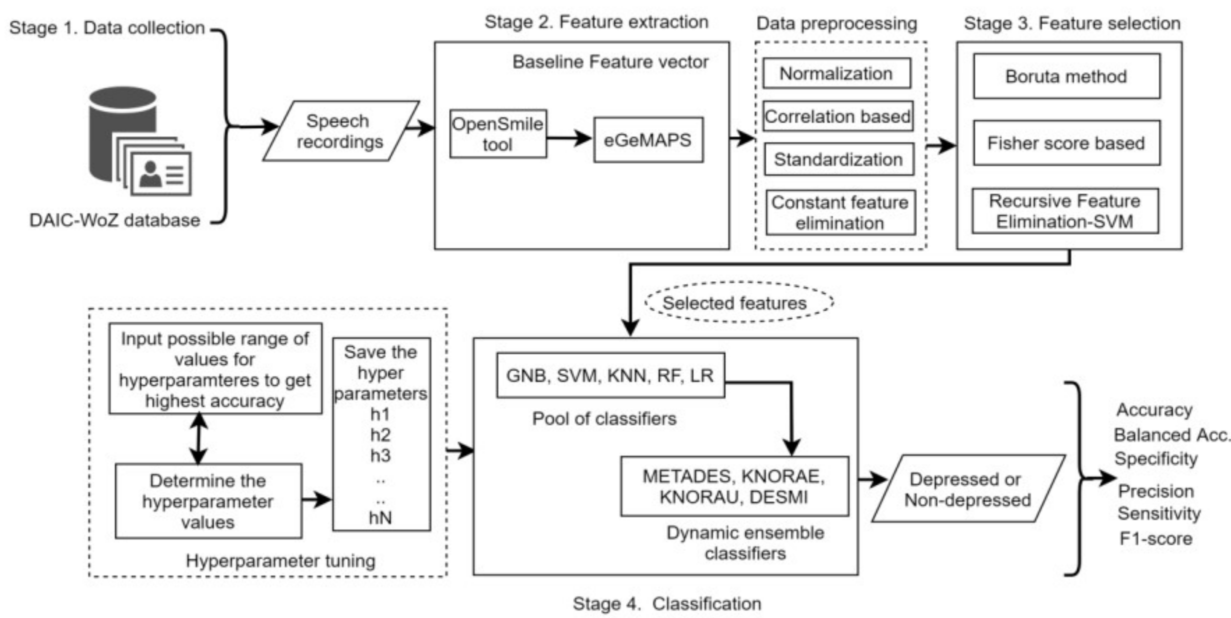

Figure 12. Architecture of features selection model (Janardhan & Kumaresh, 2022)

In a study by Janardhan and Kumaresh (2022), the researchers explored the potential of speech recognition prediction in enhancing depression prediction. They proposed and utilized four classifiers along with three feature selection techniques: Boruta FS, recursive feature elimination using support vector machine (SVM-RFE), and fisher score-based FS. The researchers achieved an accuracy of 81% by constructing a classifier model from Gaussian Naïve Bayes, SVM, KNN, LR, and RF using multiple Dynamic Ensemble Selection (DES) methods. Notably, KNN outperformed the other classifiers when using a subset of 15 features. The dataset employed in the study was obtained from the DAIC-WOZ database. Janardhan and Kumaresh (2022) implemented a four-stage classification modeling approach, incorporating the Open Smile feature selection technique. Python programming in ensembled data analysis provides an opportunity to manipulate feature selection of data, which enhances the stability of the dataset with runtime (Schowe & Morik, 2011). Figure 12 is a diagram of Janardhan and Kumaresh's (2022) architecture of their feature selection model.

SUMMARY OF FINDINGS

In the foregoing literature review, it has emerged that machine learning classification algorithms and their accompanying activities have grounded their significance in politics and election predictions, using single, ensemble, or deep learning classifiers. However, this review, which centered mainly on single or ensemble (stack) classifiers, has exposed the nitty-gritty of their application in political elections or its subsidiary election predictions.

Single/Base classifiers: These classifiers, Support Vector Machine, Random Forest, Naive Bayes, Logistic Regression, and Decision Tree, are among but a few single classifiers that have proven to be strong in handling datasets concerning election predictions so far. This was shown by the comparatively high accuracies they gave to the datasets they applied to.

Ensemble classifiers: Moreover, ensemble classifiers that researchers used had better accuracy, but few employed them in political election predictions, including deep learners. This might be due to the complexities in their applications.

Datasets: The dataset for predicting the outcome of political elections has been that of tweets and comments of people generated online.

Data-preprocessing: Imputation and bag of words were some of the main methods of data preprocessing approaches with Feature selection, like Open Smile.

Predictive Environment: Most studies did categorically state their prediction environment, but it appeared in the review that those who did use R programming, python, RapidMiner, and Weka.

CONCLUSION

The scarcity of easily accessible datasets concerning the political engagements of parents presents a notable challenge in contemporary research. Addressing this gap is pivotal for comprehensively understanding political behavior and its implications. Meanwhile, the emergence of stack classifiers holds substantial promise in political election classification. Leveraging multiple classifiers in a hierarchical structure, stack classifiers have demonstrated the potential to enhance the accuracy and robustness of election-related predictions. Furthermore, researchers have employed ensemble classifiers in political election scenarios, revealing marked advancements in machine learning performance metrics. These ensemble methods, which amalgamate predictions from diverse classifiers, notably boost overall predictive capabilities. Consequently, the integration of stack and ensemble classifiers addresses data limitations and signifies a progressive leap toward more effective and nuanced analyses of political activities and elections.

In conclusion, a significant gap exists in the research focused on gathering data concerning the political activities of parents and their potential influence on the voting preferences of their adult children, both in Ghana and globally. This critical area remains largely unexplored, representing a missed opportunity to understand the intricate dynamics shaping voting patterns across generations. To bridge this knowledge void, future studies should prioritize comprehensive data collection on parental political engagement and its impact on the voting behavior of their offspring.

Moreover, the potential of machine learning classification algorithms, particularly stack ensembles, in predicting political elections remains underutilized. Embracing these advanced computational tools can offer more accurate and robust predictions, enhancing the overall efficacy of electoral forecasting models. Additionally, the incorporation of machine learning techniques in data preprocessing, specifically in feature selection for political datasets, holds the potential to bolster dataset integrity. This strategic use of technology can instill greater confidence in the predictive outcomes, contributing to a more nuanced understanding of the factors influencing political landscapes. As the field of political science continues to evolve, leveraging machine learning tools becomes imperative for refining analytical approaches and gaining deeper insights into the intricate interplay of familial influences and electoral choices.

FURTHER RESEARCH DIRECTION

While Twitter can provide valuable insights into public sentiment, relying solely on it for political election prediction may not be sufficient. To ensure a more comprehensive analysis, alternative data sources can be explored. News articles, opinion polls, surveys, and social media platforms beyond Twitter could be considered. Additionally, academic research, demographic data, historical voting patterns, and economic indicators can offer a broader perspective. Integrating multiple data sources would provide a more diverse and reliable dataset for political election prediction, minimizing potential biases and capturing a wider range of public opinions and trends.

Additionally, ensemble classification algorithms have indeed demonstrated superior performance when carefully chosen and combined. Given their success, further research should be dedicated to advancing ensemble methods for machine learning classification. By exploring novel techniques for classifier selection, fusion, and diversity generation, researchers can unlock even greater potential. Improved ensemble algorithms could enhance model accuracy, robustness, and generalization, making them valuable tools across various domains. Additionally, investigations into ensemble learning's theoretical foundations and interpretability could lead to a better understanding of its inner workings. Continued research in ensemble classification algorithms promises to push the boundaries of machine learning and contribute to more reliable and effective classification solutions.

REFERENCES

Alhamid, M. (2022, May 12). Ensemble models: What are they and when should you use them? *Built In*. https://builtin.com/machine-learning/ensemble-model

Ali, H., Farman, H., Yar, H., Khan, Z., Habib, S., & Ammar, A. (2022). Deep learning-based election results prediction using Twitter activity. *Soft Computing, 26,* 7535–7543. https://doi.org/10.1007/s00500-021-06569-5

Althnian, A., AlSaeed, D., Al-Baity, H., Samha, A., Dris, A. B., Alzakari, N., Abou Elwafa, A., & Kurdi, H. (2021). Impact of dataset size on classification performance: An empirical evaluation in the medical domain. *Applied Sciences*, *11*(2), 796. https://doi.org/10.3390/app11020796

Alwarthan, S. A., Aslam, N., & Khan, I. U. (2022). Predicting student academic performance at higher education using data mining: A systematic review. *Hindawi*, *2022*, Article 8924028. https://doi.org/10.1155/2022/8924028

Anusha, H. (2022, January 28). Bootstrapped aggregation (bagging). *Medium*. https://medium.com/@hemaanushatangellamudi/bootstrapped-aggregation-bagging-481f4812e3ea

Awais, M., Hassan, S.-U., & Ahmed, A. (2019). Leveraging big data for politics: Predicting general election. *Journal of Ambient Intelligence and Humanized Computing*, *12*, 4305–4313. https://doi.org/10.1007/s12652-019-01378-z

Bach, R. L., Kern, C., Amaya, A., Keusch, F., Kreuter, F., Hecht, J., & Heinemann, J. (2021). Predicting voting behaviour using digital trace data. *Social Science Computer Review, 39*(5), 862-883. https://doi.org/10.1177/0894439319882896

Bakhshipour, A. (2021). Cascading feature filtering and boosting algorithm for plant type classification based on image features. *IEEE Access*, *9*, 82021-82030. https://doi.org/10.1109/ACCESS.2021.3086269

Batra, P. K., Saxena, A., Shruti, S., & Goel, C. (2020, November). Election result prediction using Twitter sentiments analysis. *Proceedings of the Sixth International Conference on Parallel, Distributed and Grid Computing, Waknaghat, India,* 182-185. https://doi.org/10.1109/PDGC50313.2020.9315789

Beqiri, L. Z., Zdravev, Z., Fetaji, M., & Fetaji, B. (2021). Analyses of methods for prediction of elections using software systems. *International Journal of Sciences: Basic and Applied Research*, *57*(2), 86-92. https://eprints.ugd.edu.mk/29672/1/12571-Article%20Text-37728-1-10-20210424.pdf

Bhandari, A. (2023, May 2). Understanding & interpreting confusion matrix in machine learning. *Analytics Vidhya*. https://www.analyticsvidhya.com/blog/2020/04/confusion-matrix-machine-learning/#:

Brito, K. d. S., & Adeodato, P. J. L. (2020). Predicting Brazilian and U.S. elections with machine learning and social media data. *IEEE*, 8. https://doi.org/10.1109/IJCNN48605.2020.9207147

Friedman, J., Hastie, T., & Tibshirani, R. (2000). Additive logistic regression: A statistical view of boosting. *Annals of Statistics*, *28*(2), 337-407. https://doi.org/10.1214/aos/1016218223

Fudenberg, D., & Liang, A. (2019). Predicting and understanding initial play. *American Economic Review*, *109*(12), 4112-41. https://doi.org/10.1257/aer.20180654

Gayo-Avello, D. (2013). A meta-analysis of state-of-the-art electoral prediction from Twitter data. *Social Science Computer Review*, *31*(6), 649-679.. https://doi.org/10.1177/0894439313493979

Geeksforgeeks. (2023, November 14). *Bagging vs boosting in mcahine learning.* https://www.geeksforgeeks.org/bagging-vs-boosting-in-machine-learning/

Grimmer, J., Roberts, M. E., & Stewart, B. M. (2021). Machine learning for social science: An agnostic approach. *Annual Review of Political Science*, *24*(1), 395-419. https://doi.org/10.1146/annurev-polisci-053119-015921

Gu, E. X. (2021, January). Convolutional neural network based Kannada-MNIST classification. *Proceedings of the IEEE International Conference on Consumer Electronics and Computer Engineering, Guangzhou, China,* 180-185. https://doi.org/10.1109/ICCECE51280.2021.9342474

Guan, L., & Mani, G. (2022). Election forecasting using macroeconomic and social indicators via machine learning. Guan, L., & Mani, G. (2022). Election Forecasting Using Macroeconomic and Social Indicators via Machine Learning. *Journal of Student Research*, *11(3).* https://doi.org/10.47611/jsrhs.v11i3.3508

Gumus, O., Yasar, E., Pinar, P. G., & Ertas, H. (2019). Comparison of different classification algorithms to identify geographic origins of olive oils. *Journal of Food Science and Technology*, *57*, 1535–1543. https://doi.org/10.1007/s13197-019-04189-4

Habashi, J., & Worley, J. A. (2013). Children's projected political preference: Transcending local politics. *Children's Geographies*, *12*(2), 205-218. https://doi.org/10.1080/14733285.2013.812306

Haralabopoulos, G., Anagnostopoulos, I., & McAuley, D. (2020). Ensemble deep learning for multilabel binary classification of user-gene. *Algorithms*, *13*(4), 83. https://doi.org/10.3390/a13040083

Hema, D., & Kannan, S. (2021). Patch-SIFT: Enhanced feature descriptor to learn human facial emotions using an ensemble approach. *Indian Journal of Science and Technology*, *14*(21), 1740-1747. https://doi.org/10.17485/IJST/v14i21.2261

Ho, T. K. (1995). Random decision forests. Proceedings of 3rd *International Conference on Document Analysis and Recognition*, Montreal, QC, Canada, pp. 278-282. https://doi.org/10.1109/ICDAR.1995.598994

Houshmand, S. M., Rothschild, D., Sharad, G., & Gelman, A. (2018). Disentangling bias and variance in election polls. *Journal of the American Statistical Association*, 113(522), 607-614. http://www.stat.columbia.edu/~gelman/research/published/polling-errors.pdf

Hunt, J. O., Myers, J. N., & Myers, L. A. (2019). *Improving earnings predictions with machine learning* [Working Paper]. https://zicklin.baruch.cuny.edu/wp-content/uploads/sites/10/2019/12/Improving-Earnings-Predictions-with-Machine-Learning-Hunt-Myers-Myers.pdf

Immanuel Jeo Sherin, V. J., & Radhika, N. (2022). Stacked Ensembled-IDSI using NSL-KDD dataset. *Journal of Pharmaceutical Negative Results*, *13*(3), 351-356. https://doi.org/10.47750/pnr.2022.13.S03.057

Immer, A., Kristof, V., Grossglauser, M., & Thiran, P. (2020). Sub-matrix factorization for real-time vote prediction. *Proceedings of the 26th SIGKDD International Conference on Knowledge Discovery* (pp. 2280–2290). ACM. https://doi.org/10.1145/3394486.3403277

Janardhan, N., & Kumaresh, N. (2022). Improving depression prediction accuracy using fisher score-based feature selection and dynamic ensemble selection approach based on acoustic features of speech. *Traitement du Signal, 39*(1), 87-107. https://doi.org/10.18280/ts.390109

Janmaat, J. G., & Hoskins, B. (2022). The changing impact of family background on political engagement during adolescence and early adulthood. *Social Forces, 101*(1), 227–251. https://doi.org/10.1093/sf/soab112

Kalirane, M. (2023, November 17). Ensemble learning in machine learning: Bagging, boosting and stacking. *Analytic Vidhya*. https://www.analyticsvidhya.com/blog/2023/01/ensemble-learning-methods-bagging-boosting-and-stacking/

Khan, A. A., Chaudhari, O., & Chandra, R. (2024). A review of ensemble learning and data augmentation models for class imbalanced problems: Combination, implementation and evaluation. *Expert Systems with Applications*, *244*, 122778. https://doi.org/10.1016/j.eswa.2023.122778

Kiingati, J. K. (2021). *A Bayesian model for forecasting the choice of candidate in a presidential election* [Doctoral dissertation, Jomo Kenyatta University of Agriculture and Technology]. http://ir.jkuat.ac.ke/bitstream/handle/123456789/5772/JEREMIAH%20KIINGATI%20PRINT%20FINAL.pdf?sequence=1&isAllowed=y

Kim, H., & Lim, E. (2019). A cross-national study of the influence of parental education on intention to vote in early adolescence: The roles of adolescents' educational expectations and political socialization at home. *International Journal of Adolescence and Youth, 24*(1), 85-101. https://doi.org/10.1080/02673843.2018.1470993

Kristensen, N. N., & Solhaug, T. (2017). Party choice and family influence in the age of late modernity – Reflections as first-time voters in a Norwegian parliamentary election. *Nordidactica – Journal of Humanities and Social Science Education*, *7(4)*, 48-69. https://journals.lub.lu.se/nordidactica/article/view/19057

Kulkarni, A. A., Dhopeshwarkar, M. G., & Jaiswal, P. S. (2021). Analyzing political opinions and prediction of election result of the Indian election using data mining approach. *International Journal of Creative Research Thoughts*, *9*(12), b494-b503. https://ijcrt.org/papers/IJCRT2112152.pdf

Liu, H. (2021, February 26). Chapter 5 - Data mining and processing for train unmanned driving systems. In H. Liu, *Unmanned Driving Systems for Smart Trains* (pp. 211-252). https://doi.org/10.1016/B978-0-12-822830-2.00005-2

Madhu, G., Bharadwaj, B., Nagachandrika, G., & Vardhan, K. (2019, November). A novel algorithm for missing data imputation on machine learning. *Proceedings of the International Conference on Smart Systems and Inventive Technology, Tirunelveli, India*, 173-177. https://doi.org/10.1109/ICSSIT46314.2019.8987895

Mahendrakar, S. (2022). Feature engineering for election results prediction using machine learning. *International Journal of Innovative Research in Technology*, *9*(1), 1248-1251. https://ijirt.org/master/publishedpaper/IJIRT155644_PAPER.pdf

Mallavarapu, C., Mandava, R., Kc, S., & Holt, G. m. (2018). Political profiling using feature engineering and NLP. *SMU Data Science Review*, *1*(4), Article 10. https://scholar.smu.edu/datasciencereview/vol1/iss4/10

Mbaabu, O. (2020, December 11). *Introduction to random forest in machine learning*. Webscale. https://www.webscale.com/engineering-education/introduction-to-random-forest-in-machine-learning/

Narassiguin, A. (2018). *Ensemble learning, comparative analysis, and further improvements with dynamic ensemble selection* [Doctoral dissertation, Université de Lyon]. https://theses.hal.science/tel-02146962/document

Olteanu, A., Cernian, A., & Gâgă, S.-A. (2022). Leveraging machine learning and semi-structured information to identify political views from social media posts. *Applied Sciences*, *12*(24), 12962. https://doi.org/10.3390/app122412962

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 the 2019 Nigeria presidential election using a voting ensemble approach*. Research Square. https://doi.org/10.21203/rs.3.rs-2104488/v1

Pan, F., Yuan, Y., & Song, Y. (2020). Students' classification model based on stacking algorithm. *Journal of Physics: Conference Series*, *1486*, 032020. https://doi.org/10.1088/1742-6596/1486/3/032020

Ping, W., Li, Y., & Reddy, C. K. (2019). Machine learning for survival analysis: A survey. *ACM Computing Surveys*, *51*(6), Article 110. https://doi.org/10.1145/3214306

Polat, H., & Korpe, M. (2022). Estimation of demographic traits of the deputies through parliamentary debates using machine learning. *Electronics*, *11*(15), 2374. https://doi.org/10.3390/electronics11152374

Polikar, R. (2009, February 26). *Scholarpedia*. http://www.scholarpedia.org/article/Ensemble_learning

Rajput, D., Wang, W. J., & Chen, C. C. (2023). Evaluation of a decided sample size in machine learning applications. *BMC Bioinformatics*, *24*, Article 48. https://doi.org/10.1186/s12859-023-05156-9

Redl, C., & Hlatshwayo, S. (2021). Forecasting social unrest: A machine learning approach. *IMF Working Papers*, *2021*(263). https://doi.org/10.5089/9781557758873.001

Rocca, J. (2019, April 13). Ensemble methods: Bagging, boosting, and stacking. *Medium*. https://towardsdatascience.com/ensemble-methods-bagging-boosting-and-stacking-c9214a10a205

Sanlı, T., Sıcakyüz, Ç., & Yüregir, O. H. (2020). Comparison of the accuracy of classification algorithms on three datasets in data mining: Example of 20 classes. *International Journal of Engineering, Science and Technology*, *12*(3), 81-89. https://doi.org/10.4314/ijest.v12i3.8

Schowe, B., & Morik, K. (2011). Fast-ensembles of minimum redundancy. In O. Okun, G. Valentini, & M. Re (Eds.), *Ensemble in Machine Learning Application* (pp. 75-95). Springer. https://doi.org/10.1007/978-3-642-22910-7_5

Skoric, M. M., Liu, J., & Jaidka, L. (2020). Electoral and public opinion forecasts with social media: A meta-analysis. *Information*, *11*(4), 187. https://doi.org/10.3390/info11040187

Suhaid, M., Kadir, R. A., & Tiun, S. (2021). A review of feature extraction methods on machine learning. *Journal of Information Systems and Management*, *6*(22), 51-59. https://doi.org/10.35631/JISTM.622005

Sushma Koushik, N., & Deepu, R. (2021). Analyzing mental health evaluation techniques from a statistical perspective. *Proceedings of the 3rd International Conference on Integrated Intelligent Computing Communication & Security* (pp. 373-382). Atlantis Press. https://doi.org/10.2991/ahis.k.210913.046

Tarimo, C. S., Bhuyan, S. S., Zhao, Y., Ren, W., Akram, M., Li, Q., Gardner, M., Mahande, M. J., Yuhui, W., & Wu, J. (2022). Prediction of low Apgar score at five minutes following labor induction intervention in vaginal deliveries: Machine learning approach for imbalanced data at a tertiary hospital in North Tanzania. *BMC Pregnancy Childbirth*, *22*, Article 275. https://doi.org/10.1186/s12884-022-04534-0

Thockchom, N., Singh, M. M., & Nandi, U. (2023). A novel ensemble learning-based model for network intrusion detection. *Complex & Intelligent Systems*, *9*, 5693–5714. https://doi.org/10.1007/s40747-023-01013-7

Vendeville, A., Guedj, B., & Zhou, S. (2021). Forecasting elections results via the voter model with stubborn nodes. *Applied Network Science*, *6*, Article 1. https://doi.org/10.1007/s41109-020-00342-7

Walia, M. S. (2021, April 27). Best boosting algorithm in machine learning in 2021. *Analytic Vihya*. https://www.analyticsvidhya.com/blog/2021/04/best-boosting-algorithm-in-machine-learning-in-2021/

Wang, Y., & Kosinski, M. (2018). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology*, *114*(2), 246-257. https://doi.org/10.1037/pspa0000098

Wordliczek, Ł. (2022). Neural networks and political science: Testing the methodological frontiers. *EMPIRIA. Revista de Metodología de Ciencias Sociales*, *57*, 37-62. https://www.redalyc.org/journal/2971/297176235002/297176235002.pdf

Wu, H., Liu, M., Zhang, P., & Zhang, H. (2023). iEnhancer-SKNN: A stacking ensemble learning-based method for enhancer identification and classification using sequence information, *22*(3), 302-311. https://doi.org/10.1093/bfgp/elac057

Yang, Y. (2017). Chapter 4 - Ensemble Learning. In Y. Yang, *Temporal data mining via unsupervised ensemble learning* (pp. 35-36). https://doi.org/10.1016/B978-0-12-811654-8.00004-X

Yaqoob, A., Musheer Aziz, R., & Verma, N. K. (2023). Applications and techniques of machine learning in cancer classification: A systematic review. *Human-Centric Intelligent Systems*, *3*, 588-615. https://doi.org/10.1007/s44230-023-00041-3

Yousuf, A., & Khan, U. (2021). Ensemble classifier for plant disease detection. *International Journal of computer Science and mobile Computing*, *10*(1), 14-22. https://doi.org/10.47760/ijcsmc.2021.v10i01.003

Zhang, Y., Liu, J., & Shen, W. (2022). A review of ensemble learning algorithms used in remote sensing applications. *Applied Sciences, 12*(17)*,* 8654. https://doi.org/10.3390/app12178654

Zhou, Z., & Makse, H. A. (2019). *Artificial intelligence for elections: The case of 2019 Argentina primary and presidential election*. arXiv:1910.11227. https://www.semanticscholar.org/reader/39f79339bb8002aa7e77b10cbfa28d306175085a

Ziweritin, S. (2022). Stacked generalization of random forest and decision tree techniques for library data visualization. *International Journal of Engineering and Applied Computer Science*, *4*(4), 1-9. https://doi.org/10.24032/IJEACS/0404/005

AUTHORS

Godwin Elo (MPhil) is at the Department of ICT Education, Winneba, Ghana. His research interests are in the areas of Computer Networking, Machine Learning, ICT education, and so forth.

Prof. Benjamin Ghansah is a Computer Scientist, passionate about teaching and mentoring young people. He holds a BSc degree in Computer Science, an MSc in Computer Science, and a PhD in Computer Science. Currently, he is an Associate Professor in the Department of ICT Education, University of Education, Winneba. Prof. Ghansah has worked as an adjunct lecturer in some prestigious universities at home and abroad: Jiangsu University, Sikkim Manipal University, University of Ghana, Valley View University, Central University, Ghana Technological University College, and Blue Crest University College. He also serves as an assessor for the Computer Science department at the University of Malta, the University of Ghana, the University of Cape Coast, C.K. Tandem University, and the Computer Engineering Department at KNUST. Prof Ghansah has for the past 15 years been teaching and conducting research in federated search, information retrieval, information security, and machine learning. He is also an

assessor for the Ghana Tertiary Education Commission. He has over 40 peer-reviewed journal articles to his credit. Prof. Ghansah also serves as a reviewer for several journals.

Dr. Ephrem Kwaa-Aidoo holds a PhD in computing from the University of Bradford (UK). He is a Senior Lecturer in ICT at the Department of ICT Education and was the Director of ICT Services at the University of Education, Winneba. Ephrem has extensive work experience in the private and public sectors in various industries ranging from the Manufacturing, Health, Banking, and Education sectors in Ghana and the UK. His main area of interest is IT governance, including the use, security, and control of data and ICT tools to support the achievement of organizational goals. He has several publications to his credit. He has led the implementation of various Educational ICT projects, including ERP projects and Learning Management Systems. He has also worked and consulted on various projects training teachers in ICT integration and developing curricula and research for various organizations, including TTEL-DFID and UNESCO. He is an assessor for the Ghana Tertiary Education Commission.