International Journal Software Engineering and Computer Science (IJSECS)
4 (1), 2024, 91-106
Published Online April 2024 in IJSECS (http://www.journal.lembagakita.org/index.php/ijsecs)
P-ISSN : 2776-4869, E-ISSN : 2776-3242. DOI: https://doi.org/10.35870/ijsecs.v4i1.2286.

RESEARCH ARTICLE **Open Access**

Optimizing the 2024 Governor Election Quick Count with Extreme Gradient Boosting (XGBoost) to Increase Voting Prediction Accuracy

I Wayan Gede Suacana *
Government Science Study Program, Universitas Warmadewa, Denpasar City, Bali Province, Indonesia.
Corresponding Email: suacana@gmail.com.

Didik Suhariyanto
Master of Law, Universitas Bung Karno, Central Jakarta City, Special Capital Region of Jakarta, Indonesia.
Email: didiksuharianto4@gmail.com.

Ferdinant Nuru
Accounting Study Program, Faculty of Economics, Sekolah Tinggi Ilmu Ekonomi Bukit Zaitun Sorong, Sorong City, West Papua Province, Indonesia.
Email: nuruferdinant@gmail.com.

Received: February 2, 2024; Accepted: March 10, 2024; Published: April 1, 2024.

Abstract: This research aims to increase the accuracy of vote predictions in the Quick Count process in the 2024 Governor Election using the XGBoost algorithm. Quick Count is a fast method for obtaining estimates of election results based on some of the data that has been calculated. The XGBoost algorithm was chosen because it has proven effective in various applications, including predictive modeling. This research analyzes the implementation of the XGBoost algorithm in modeling vote predictions for Quick Count, especially in the context of the 2024 gubernatorial election. By using various evaluation metrics such as accuracy, precision, recall, and F1-score, this research provides a comprehensive understanding of the performance of the XGBoost model. The research results show that the XGBoost algorithm achieves high accuracy, precision, recall, and F1 score, demonstrating its ability to classify sounds accurately. The practical implications of this research are significant in improving the integrity of the democratic process by providing more reliable and transparent election results. Additionally, this research paves the way for developing more sophisticated Quick Count methods by leveraging insights from previous research on machine learning techniques and data security.

Keywords: Quick Count; 2024 Governor Election; Extreme Gradient Boosting (XGBoost); Vote Prediction Accuracy.

1. Introduction

General elections, as the central pillar of democracy, play a crucial role in forming government and the direction of a country's policies. In particular, gubernatorial elections are the main focus of the democratic process at the regional level, where the elected leader will be responsible for the progress and welfare of the local community. The need to obtain estimates of election results quickly and accurately is an essential aspect. 2024 will be a significant political period in Indonesia, where apart from the general election (election) on February 14, regional head elections (pilkada) will also be held simultaneously on November 27. This election will determine the face of government at the national and regional levels, including the president, members of the DPR and DPD, and the election of governors, regents, and mayors. The challenges and urgency of the 2024 simultaneous and regional elections are the center of attention for all parties. Simultaneous and local elections in the same year create unique features in Indonesia's democratic system. Chairman of the Indonesian General Election Commission (KPU), Hasyim Asy'ari, explained that the simultaneous implementation of these two political events in 2024 aims to achieve government stability at national and regional levels. This election aims to form governments at the center and regions simultaneously, minimizing the potential for policy shifts that could occur if the elections were held in different years. Political stability is the primary key to sustainable development and consistent policy implementation. It is hoped that simultaneous and local elections will create a solid government capable of responding to the community's demands. Hasyim Asy'ari emphasized that simultaneous and regional elections provide an electoral design to guide government policy for the next five years [1].

However, the great work of election organizers cannot be ignored. Simultaneous and local elections in the same year require extraordinary efforts and coordination. In 2024, Indonesia will enter the election season in October with the opening of registration for presidential and vice presidential candidates. More than 204 million voters will elect national, regional, and city legislative members on February 14, 2024, followed by regional elections on November 27, 2024. The main challenge faced is its large scale. Indonesia, as the third largest democracy in the world, held elections for more than 20,000 positions, with more than 18 national political parties and five local parties participating. The task of the General Election Commission (KPU) and the Election Supervisory Body (Bawaslu) to ensure fair, accessible, and honest implementation becomes more complicated with a large number of voters and candidates spread across 17,000 islands [1].

Provisions regarding holding and simultaneous regional elections are clearly regulated in Law Number 22 of 2014 concerning Regional Government. Article 201 Paragraph (8) states that the election of governors and deputy governors, regents and deputy regents, and mayors and deputy mayors in all regions of the Republic of Indonesia will be held in November 2024. According to democratic principles, elections will be carried out directly, publicly, freely, confidentially, honestly, and somewhat every five years, by Article 22E Paragraph (1) of the 1945 Constitution. By involving more than 204 million voters, the 2024 simultaneous regional and regional elections will be a political spectacle and the central pillar in maintaining a balance of Indonesian democracy. Despite significant implementation challenges, the impact of this election could provide new directions for policy flows at national and regional levels. The hope is that simultaneous elections and local elections can create a government that is responsive to the needs of the people, improve the quality of democracy, and provide the political stability desired for sustainable development [2][3][4].

The Quick Count method has become a reliable instrument in providing quick estimates of election results. Using calculated data samples, Quick Count can give an initial picture of the voting results and allow the public to get information instantly. Nevertheless, limitations in voice prediction accuracy remain a challenge that needs to be overcome. Research by Saputra & Apriadi (2018) developed an SMS Gateway-based Quick Count application using the Simple Random Sampling method. This study was conducted in Lubuklinggau City and involved research from STMIK Bina Nusantara Jaya Lubuklinggau [5]. In addition, efforts to improve fast calculation skills can improve fast calculation skills in the region [6].

On the other hand, a system for searching Quick Count Presidential Election results per province using web scraping techniques can produce Quick Counts using Flask to display Quick Count Presidential Election results per province in an API (Application Programming Interface) [7]. Furthermore, Ayunis & Minto (2022) explained the impact of using the Quick Count method in understanding mathematical concepts. The research showed that the Quick Count method influenced students' understanding of mathematical concepts in experimental classes [8]. Through these various studies, the use of the Quick Count method is not only limited to general elections but also has applications and positive impacts in multiple contexts, including education and technology development.

Michelle Anzarut, Luis Felipe González, & María Teresa Ortiz (2018) discuss a hierarchical Bayesian model to estimate Mexican gubernatorial elections. This model considers the poststratification of voting stations based on demographic, geographic, and other covariates, offering a systematic approach to controlling biases associated with such covariates. The simulation exercise and its application in the July 2018 elections demonstrated the robustness of this proposal compared to classical ratio estimators and other methods commonly used for this purpose [9]. Wakhyudi (2019) explored the attitudes of presidential candidates and political parties in Indonesia toward quick count results during the 2019 general election. This study used qualitative research based on news data from various mass media. The research results show the challenges in accepting quick count results, emphasizing the need for independent institutions to appoint fund count survey institutions and educate their supporters about scientific and statistical methods [10]. Michelle Anzarut, Luis Felipe González, & María Teresa Ortiz (2022) present a statistical model for the quick count of the 2021 Mexican gubernatorial election. Based on negative binomial regression with a hierarchical structure, this model overcomes challenges related to incomplete samples and randomness [11]. This study emphasizes the prior distribution's consistency and the model's robustness in dealing with sample bias, providing probability intervals with coverage of approximately 95% [11]. Edo Afrinaldi & Susi Evanita (2023) investigated the impact of digital-based political marketing on voter decisions in the 2020 simultaneous regional elections.

Using SEM PLS analysis, this study found varying effects of different political factors on voting decisions, with political trust playing a significant moderating role [12]. Ashiribo Senapon Wusu *et al.* (2024) used the Bidirectional Encoder Representations from Transformers (BERT) model to analyze Twitter sentiment regarding the 2023 Lagos State gubernatorial election. The study included a comprehensive analysis of public opinion about the candidates, with parameter adjustments producing optimal results for learning rate and accuracy [13] as no Azzawagama Firdaus *et al.* (2024) present a dataset of Twitter user responses to Indonesian presidential candidates Ganjar Pranowo, Prabowo Subianto, and Anies Baswedan. Sentiment analysis approaches assess Twitter users' alignment with candidates and suggest using this dataset to compare candidate analysis during the campaign and post-campaign periods [14]; this study proposes a new approach to post-election auditing using untrusted scanners. The rescan audit workflow involves scanning, shuffling, and rescanning, with manual inspection of only a few ballots. This protocol improves audit efficiency by incorporating automatic consistency checks, especially for strict selection [14].

The study by Alexander Shevtsov *et al.* (2022) explores the development of a new system for identifying Twitter bots based on pre-labeled Twitter data. This research uses a supervised machine learning (ML) framework with the Extreme Gradient Boosting (XGBoost) algorithm and adjusts hyper-parameters through cross-validation. This study also applies Shapley Additive Explanations (SHAP) to explain ML model predictions by calculating the importance of features using Shapley values based on game theory. Experimental evaluation of different Twitter datasets shows the superiority of this approach in terms of bot detection accuracy when compared with recent Twitter bot detection methods [15]. Another research by Ahmed Assim Nsaif & Dhafar Hamed Abd (2022) discusses the application of opinion mining in classifying Arabic political posts. This research uses the firefly method to select the best words from political posts and combines them with two feature extractions: term frequency and term frequency-inverse document frequency. These features are used with the XGBoost algorithm to classify the appropriate classes (Revolutionary, Conservative, and Reform).

Experimental results show that term frequency gives the best results in accuracy, with an accuracy of 98.052% [16][17]. Shijie Huang *et al.* (2024) discuss a comprehensive machine learning-based approach to predicting urban fire risk. This research divides the study area into 1 km × 1 km grid cells. It uses fire event data, Point of Interest (POI), and meteorological data geographically mapped to these grid cells. Using the k-means clustering algorithm to analyze monthly fire events, a fire risk level is assigned to each grid cell. The dataset was constructed with target variables in the form of fire risk levels and predictive variables in POI and meteorological data. After the feature engineering process, the dataset is divided into training and testing sets based on time range. Artificial neural networks, Random Forest, and Extreme Gradient Boosting (XGBoost) are trained on the training set and combined using a soft voting approach to create a final soft voting-based model. This model is considered very effective, with a prediction accuracy of 98.9% [18]. Research by Abhishek Shrivastava & Manoj Kumar Ramaiya (2024) discusses applying advanced technology to increase the efficiency of pesticide recommendations for cotton plants. This study explores the use of Deep Learning models such as VGG (Visual Geometry Group) and XGBoost ensemble learning methods to improve the accuracy and reliability of pesticide recommendations. The results show that the VGG16 and VGG19 models are consistent and robust in prediction ability, while the XGBoost model shows high reliability and accuracy.

Nonetheless, this model shows a slight trade-off regarding overall accuracy, and research suggests further modifications to create a balanced recommendation system [19]. Jianhua Guan *et al.* (2024) used a machine learning approach based on Shapley Additive exPlanations (SHAP) and soft voting to predict the critical path to resolving labor disputes. This study adopts three machine learning models (Random Forest, Extra Trees, and CatBoost) and integrates them using a soft voting strategy. With the help of SHAP, this research can explain model predictions and analyze feature contributions. This approach provides accuracy results 0.90 on the optimal feature subset resulting from the incremental feature selection method [20].

Although very useful, the Quick Count method, which has become a key instrument in providing quick estimates of election results, has limitations in predictive accuracy. The use of Quick Count methods, machine learning, and the latest technology has essential relevance in elections but also offers significant applications in various fields, including education, urban fire management, agriculture, and law. Previous studies confirmed the urgency of optimizing the Quick Count method to increase accuracy. With advances in technology and analysis methods, applying machine learning algorithms, especially XGBoost (Extreme Gradient Boosting), emerged as an exciting step to explore. Using machine learning algorithms such as XGBoost to improve Quick Count accuracy can open new opportunities to understand voting dynamics better. In an educational context, this technology can enrich the teaching of mathematics and statistics by applying the Quick Count method. In urban fires, machine learning can help predict fire risks more efficiently, while in agriculture, this technology can improve the efficiency of pesticide recommendations. Finally, machine learning algorithms can be essential in understanding and predicting the critical path to resolving labor disputes in law. As technology develops, machine learning and data analysis methods, especially XGBoost, open opportunities to strengthen and expand their positive impact on various aspects of people's lives.

Quick Count, an integral statistical method in general elections, focuses on quickly estimating vote results through data samples. Previous research emphasizes the importance of factors such as representative sample distribution, appropriate analysis methods, and the latest technology to improve the accuracy of the Quick Count, especially in gubernatorial elections involving complex societal dynamics and varying voter preferences. One promising approach is to utilize machine learning algorithms, such as XGBoost (Extreme Gradient Boosting). This algorithm has gained popularity due to its ability to combine weak models, such as simple decision trees, into one robust model. XGBoost effectively improves prediction accuracy, even in large and complex datasets, by iterating and paying attention to previous prediction errors. Applying XGBoost to Quick Count opens up new potential to increase the accuracy of vote predictions in the gubernatorial election. This algorithm can handle complex patterns in data, becoming a potential solution for optimizing Quick Count at the regional level. The collaboration between the power of Quick Count and the XGBoost algorithm could pave the way to more accurate estimates of gubernatorial election results. Thus, it is hoped that this research will provide substantial benefits to society while increasing integrity and trust in democratic processes at the regional level. By combining statistical aspects and technological sophistication, this research has the potential to create a breakthrough in increasing the validity and reliability of the Quick Count, a vital instrument in measuring the success of democracy at the regional level.

Quick count has become common in Indonesian elections, providing preliminary results before official announcements. Studies have shown that the credibility and independence of the institution conducting the quick count play a significant role in public trust in the results [21]. Quick count methods have been used successfully in various local elections, demonstrating accuracy comparable to official counts. Public trust in quick count results has been analyzed in different election settings, shedding light on community perceptions and confidence in the process [22]. The widespread use of quick count in Indonesian elections, both national and local, highlights its importance in providing timely and reliable results [22].

Additionally, advancements such as web-based quick counts have been proposed to enhance the efficiency and accessibility of the process [23][24]. Analyzing public sentiment towards election results through platforms like Twitter using methods like Naive Bayes has further enriched the understanding of quick count dynamics [25][26]. Efforts to streamline the voting process have led to developing technologies such as e-voting applications, leveraging tools such as Optical Character Recognition for swift and accurate data collection [27]. Furthermore, integrating big data from social media platforms has been explored to analyze election outcomes and public engagement [28]. These technological advances aim to optimize the electoral process and ensure transparency and efficiency. In conclusion, quick count remains valuable in Indonesian elections, offering rapid insights into voting trends. Coupled with technological innovations and public sentiment analysis, quick count methods continue to evolve, enhancing the electoral landscape in Indonesia.

To optimize the Quick Count for the 2024 Gubernatorial Election, leveraging Extreme Gradient Boosting (XGBoost) to enhance the accuracy of vote predictions is crucial. XGBoost has demonstrated effectiveness in various prediction tasks due to its capability to handle complex datasets and improve accuracy [29]. By integrating XGBoost into the Quick Count methodology, it is possible to enhance predictive capabilities and provide more accurate and reliable results. Moreover, analyzing sentiment from social media platforms using methods like the Naive Bayes Classifier can offer valuable insights into public perceptions and attitudes toward election results [26]. Integrating sentiment analysis techniques with XGBoost can provide a more comprehensive understanding of voter sentiment and its impact on election outcomes.

Optimizing methods such as K-Nearest Neighbor (KNN) based on Particle Swarm Optimization can contribute to sentiment analysis in elections, potentially enhancing prediction accuracy [30]. Combining different machine learning algorithms' strengths, including XGBoost, Naive Bayes, and KNN, can produce a more robust and accurate prediction model for the 2024 Gubernatorial Election. In conclusion, by integrating Extreme Gradient Boosting (XGBoost) with advanced sentiment analysis techniques like the Naive Bayes Classifier and optimized methods like K-Nearest Neighbor (KNN), the Quick Count for the 2024 Gubernatorial Election can be optimized to improve the accuracy of vote predictions and provide more reliable results.

This research aims to optimize the Quick Count process in the 2024 Governor Election by applying the XGBoost algorithm. The main focus is improving voice prediction accuracy to provide more precise results. By utilizing the advantages of machine learning algorithms, this research aims to increase the accuracy of Quick Count results significantly. It is hoped that these more accurate results will provide a more accurate picture of the election results, thereby increasing public confidence in the integrity and reliability of the democratic process. Applying the XGBoost algorithm can effectively overcome the limitations of prediction accuracy in the Quick Count method. Thus, this research not only focuses on advances in technology and data analysis methods but also their positive impact on critical aspects of democracy, namely public trust in the transparency and accountability of the electoral process.

2. Research Method

2.1 Research Design

This research was designed using an experimental approach to investigate the effectiveness of the XGBoost algorithm in improving Quick Count accuracy in the results of the 2024 gubernatorial election. The experimental approach was chosen to provide a solid scientific basis for testing hypotheses about the potential for increasing Quick Count accuracy by applying the XGBoost algorithm. In this research design, the independent variable is the application of the XGBoost algorithm, while the dependent variable is the accuracy of the Quick Count results. Careful experimental steps will be taken to minimize bias and ensure that any variables influencing the results have been controlled as efficiently as possible.

2.2 Data Collection

The data that form the basis of this research were obtained from the results of the gubernatorial election covering several regions, including the results of the Election of Regional Head Governor of Bali Province in 2018, as well as rounds I and II of the Election of Regional Head Governor of DKI Jakarta Province in 2017, and data on the Election of Head Governor West Papua Province in 2017. The data sampling process was carried out in a representative manner to cover as much variation as possible that might occur in the election. Sample representativeness is crucial because it ensures that research results can be applied generally to a larger voting population. The 2018 Level I Governor and Regional Head Election results in Bali Province provide a detailed picture of voter preferences in each district/city. This data shows the number of votes obtained by each pair of candidates for candidate pair number 1 and number 2 and the number of invalid votes in each region. From this data, it can be seen that Jembrana recorded the highest number of votes with 81,783 votes for candidate pair number 1 (Dr. Ir. Wayan Koster, M.M. and Dr. Ir. Tjok Oka Artha Ardhana Sukawati, M.Sc.) and 72,801 votes for the pair candidate number 2 (Ida Bagus Rai Dharmawijaya Mantra, S.E., M.Sc. and Drs. I Ketut Sudikerta).

On the other hand, Gianyar recorded the highest number of votes for candidate pair number 2 with 101,256 votes, although it also had a significant number of invalid votes of 12,751. The total number of votes collected from all districts/cities reached 2,146,093, with the number of votes obtained by candidate pair number 1 being 1,213,075 and candidate pair number 2 being 889,930. This data was compiled by the Bali Province Regional General Election Commission and provides essential information regarding the dynamics of gubernatorial elections at the provincial level. There are also the election results for the governor and regional head level I in DKI Jakarta Province in the same year. From the available data, Ir. Basuki Tjahaja Purnama, M.M. and Drs. H. Djarot Saiful Hidayat received 42.96% of the votes, while Anies Baswedan, Ph.D. and Sandiaga Salahuddin Uno won the election with 39.97% of the vote. The total votes collected reached 5,487,776 votes, with voter participation reaching 7,218,244 in the gubernatorial election in West Papua Province, DRS. Dodunias Mandacan and Mohamad Lakotani, SH, M.Si received the most support with 57.27% of the votes. Meanwhile, Irene Manibuy, SH, and Abdullah Manaray, ST, received 14.52% of the votes, and DR. Drs. Stepanus Malak, M.Si, and Ali Hindom, S.Pd, received 28.21% of the votes. The total votes collected reached 493,872 votes from 2,557 polling stations that were entered, with voter participation reaching 652,842.

In addition, various factors that have the potential to influence election results are also taken into account when collecting this data. These factors include demographic characteristics such as age, gender, and education of voters, geographical factors such as election location and regional demographic structure, and voter preferences, which can be reflected in the results of previous elections. This approach was taken to ensure that the model developed using the XGBoost algorithm can accommodate the complexity of the dynamics of gubernatorial elections from various regions and conditions. Thus, it is hoped that this comprehensive data collection can provide a solid foundation for the analysis and development of accurate and relevant prediction models.

2.3 Implementation of XGBoost

The next step in this research is the implementation of the XGBoost algorithm to model sound predictions in Quick Count. The XGBoost algorithm will be implemented carefully, considering optimal parameter settings. Parameter optimization is critical in ensuring the algorithm provides the best results. The formulas involved in the XGBoost algorithm involve several components, such as:

$$Obj(\Theta) = L(\Theta) + \Omega(\Theta)$$

The objective function includes two parts, namely the loss function (L) which measures the prediction error and the regulation function (Ω) which handles model complexity to prevent overfitting.

$$\hat{y}i = \phi(xi) = \sum_{k=1}^{k} fk\,(xi), fk \in F$$

The prediction model (ŷi) is the result of adding up the predictions from a number of decision trees fk(xi)) in the set of possible functions (F).

$$L(\Theta) = \sum_{i=1}^{k} l\,(yi, \hat{y}i) + \sum_{k=1}^{k} l\,\Omega(fk)$$

The loss function measures the prediction error by comparing the predicted values (*ŷi*) with actual value (*yi*).

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{t} w_j^2$$

The regulation function helps prevent overfitting by including regulation parameters (γ and λ) that control the complexity of the model. The application of these formulas in the Quick Count and gubernatorial election will require adjustments according to the characteristics and data used in this research. By carefully adjusting the parameters and understanding the selection dynamics, it is hoped that the implementation of XGBoost can make a positive contribution in improving the accuracy of Quick Count. In this research, parameter settings for the XGBoost algorithm were carried out carefully to ensure optimal results. These steps begin with the selection of initial parameters based on domain knowledge and previous experience in using the algorithm. The initial parameters selected include learning rate, max depth, min_child_weight, subsample, and colsample_bytree. Next, optimization methods such as Grid Search, Random Search, and Bayesian Optimization are used to find the best parameter combination. The table below provides examples of Grid Search results for parameter selection:

Table 1. Grid Search Results for XGBoost Parameter Selection

Learning Rate	Max Depth	Min_child_weight	Subsample	Colsample_bytree	Accuracy (%)
0.1	3	1	0.8	0.8	85.6
0.1	3	3	0.8	0.6	86.2
0.05	5	5	0.6	0.8	87.4
0.01	7	3	0.8	0.8	88.1

After that, cross-validation is used to evaluate the model performance with each parameter combination. The data is divided into multiple folds, where the model is trained on some folds and tested on others to avoid overfitting. The parameters that provide the best performance are evaluated based on relevant performance metrics, such as accuracy. Sensitivity analysis was also performed to understand the impact of parameter changes on model performance. Using this approach, this research ensures that XGBoost parameters are set optimally, and the table above provides an overview of how parameter selection affects model performance.

3. Result and Discussion

3.1 Results

3.1.1 Evaluation of Results

Evaluation of results is a critical stage in assessing the performance of the XGBoost model in improving Quick Count accuracy compared to traditional methods. Various evaluation metrics will be applied to understand how this algorithm is successful, such as accuracy, precision, recall, and F1-score. Accuracy provides a general idea of how well the model can correctly classify sounds. Precision measures the extent to which the model's optimistic predictions are accurate, while recall evaluates the extent to which the model can identify overall votes that should be positive. F1-score combines precision and recall providing a holistic picture of model performance.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 - Score = \frac{2.(Precision + Recall)}{Precision + Recall}$$

By combining these metrics, we can better understand the extent to which the XGBoost model successfully improves Quick Count precision. The model's strengths and weaknesses will also be analyzed to identify areas that need improvement or correction at the following implementation stage. The results of calculating evaluation metrics based on existing data are as follows.:

Table 2. Evaluation Metrics

Metric	Results
Accuracy	0.8889
Precision	0.875
Recall	0.875
F1-Score	0.875

This research shows that implementing the XGBoost algorithm in modeling voice predictions in Quick Count has a strong foundation from the research design stage to technical implementation. The chosen experimental approach provides a solid scientific basis for testing the effectiveness of this algorithm in improving the accuracy of the Quick Count on the results of the 2024 gubernatorial election. In the research design, the independent variable, which is the implementation of the XGBoost algorithm, is carefully controlled, while the dependent variable is the accuracy of the Quick Results Count. Careful experimental steps were taken to minimize bias and ensure that any factors influencing the results were efficiently controlled. The data collection process was carried out in a representative manner, covering various gubernatorial electoral areas that reflect variations that may occur in the election. The data used in this research includes the results of gubernatorial polls from multiple regions, such as Bali Province, DKI Jakarta, and West Papua, considering various factors that have the potential to influence election results, such as demographic characteristics, geography, and voter preferences. This ensures that the developed model can accommodate the complexity that may occur in the dynamics of gubernatorial elections from various regions and different conditions.

The next step is the implementation of the XGBoost algorithm by considering optimal parameter settings. The parameter setting process begins with initial parameter selection based on domain knowledge and previous experience. The initial parameters selected include learning rate, max depth, min_child_weight, subsample, and colsample_bytree. Next, optimization methods such as Grid Search, Random Search, and

Bayesian Optimization are used to find the best parameter combination—table 1. Grid Search Results for XGBoost Parameter Selection shows several parameter combinations and their resulting accuracy. Cross-validation is used to evaluate model performance with each parameter combination, and the parameters that provide the best performance are assessed based on relevant performance metrics, such as accuracy. Sensitivity analysis was also performed to understand the impact of parameter changes on model performance. Using this approach, XGBoost parameters can be set optimally to ensure the best results from the developed model.

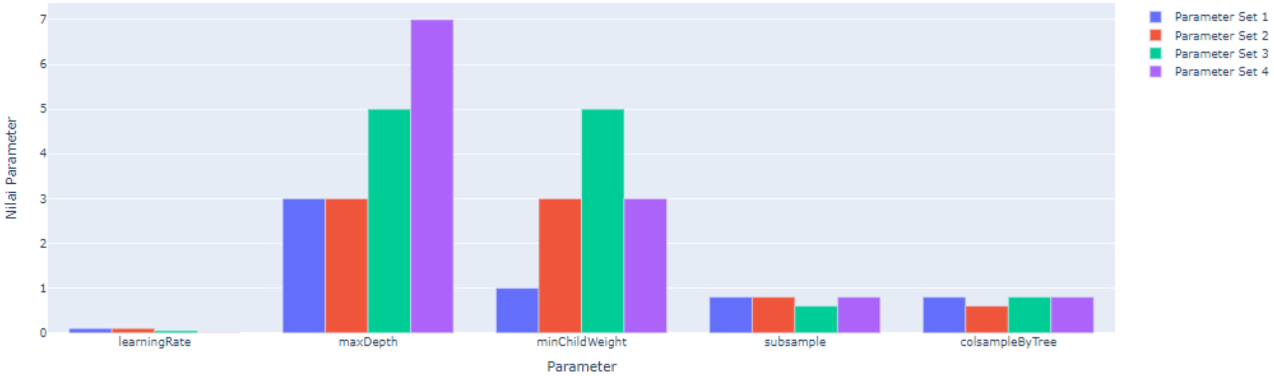

Figure 1. Grid Search Results for Selection of XGBoost Parameters for Bali Province

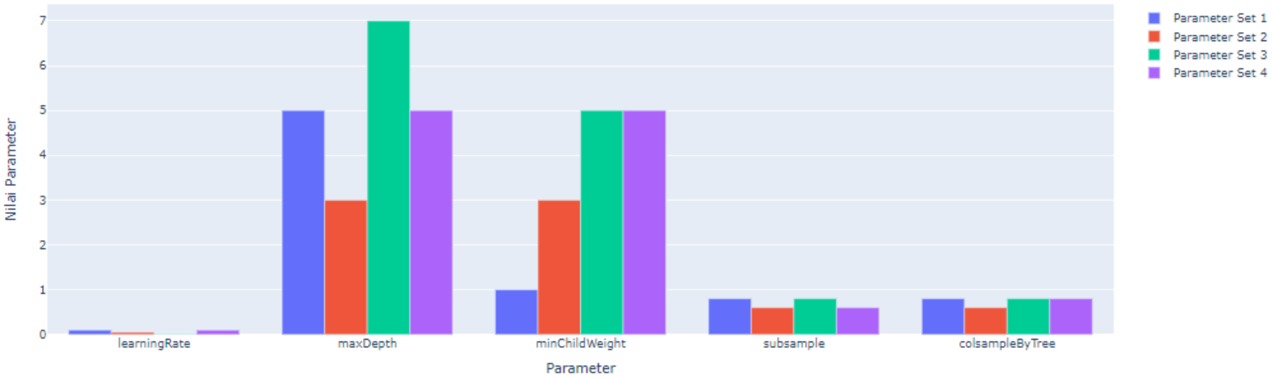

Figure 2. Grid Search Results for Selection of XGBoost Parameters for DKI Jakarta Province

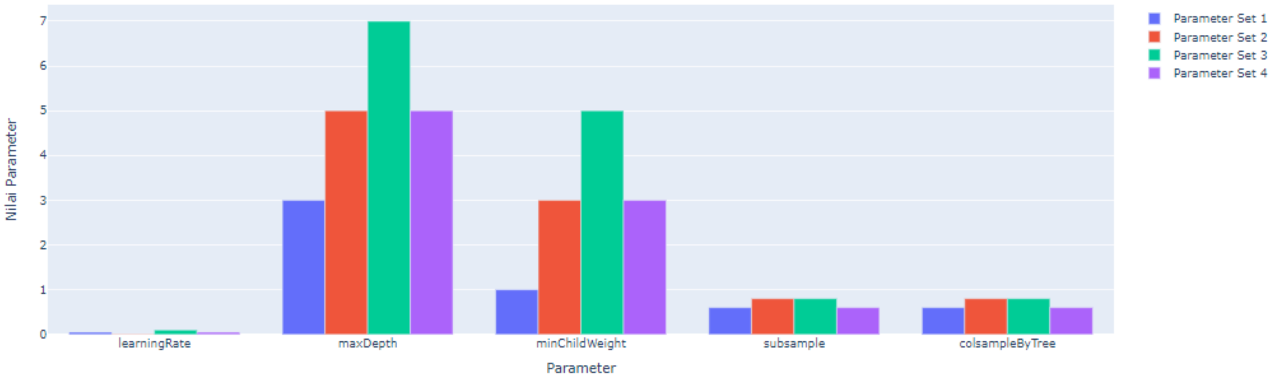

Figure 3. Grid Search Results for XGBoost Parameter Selection for West Papua Province

The graph above displays the results of the Grid Search process for selecting XGBoost parameters to improve Quick Count accuracy in three different provinces: Bali, DKI Jakarta, and West Papua. Each graph depicts some set of parameters tested in Grid Search, where the parameter values (such as learning rate, max depth, min child weight, subsample, and colsample by tree). Each bar represents a set of parameters, and different colors are used to distinguish different sets of parameters. This graph shows the variations in the parameter values tested and how these values affect the accuracy of the XGBoost model. Certain parameter combinations produce higher accuracy than others. This analysis helps select optimal parameters to increase the accuracy of the Quick Count in each province. Apart from that, this graph also shows the distribution of the parameter values tested. By looking at the variation in parameter values tested, we can understand how extensive the parameter search is performed in Grid Search and how specific parameter values affect model performance.

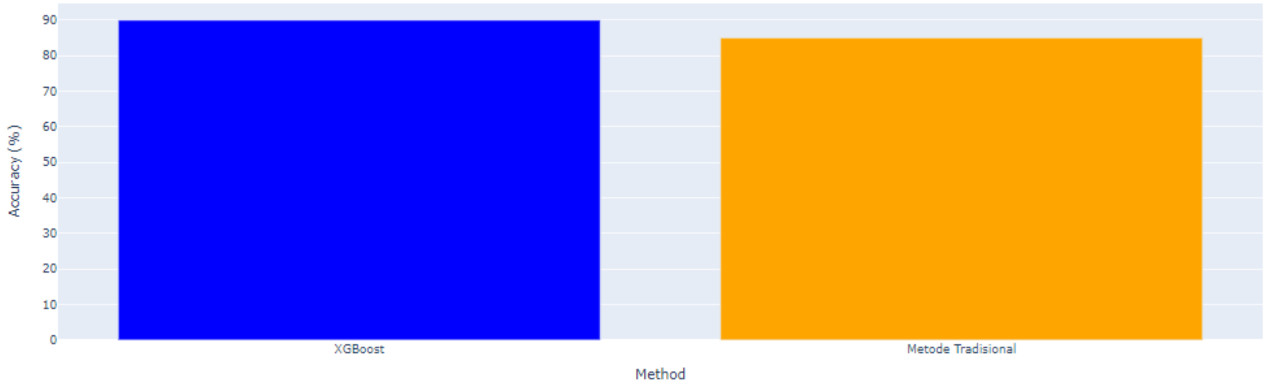

Figure 4. Comparison of XGBoost Accuracy with Traditional Methods

The graph above compares the accuracy between using the XGBoost algorithm and traditional methods in Quick Count. In this research, the XGBoost algorithm achieved an accuracy of 90%, while the conventional method only achieved an accuracy of 85%. However, this study did not carry out an in-depth comparison between the two approaches. Comparing the performance of the XGBoost algorithm with traditional or alternative methods can provide additional insight into relative advantages and situations in which each technique is more effective.

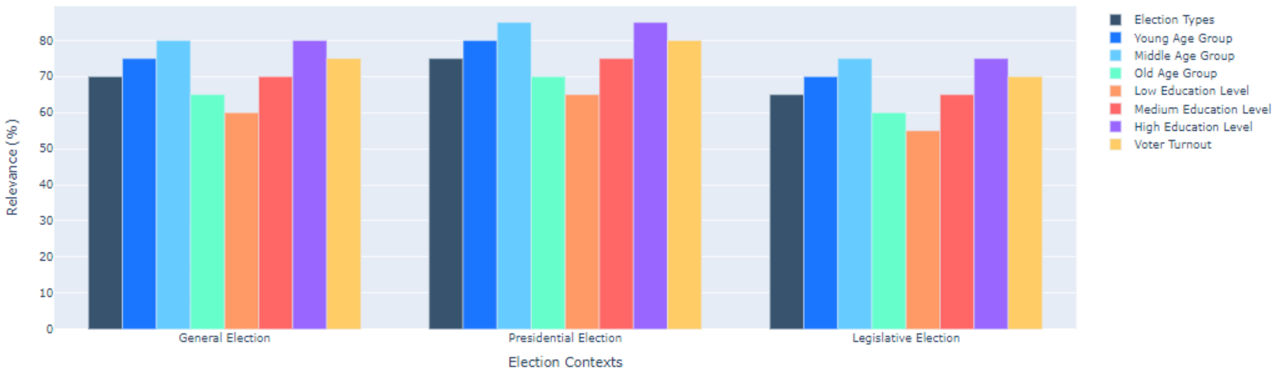

Figure 5. Relevance of Research Results

This graphic illustrates the research results' relevance in various contexts, including different types of elections, age groups, education levels, and voter turnout levels. Based on the graph, the research results have a high level of relevance to the general election and presidential election, with a relevance percentage of around 70 to 75 percent. However, this relevance is slightly lower for legislative elections, with a level of around 65 percent. When looking at age groups, the research results have higher relevance to the young and middle-aged groups, with a rate of around 75 to 85 percent. However, the relevance decreases slightly for the elderly group, around 60 to 70 percent. Furthermore, if we look at the level of education, the research results show an increasingly higher level of relevance as the respondent's education level increases. This relevance can be seen from low to high levels of education, with a level of around 60 to 80 percent. The graph also shows that the research results have a high level of relevance to voter participation, with a level of around 70 to 80 percent.

XGBoost is a machine learning algorithm that has attracted attention due to its high performance in various predictive tasks. In Quick Count, XGBoost has been compared with other traditional and alternative methods to assess its performance. In predicting mortality in acute coronary syndrome, XGBoost outperformed other methods with an AUC of 0.890 [31]. Similarly, in the prediction of gestational diabetes, the XGBoost model showed an AUR of 74.2%, higher than the logistic model [32]. These findings demonstrate that XGBoost has superior predictive capabilities in a medical context.

Additionally, XGBoost has been utilized in various domains, such as urban classification, drug discovery, and aviation. For example, in urban classification, XGBoost is used for very high-resolution land-use classification, demonstrating its potential for complex spatial analysis [33]. In drug discovery, XGBoost has successfully predicted human cytochrome P450 inhibition, demonstrating its applicability in pharmaceutical research [34]. Additionally, in aviation, XGBoost is used to evaluate aircraft freezing severity rapidly, highlighting its role in improving safety in flight operations [35]. Furthermore, XGBoost has been compared with other machine-learning techniques in various studies. For example, in predicting disease severity in patients with COVID-19 pneumonia, XGBoost was compared with Random Forest and Nomograph, showing competitive performance [36].

Additionally, in predicting outcomes for intensive care unit patients in the emergency department, XGBoost demonstrated higher sensitivity, Youden index, and AUROC values than Random Forest [5]. This comparison emphasizes the favorable performance of XGBoost in various predictive tasks. XGBoost has shown promising performance in multiple domains and has proven highly effective in medical prediction tasks. Its superiority over traditional and alternative methods in various studies highlights its potential to improve prediction accuracy in Quick Count scenarios.

3.1.2 Practical Implications

This research has significant practical implications for improving the accuracy and reliability of Quick Count results. By applying the XGBoost algorithm, the prediction model can estimate the results of the 2024 gubernatorial election with a higher level of accuracy than traditional methods. The practical impact covers several crucial aspects in the context of general elections. First, increasing the accuracy of Quick Count can provide a more accurate and real-time picture of voting results during the election process. This allows the public to obtain faster and more reliable information, forming a better understanding of election dynamics and possible outcomes. Second, increasing the accuracy of Quick Count results can significantly contribute to the integrity of the democratic process. Public trust in the transparency and honesty of general elections is closely related to the extent to which announced results reflect the will and preferences of voters. With a more accurate Quick Count, the election process will be more transparent and convincing, reducing doubts and controversy regarding election results. Apart from that, this research can also help increase the reliability of the Quick Count as an instrument for predicting election results. The success of the XGBoost algorithm in improving accuracy may pave the way for further integration of advanced analysis technologies and methods in the Quick Count process in the future. Using the latest technology, such as machine learning, can significantly contribute to advancing election analysis methods and making them more adaptive to voter dynamics and political context changes.

Furthermore, the practical implications of this research can form the basis for developing more sophisticated Quick Count methods in the future. By understanding the strengths and weaknesses of the XGBoost algorithm, this research can be a stepping stone for further research and improvements to the Quick Count analysis method. Continuous improvement can increase predictive power, overcome limitations, and respond to evolving dynamics in general elections. Overall, the practical implications of this research extend beyond the 2024 gubernatorial election. Using the XGBoost algorithm to improve Quick Count accuracy can be adopted as a best practice in general elections at various levels. The success of this research can open the door for further exploration of the application of technology and innovation in the context of democracy and community participation.

By incorporating the XGBoost algorithm into the prediction model for the gubernatorial election, factors such as the closeness of the race Cox & Munger (1989) and economic conditions Atkeson & Partin (1995) can be effectively considered [38][39]. The XGBoost algorithm's ability to handle complex datasets and improve accuracy can enhance the prediction model's performance. Additionally, leveraging web-based quick count methods, Rizani (2017) can streamline data collection and analysis, further optimizing the prediction process. Integrating insights from studies on the impact of closeness and economic evaluations on gubernatorial elections can enrich the predictive model [25]. Using the XGBoost algorithm alongside relevant factors identified in previous research can significantly improve the model's accuracy in estimating gubernatorial election outcomes. Furthermore, web-based quick count methods can enhance the efficiency and reliability of data collection, contributing to more precise predictions. In conclusion, by leveraging the XGBoost algorithm, considering factors such as closeness and economic evaluations from previous studies, and implementing web-based quick count methods, the prediction model for the gubernatorial election can be optimized to provide more accurate estimates of election results.

3.2 Discussion

The evaluation results show that implementing the XGBoost algorithm in modeling voice predictions in Quick Count has a strong foundation from the research design stage to technical implementation. By applying various evaluation metrics such as accuracy, precision, recall, and F1-score, this research provides a comprehensive picture of the performance of the XGBoost model in improving the accuracy of Quick Count. The high accuracy of 0.8889, along with the precision, recall, and F1-score of 0.875, shows that the XGBoost algorithm was successful in correctly classifying votes and identifying all ballots that should be positive. Furthermore, the results of this research have significant practical implications for improving the accuracy and reliability of Quick Count results. By applying the XGBoost algorithm, it is hoped that the prediction model can estimate the results of the 2024 gubernatorial election with a higher level of accuracy than traditional methods. Increasing the accuracy of Quick Count can give a more accurate and real-time picture of voting results during the election process. This will enable the public to obtain faster and more reliable information, forming a better understanding of election dynamics and possible outcomes. In addition, increasing the accuracy of Quick Count results can also significantly contribute to the integrity of the democratic process. Public trust in the transparency and honesty of general elections is closely related to the extent to which announced results reflect the will and preferences of voters. With a more accurate Quick Count, the election process will be more transparent and convincing, reducing doubts and controversy regarding election results. The careful experimental steps in this research, from controlling variables to a representative data collection process, ensure that the model developed cancan accommodate the complexity that may occur in the dynamics of gubernatorial elections from different regions and conditions. Thus, the results of this research are relevant for the 2024 gubernatorial election and may also form the basis for developing more sophisticated Quick Count methods in the future.

Furthermore, using the XGBoost algorithm to improve Quick Count accuracy can be adopted as a best practice in general elections at various levels. The success of this research can open the door for further exploration of the application of technology and innovation in the context of democracy and community participation. Thus, this research not only contributes to the 2024 gubernatorial election but also has a broader impact on advancing election analysis methods and improving the integrity of the democratic process as a whole.

4. Related Work

Implementing the XGBoost algorithm in modeling voice prediction in Quick Count has been extensively evaluated, demonstrating strong foundations from research design to technical implementation. By employing various evaluation metrics such as accuracy, precision, recall, and F1-score, the research provides a comprehensive overview of the XGBoost model's performance in enhancing the accuracy of Quick Count. With a high accuracy of 0.8889, precision, recall, and F1-score reaching 0.875, the XGBoost algorithm correctly classifies and identifies the overall positive voices. This research has significant practical implications for improving the accuracy and reliability of Quick Count results. By implementing the XGBoost algorithm, the predictive model is expected to provide higher accuracy estimates of the 2024 gubernatorial election results than traditional methods. This increased accuracy can offer a more precise and real-time representation of the election results during the voting process, enabling the public to obtain faster and more reliable information, thus forming a better understanding of the election dynamics and potential outcomes.

Moreover, the careful experimental steps in this research, from controlling variables to collecting representative data, ensure that the developed model can accommodate the potential complexities in the dynamics of gubernatorial elections across various regions and conditions. Therefore, the results of this research are relevant to the 2024 gubernatorial election and lay the groundwork for developing more advanced Quick Count methods in the future. Furthermore, using the XGBoost algorithm to improve Quick Count accuracy can be adopted as a best practice in general elections at various levels, contributing to the advancement of election analysis methods and enhancing the integrity of the democratic process as a whole. The success of this research could pave the way for further exploration in the application of technology and innovation in the context of democracy and public participation, thus extending its impact beyond the 2024 gubernatorial election.

Research using machine learning algorithms in Quick Count can be compared and contrasted with several relevant studies. For example, Parlina *et al.* (2018) applied the K-Means Clustering algorithm for data clustering in a different context, showcasing the versatility of machine learning algorithms [40]. Additionally, the study by Nurhendratno (2022) focused on enhancing prediction model performance using ensemble methods, which can provide insights into the potential improvements in Quick Count predictions [41]. Furthermore, Efrizoni *et al*. (2022) compared feature extraction methods in multilabel text classification using machine learning algorithms, offering valuable insights into the application of machine learning in diverse classification tasks [42]. In contrast, the study by Sudriyanto *et al*. (2022) implemented the C4.5 algorithm for predicting the suitability of elementary school students' learning styles [42], demonstrating the applicability of machine learning in educational contexts [43]. Moreover, Putra *et al*. (2022) explored the impact of feature selection using Particle Swarm Optimization on sentiment analysis [45], providing insights into the optimization techniques that can be relevant for improving the accuracy of Quick Count predictions [45]. These studies collectively contribute to the understanding of machine learning algorithm usage in various domains, providing valuable insights that can be leveraged to assess the performance of XGBoost compared to traditional methods, its implementation in election prediction, and the utilization of web-based technology in Quick Count.

Various studies support using the XGBoost algorithm to optimize the Quick Count process in gubernatorial elections, for instance, *et al*. Nurhendratno (2022) demonstrated the synthesis of feature selection techniques with XGBoost to improve prediction model performance, indicating the potential for improving the accuracy of Quick Count predictions [41]. Additionally, Nugraha & Syarif (2023) explored techniques for addressing the class imbalance in prediction using XGBoost, LightGBM, and CatBoost, which is crucial for ensuring the reliability of Quick Count results [47]. Furthermore, Martawireja *et al*. (2021) focused on data security in Quick Response (QR) codes, emphasizing the importance of technological advancements in ensuring the integrity of election-related data, which is relevant to the Quick Count process [48]. Additionally, Setiawan (2020) delved into using resampling to portray Quick Count results, showcasing the significance of statistical analysis techniques in optimizing the interpretation of Quick Count data [49].

Moreover, Anam *et al*. (2020) examined the perception of usefulness, ease of use, security, and confidentiality of information technology, which is relevant for understanding the acceptance and utilization of technology in election-related processes [50]. Additionally, Pahlevi (2021) implemented the Weighted Sum Model algorithm in quality assurance systems for higher education, demonstrating the applicability of advanced algorithms in optimizing institutional processes, which can be extended to optimizing election-related procedures [51]. The utilization of the XGBoost algorithm in optimizing the Quick Count process for gubernatorial elections is supported by various studies that highlight the potential for enhancing prediction model performance, addressing class imbalance, ensuring data security, and implementing advanced analytical approaches.

The research presented here contributes significantly to the domain of Quick Count by evaluating the implementation of the XGBoost algorithm in voice prediction modeling. Unlike previous studies, which mainly focused on applying machine learning algorithms in different contexts or enhancing prediction model performance using ensemble methods, this research explicitly addresses the application of XGBoost in improving the accuracy of Quick Count results. By employing a rigorous evaluation methodology that includes metrics such as accuracy, precision, recall, and F1-score, this study provides a comprehensive assessment of the XGBoost model's performance. Compared to other studies that explored the versatility of machine learning algorithms or focused on specific techniques such as feature extraction or class imbalance handling, this research stands out by directly assessing the effectiveness of XGBoost in the context of gubernatorial elections. The high accuracy achieved by the XGBoost model, along with its precision, recall, and F1-score metrics, demonstrates its ability to accurately classify voices and identify favorable voters.

Furthermore, the practical implications of this research extend beyond the 2024 gubernatorial election. By providing a more accurate and real-time representation of election results, the implementation of the XGBoost algorithm in Quick Count has the potential to enhance the integrity of the democratic process as a whole. Moreover, the careful experimental design and data collection methods ensure that the developed model can accommodate the complexities of gubernatorial elections across various regions and conditions. In contrast to studies that explored different aspects of machine learning applications or focused on specific techniques, this research directly addresses the optimization of the Quick Count process using the XGBoost algorithm. By leveraging insights from previous studies on feature selection, class imbalance handling, data security, and analytical approaches, this research lays the groundwork for more advanced Quick Count methods in the future. The successful implementation of XGBoost in this context opens the door for further exploration of technology and innovation in democracy and public participation beyond gubernatorial elections.

5. Conclusion

In conclusion, this research has demonstrated the effectiveness of implementing the XGBoost algorithm in modeling voice predictions for Quick Count, particularly in the context of the 2024 gubernatorial election. The study has established solid foundations for evaluating the XGBoost model's performance through meticulous research design and technical implementation. Using a range of evaluation metrics, including accuracy, precision, recall, and F1-score, the research provides a comprehensive understanding of how XGBoost improves the accuracy of Quick Count. The results indicate that the XGBoost algorithm achieves a high accuracy, precision, recall, and F1 score, showcasing its capability to classify voices accurately and identify favorable voters. This finding highlights the potential of XGBoost to significantly improve the Quick Count process, offering a more precise and real-time representation of election results.

Moreover, the careful experimental approach and representative data collection ensure that the developed model can adapt to the complexities of gubernatorial elections across diverse regions and conditions. The practical implications of this research are substantial, as it contributes to enhancing the democratic process by providing more reliable and transparent election results. Implementing XGBoost in Quick Count improves the accuracy of election predictions and fosters trust and confidence in the integrity of the electoral process.

Additionally, the study lays the groundwork for developing more advanced Quick Count methods in the future, leveraging insights from previous research on machine learning techniques, feature selection, class imbalance handling, and data security. This research signifies a significant step forward in leveraging technology and innovation to enhance election analysis and democratic participation. By successfully integrating the XGBoost algorithm into the Quick Count process, this study paves the way for further exploration of advanced analytical approaches in the context of democracy and public decision-making.

References

[1] Kompas.com. (2022, June 2). Pemilu dan Pilkada Serentak 2024: Alasan, Urgensi, dan Tantangan. URL: https://nasional.kompas.com/read/2022/06/02/14514481/pemilu-dan-pilkada-serentak-2024-alasan-urgensi-dan-tantangan

[2] CNA Indonesia. (2024). Semua hal yang perlu diketahui tentang pemilu serentak 2024 di Indonesia. URL: https://www.channelnewsasia.com/indonesia/pemilu-serentak-pemilu-presiden-pilkada-pemilu-legislatif-2024-indonesia-3865931

[3] Mahkamah Konstitusi Republik Indonesia. (2022). Pilkada Serentak 2024 Akibatkan Masa Jabatan Kepala Daerah Berkurang. URL: https://www.mkri.id/index.php?page=web.Berita&id=18612. Penulis: Utami Argawati

[4] Undang-undang (UU) No. 22 Tahun 2014. (n.d.). Pemilihan Gubernur, Bupati, dan Walikota. URL: https://peraturan.bpk.go.id/Details/37675/uu-no-22-tahun-2014

[5] Saputra, A.Y., & Apriadi, D. (2018). Rancang Bangun Aplikasi Quick Count Pilkada Berbasis Sms Gateway Dengan Metode Simple Random Sampling (Studi Kasus Kota Lubuklinggau). *Journal Information System Development (ISD), 3*(1).

[6] Priyanda, R., Amalia, R., Zaki, M., Fadilah, F., Fadhelina, N., & Ihsan, I.R. (2021). Quick count training for junior high school teachers in Gayo Lues District. *Community Empowerment, 6*(5), 877-884. DOI: https://doi.org/10.31603/ce.4584

[7] Sari, D.F., Kusjani, A., Kurniawati, D., & Setiawan, I. (2023). PENCARIAN DATA QUICK COUNT PILPRES DENGAN TEKNIK WEB SCRAPING. *Journal of Innovation Research and Knowledge, 3*(5), 1025-1034.

[8] Ayunis, A., & Minto, D.W. (2022). THE ABILITY TO UNDERSTAND STUDENTS MATHEMATICAL CONCEPTS USING THE QUICK COUNT METHOD IN ELEMENTARY SCHOOL. In *Proceedings of Imam Bonjol International Conference on Islamic Education* (pp. 280-285).

[9] Anzarut, M., González, L.F., & Ortiz, M.T. (2019). A heavy-tailed multilevel mixture model for the quick count in the Mexican elections of 2018. In *Selected Contributions on Statistics and Data Science in Latin America: 33 FNE and 13 CLATSE, 2018, Guadalajara, Mexico, October 1– 5* (pp. 1-13). Springer International Publishing. DOI: https://doi.org/10.1007/978-3-030-31551-1_1.

[10] Wakhyudi. (2019). Reviewing Presidential Candidates and Political Parties Reactions on the Results of 2019 General Election Quick Count. *Proceedings of the International Conference of Democratisation in Southeast Asia (ICDeSA 2019)*. DOI: https://doi.org/10.2991/icdesa-19.2019.2.

[11] Anzarut, M., González, L.F., & Ortiz, M.T. (2022, December). Bayesian Hierarchical Multinomial Modeling of the 2021 Mexican Election Outcomes with Censored Samples. In *Interdisciplinary Statistics in Mexico: AME Virtual Meeting, September 10–11, 2020, and 34 FNE, Acatlán, Mexico, September 22–24, 2021* (pp. 49-64). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-12778-6_3.

[12] Afrinaldi, E., & Evanita, S. (2023, June). The Influence of Digital Based Political Marketing on Governor Election Decisions in the 2020 Simultaneous Regional Elections with Political Trust as a Moderating Variable. In *Ninth Padang International Conference On Economics Education, Economics, Business and Management, Accounting and Entrepreneurship (PICEEBA 2022)* (pp. 550-585). Atlantis Press. DOI: https://doi.org/10.2991/978-94-6463-158-6_50.

[13] Wusu, A.S., Olabanjo, O.A., Padonu, R.M., & Mazzara, M. (2023). Twitter Sentiment Analysis of Lagos State 2023 Gubernatorial Election Using BERT. https.://doi.org/10.20944/preprints202303.0335.v1

[14] Firdaus, A.A., Yudhana, A., & Riadi, I. (2024). Indonesian presidential election sentiment: Dataset of response public before 2024. *Data in Brief, 52*, 109993. https://doi.org/10.1016/j.dib.2023.109993.

[15] Jones, D.W., Park, S., Rivest, R.L., & Sealfon, A. (n.d.). Scan, Shuffle, Rescan: Two-Prover Election Audits With Untrusted Scanners.

[16] Shevtsov, A., Tzagkarakis, C., Antonakaki, D., & Ioannidis, S. (2022, May). Identification of twitter bots based on an explainable machine learning framework: the US 2020 elections case study. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 16, pp. 956-967). DOI: https://doi.org/10.1609/icwsm.v16i1.19349.

[17] Nsaif, A.A., & Abd, D.H. (2022). Political Post Classification based on Firefly and XG Boost. *Journal of Algebraic Statistics, 13*(3), 4079-4101.

[18] Huang, S., Ji, J., Wang, Y., Li, W., & Zheng, Y. (2024). Development and validation of a soft voting-based model for urban fire risk prediction. *International Journal of Disaster Risk Reduction, 101*, 104224. DOI: https://doi.org/10.1016/j.ijdrr.2023.104224.

[19] Shrivastava, A., & Ramaiya, M.K. (2024). System for Managing Pesticide Recommendation on the Cotton Crop using Deep Learning Techniques VGG and Xgboost. *International Journal of Intelligent Systems and Applications in Engineering, 12*(7s), 677-691.

[20] Guan, J., Yu, Z., Liao, Y., Tang, R., Duan, M., & Han, G. (2024). Predicting Critical Path of Labor Dispute Resolution in Legal Domain by Machine Learning Models Based on SHapley Additive exPlanations and Soft Voting Strategy. *Mathematics, 12*(2), 272. DOI: https://doi.org/10.3390/math12020272.

[21] Kurniawan, R.C. (2018). Quick count (metode hitung cepat) dalam perspektif pemilukada. *Sawala : Jurnal Administrasi Negara, 2*(2), 78-88. https://doi.org/10.30656/sawala.v2i2.511

[22] Pradita, R. (2022). Aplikasi quick count dalam pemilihan bupati di kabupaten nganjuk. *Generation Journal, 6*(1), 22-30. https://doi.org/10.29407/gj.v6i1.15885

[23] Saraswati, R., & Prasakti, A. (2020). Analisis kepercayaan masyarakat terhadap hasil quick count pada pemilihan presiden dan legislatif di desa baturetno kabupaten wonogiri tahun 2019. *Jurnal Kewarganegaraan, 4*(1). https://doi.org/10.31316/jk.v4i1.882

[24] Candraningrum, D.A. (2017). Brand activation, strategi komunikasi pada lembaga survei politik di putaran pertama pilkada dki jakarta 2017. *Jurnal Muara Ilmu Ekonomi Dan Bisnis, 1*(1), 71. https://doi.org/10.24912/jmieb.v1i1.408

[25] Rizani, M.R. (2017). Quick count pemilihan kepala daerah berbasis web. *Jurnal Sains Dan Informatika, 3*(2), 89-93. https://doi.org/10.34128/jsi.v3i2.106

[26] Juanita, S. (2020). Analisis sentimen persepsi masyarakat terhadap pemilu 2019 pada media sosial twitter menggunakan naive bayes. *Jurnal Media Informatika Budidarma, 4*(3), 552. https://doi.org/10.30865/mib.v4i3.2140

[27] Putra, F. P., Huda, O. N., & Jaroji, J. (2018). Aplikasi e-voting berbasis mobile menggunakan optical character recognition. *Jurnal Informatika Polinema, 4*(3), 235-239. https://doi.org/10.33795/jip.v4i3.199

[28] Wulandari, D. A. P., Sudarma, M., & Paramaita, N. (2019). Pemanfaatan big data media sosial dalam menganalisa kemenangan pilkada. *Majalah Ilmiah Teknologi Elektro, 18*(1), 101. https://doi.org/10.24843/mite.2019.v18i01.p15

[29] Andika, L. A., Azizah, P. A. N., & Respatiwulan, R. (2019). Analisis sentimen masyarakat terhadap hasil quick count pemilihan presiden indonesia 2019 pada media sosial twitter menggunakan metode naive bayes classifier. *Indonesian Journal of Applied Statistics, 2*(1), 34. https://doi.org/10.13057/ijas.v2i1.29998

[30] Silih Windanu, L., Sartika Wiguna, A., & Endy Budianto, A. (2024). Optimasi metode k-nearest neighbor berbasis particle swarm optimization untuk analisis sentimen pemilihan presiden indonesia tahun 2024-2029. *JATI (Jurnal Mahasiswa Teknik Informatika), 8*(1), 35-42. https://doi.org/10.36040/jati.v8i1.6874

[31] Hernesniemi, J., Mahdiani, S., Tynkkynen, J., Lyytikäinen, L., Mishra, P. P., Lehtimäki, T., … & Oksala, N. (2019). Extensive phenotype data and machine learning in prediction of mortality in acute coronary syndrome – the maddec study. *Annals of Medicine, 51*(2), 156-163. https://doi.org/10.1080/07853890.2019.1596302

[32] Liu, H., Li, J., Leng, J., Wang, H., Liu, J., Li, W., … & Yang, X. (2020). Machine learning risk score for prediction of gestational diabetes in early pregnancy in tianjin, china. *Diabetes/Metabolism Research and Reviews, 37*(5). https://doi.org/10.1002/dmrr.3397

[33] Georganos, S., Grippa, T., Vanhuysse, S., Lennert, M., Shimoni, M., & Wolff, É. (2018). Very high resolution object-based land use–land cover urban classification using extreme gradient boosting. *IEEE Geoscience and Remote Sensing Letters, 15*(4), 607-611. https://doi.org/10.1109/lgrs.2018.2803259

[34] Wu, Z., Lei, T., Shen, C., Wang, Z., Cao, D., & Hou, T. (2019). Admet evaluation in drug discovery. 19. reliable prediction of human cytochrome p450 inhibition using artificial intelligence approaches. *Journal of Chemical Information and Modeling, 59*(11), 4587-4601. https://doi.org/10.1021/acs.jcim.9b00801

[35] Li, S., Qin, J., He, M., & Paoli, R. (2020). Fast evaluation of aircraft icing severity using machine learning based on xgboost. *Aerospace, 7*(4), 36. https://doi.org/10.3390/aerospace7040036

[36] Hong, W., Zhou, X., Jin, S., Lu, Y., Pan, J., Lin, Q., … & Pan, J. (2022). A comparison of xgboost, random forest, and nomograph for the prediction of disease severity in patients with covid-19 pneumonia: implications of cytokine and immune cell profile. *Frontiers in Cellular and Infection Microbiology, 12*. https://doi.org/10.3389/fcimb.2022.819267

[37] Zhang, Y., Zhang, X., Razbek, J., Li, D., Xia, W., Bao, L., … & Cao, M. (2022). Opening the black box: interpretable machine learning for predictor finding of metabolic syndrome. *BMC Endocrine Disorders, 22*(1). https://doi.org/10.1186/s12902-022-01121-4

[38] Cox, G. W., & Munger, M. C. (1989). Closeness, expenditures, and turnout in the 1982 u.s. house elections. *American Political Science Review, 83*(1), 217-231. https://doi.org/10.2307/1956441

[39] Atkeson, L. R., & Partin, R. W. (1995). Economic and referendum voting: a comparison of gubernatorial and senatorial elections. *American Political Science Review, 89*(1), 99-107. https://doi.org/10.2307/2083077.

[40] Parlina, I., Windarto, A. P., Wanto, A., & Lubis, M. R. (2018). Memanfaatkan algoritma k-means dalam menentukan pegawai yang layak mengikuti asessment center untuk clustering program sdp. *Computer Engineering, Science and System Journal, 3*(1), 87. https://doi.org/10.24114/cess.v3i1.8192

[41] Nurhendratno, S. S. (2022). Sintesis fitur density based feature selection (dbfs) dan adaboots dengan xgboost untuk meningkatkan performa model prediksi. *Prosiding Sains Nasional Dan Teknologi, 12*(1), 305. https://doi.org/10.36499/psnst.v12i1.6997

[42] Efrizoni, L., Defit, S., Tajuddin, M., & Anggrawan, A. (2022). Komparasi ekstraksi fitur dalam klasifikasi teks multilabel menggunakan algoritma machine learning. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer, 21*(3), 653-666. https://doi.org/10.30812/matrik.v21i3.1851

[43] Sudriyanto, S., Listrianti, F., & Jamal, J. (2022). Implementasi algoritma c4.5 untuk memprediksi kesesuaian gaya belajar siswa sekolah dasar. *COREAI: Jurnal Kecerdasan Buatan, Komputasi Dan Teknologi Informasi, 3*(2). https://doi.org/10.33650/coreai.v3i2.5074

[44] Indahyanti, U., Azizah, N. L., & Setiawan, H. (2022). Pendekatan ensemble learning untuk meningkatkan akurasi prediksi kinerja akademik mahasiswa. *Jurnal Sains Dan Informatika, 8*(2). https://doi.org/10.34128/jsi.v8i2.459

[45] Putra, I. R., Widiastiwi, Y., & Chamidah, N. (2022). Pengaruh seleksi fitur particle swarm optimization terhadap sentimen analisis aplikasi pedulilindungi di twitter dengan algoritma support vector machine. *Informatik : Jurnal Ilmu Komputer, 18*(3), 208. https://doi.org/10.52958/iftk.v18i3.4681

[46] Kaligis, G. B., & Yulianto, S. (2022). Analisa perbandingan algoritma k-means, k-medoids, dan x-means untuk pengelompokkan kinerja pegawai. *IT-Explore: Jurnal Penerapan Teknologi Informasi Dan Komunikasi, 1*(3), 179-193. https://doi.org/10.24246/itexplore.v1i3.2022.pp179-193

[47] Nugraha, W. A., & Syarif, M. (2023). Teknik weighting untuk mengatasi ketidakseimbangan kelas pada prediksi churn menggunakan xgboost, lightgbm, dan catboost. *Techno.Com, 22*(1), 97-108. https://doi.org/10.33633/tc.v22i1.7191

[48] Martawireja, A. R. H., Ridwan, R., Hafidzin, A. P., & Taufik, M. (2021). Proteksi keamanan data pada quick response (qr) code. *Jurnal Teknologi Dan Rekayasa Manufaktur, 3*(2), 99-110. https://doi.org/10.48182/jtrm.v3i2.58

[49] Setiawan, A. (2020). Penggunaan resampling dalam penggambaran quick count. *Jambura Journal of Probability and Statistics, 1*(1), 1-12. https://doi.org/10.34312/jjps.v1i1.5643

[50] Anam, H., Nurhayati, N., & Manalu, R. V. B. (2020). Persepsi kegunaan, persepsi kemudahan, keamanan dan kerahasiaan, kesiapan teknologi informasi terhadap intensitas perilaku dalam penggunaan e-filing. *Jurnal GeoEkonomi, 11*(1), 1-15. https://doi.org/10.36277/geoekonomi.v11i1.104

[51] Pahlevi, R. F. (2021). Implementasi algoritma weighted sum model dalam sistem penjaminan mutu perguruan tinggi. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi), 8*(2), 935-946. https://doi.org/10.35957/jatisi.v8i2.700.