ABUAD Journal of Engineering Research and Development (AJERD)
ISSN (online): 2645-2685; ISSN (print): 2756-6811

Volume 8, Issue 1, 122-128

Machine Learning-Based Multimodal Biometric Authentication System (Facial and Fingerprint Recognition) for Online Voting Systems

Solomon OMOZE, Samuel OMAJI, Glory Nosawaru EDEGBE

Department of Computer Science, Edo State University, Uzairue
omozesolomon@yahoo.com/omaji.samuel@edouniversity.edu.ng/edegbe.glory@edouniversity.edu.ng

Corresponding Author: omaji.samuel@edouniversity.edu.ng, +2348061932007
Date Submitted: 12/08/2024
Date Accepted: 08/02/2025
Date Published: 11/02/2025

Abstract: *Online voting systems offer many advantages over traditional voting methods, such as paper ballots, this is because paper ballots face some well-known challenges, ranging from logistics, susceptibility to tampering, and the requirement for voters to be physically present at polling stations. In contrast, online voting systems offer the potential to overcome these challenges by providing a convenient and accessible means for citizens to cast their votes from anywhere. However, online voting systems must address significant security and authentication challenges to ensure that each vote is cast by a legitimate and unique voter, maintaining the integrity of the electoral process. This project proposes the development of a machine learning authentication module that can be integrated into an online voting system using facial recognition and fingerprint recognition to enhance the security of online voting. The system therefore consists of two main components; the machine learning-based authentication component and the web-based voting platform. The authentication component uses machine learning algorithms to accurately and reliably verify the identities of voters based on their biometric data. The web-based platform facilitates voter registration, authentication, and voting processes, ensuring a seamless and secure user experience. These two components were implemented first by obtaining a comprehensive database of user biometric data, training the machine learning module, and implementing a user-friendly web interface using Java Server Pages (JSP) and a MySQL database. The system's performance was evaluated using established metrics, including accuracy, precision, recall, and R2 Score with the following values 98%, 1.0, 0.8 and 0.78 respectively.*

Keywords: *Online Voting, Biometric Authentication, Machine Learning, Facial Recognition, Fingerprint Recognition, JSP, MySQL*

1. INTRODUCTION

As humans depend more and more on computers and digital tools, there has been a growing interest in developing secure and efficient online voting systems. Traditional voting methods, including paper ballots and electronic voting machines, have several limitations, such as logistical challenges, susceptibility to tampering, and the need for physical presence at polling stations. Online voting systems on the other hand, offer the potential to overcome most of these challenges by providing a convenient and accessible means for citizens to cast their votes from anywhere.

However, online voting systems face significant security and authentication challenges. Ensuring that each vote is cast by a legitimate and unique voter is critical to maintaining the integrity of the electoral process. Traditional authentication methods, such as usernames and passwords, are insufficient due to their vulnerability to phishing, hacking, and other forms of cyber-attacks. Consequently, there is a need for more robust authentication mechanisms that can provide higher levels of security and reliability. Deploying online voting systems therefore presents new challenges, chief among them is safety, with risks of cyber-attacks, hacking, and tampering. Checking who voters are online is tricky; as usual methods like passwords as earlier stated can be cracked. Keeping trust and openness in how people vote how votes are saved, and counted is key for public trust in the electoral process [1].

Biometric authentication, which uses unique physiological or behavioral characteristics to verify an individual's identity, has emerged as a promising solution for enhancing the security of online voting systems. Multimodal biometric systems combine multiple biometric modalities (e.g., facial recognition, fingerprint recognition), offering even greater security by requiring multiple forms of biometric verification [2]. These systems makes use machine learning algorithms to improve the accuracy and reliability of biometric recognition, making them well-suited for high-security applications like online voting.

To tackle these issues, this research aims to develop a machine learning-based multimodal biometric authentication system that uses facial and fingerprint recognition to enhance the security and reliability of online voting. This method improves security by giving us a strong way to check that voters are who they say they are.

2. RELATED WORKS

Current research works are beginning to focus on online voting, this has been on the increase because of new technological breakthroughs in computing and AI and the search for better ways to include everyone in democracy. The

big issue is how to make sure voters are who they say they are online, which is key to keep elections fair and real. In this section a few related works are presented providing an overview of current research and perspectives on the challenges and solutions related to enhancing authentication in online voting systems, highlighting the potential of bimodal bio-metrics and machine learning to address these challenges effectively.

The work by [3] outlines key advantages of online voting when compared to paper voting, these are in making it easier for people to vote faster to count, and getting more people out to vote. It lets people vote from anywhere so distance isn't a problem and we don't need as many voting places. This helps people with disabilities or those who live far [4]. Despite these advantages, [5] Outlines key challenges that a full deployment of online voting system would face. This majorly includes security and trust issues. Safety weak spots like cyber-attacks and hacking risks continue to pose big threats.

Recent progress in biometric technology has made possible the combination of two-mode bio-metrics (combining different biometric types) and machine learning algorithms. Biometric types like fingerprints and face scans give unique and trustworthy IDs that are hard to fake or copy [6]. The works in [7, 8, 9] showed the different techniques of multimodal bio-metrics systems for online voting. Other recent applications of machine learning applications for solving contemporary societal problems in business, commerce etc has also been reported in [10, 11, 12].

The research works in [13, 14] showed the application of Machine learning methods with biometric data for online voting system. Their work showed accuracy and adaptability over time. This ensures that online voting systems stay strong against new threats from potential attackers.

This work proposes a combination of more than one bio-metrics, in this case finger and facial recognition to build a more secure online voting system. The works further incorporates the machine learning authentication system into a web application to demonstrate its practicability.

2.1 Background of Selected Machine Learning Algorithm

In this sub-section, we present the materials needed to model the network.

2.1.1 Dataset

The finger print and facial pictures dataset where obtained from kaggle.com and converted to a128-dimensional feature vectors Number of users. These two datasets were combined using numpy's hstack method to allow further analysis and data preprocessing activities. Saving these images dimensional vectors rather than jpeg and PNG images allows the use of random forest classifier to build the machine learning model. Figure 1 shows a snapshot of the dataset while Figure 2 shows its distribution.

```
# Load the dataset
df = pd.read_csv('biometric_data.csv')
df.head()
```

	feature_0	feature_1	feature_2	feature_3	feature_4	feature_5	feature_6	feature_7	feature_8	feature_9	...	feature_247	feature_248	f
0	0.101904	0.016009	0.765028	0.019518	0.556789	0.806352	0.994558	0.998396	0.453349	0.084416	...	0.148611	0.006376	
1	0.061437	0.341423	0.451205	0.511674	0.034419	0.507581	0.715682	0.496565	0.132686	0.651298	...	0.822347	0.030823	
2	0.833198	0.343513	0.736204	0.227143	0.624250	0.728805	0.958180	0.521640	0.456838	0.815795	...	0.548021	0.201706	
3	0.802557	0.042862	0.316285	0.871787	0.529877	0.093783	0.873991	0.165015	0.203585	0.327283	...	0.818036	0.125129	
4	0.523933	0.920984	0.035691	0.952185	0.498743	0.997147	0.562154	0.630680	0.435059	0.742543	...	0.286864	0.165519	

5 rows × 257 columns

Figure 1: Snapshot of converted dataset

The dataset consists of two hundred and fifty seven (257) variables representing the different features of the images of the biometric capture (finger prints and facial images).

2.2 Methods

Figure 3 shows the overall flow of the proposed system, showing the two major components which are;

i. The machine learning authentication component
ii. The web-based voting platform.

2.3 Random Forest Classifier

For decision tree-based classification, regression, and other uses, supervised machine learning techniques like the Random Forest, also called the Random Decision Forest, are employed. It works particularly well for handling big, complicated datasets, high-dimensional feature spaces, and providing feature importance insights [15]. This algorithm's capacity to maintain high predicted accuracy while lowering over fitting makes it popular in a variety of industries, including banking, healthcare, and image analysis. A collection of decision trees is produced by the Random Forest classifier using randomly selected portions of the training set. From these random subsets, a set of decision trees is built, and the final forecast is obtained by adding the votes from each tree [16]. In addition, the Random Forest classifier can

handle jobs involving both regression and classification. It is a priceless tool for determining the significance of different variables in a dataset due to its feature importance ratings. Figure 4, shows the structure of the random forest model.

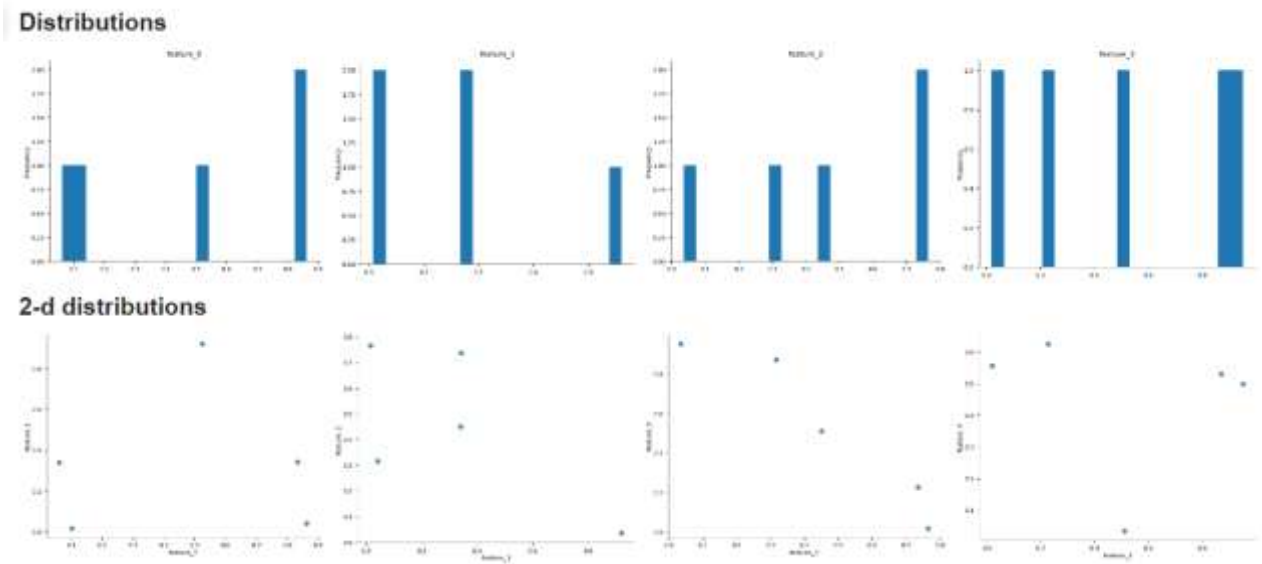

Figure 2: Feature distribution

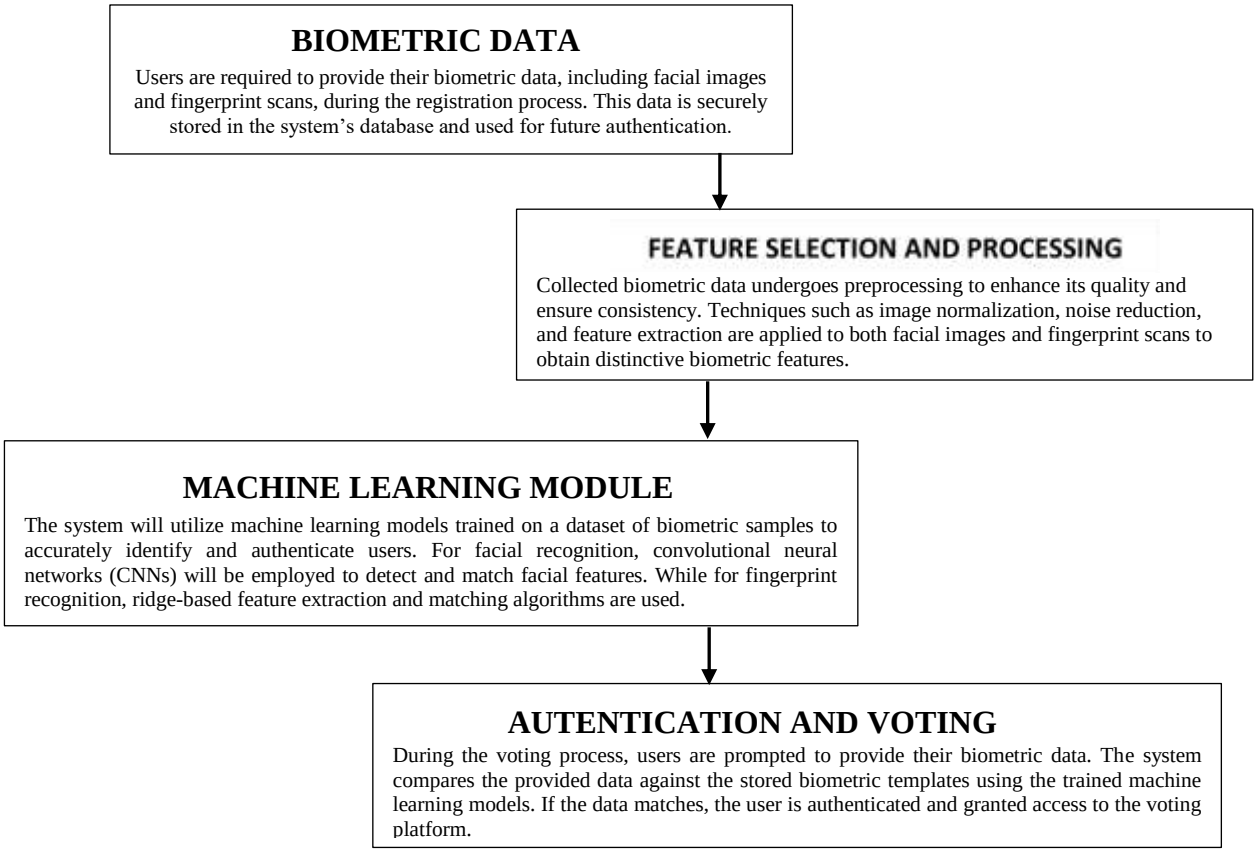

Figure 3: System flow diagram

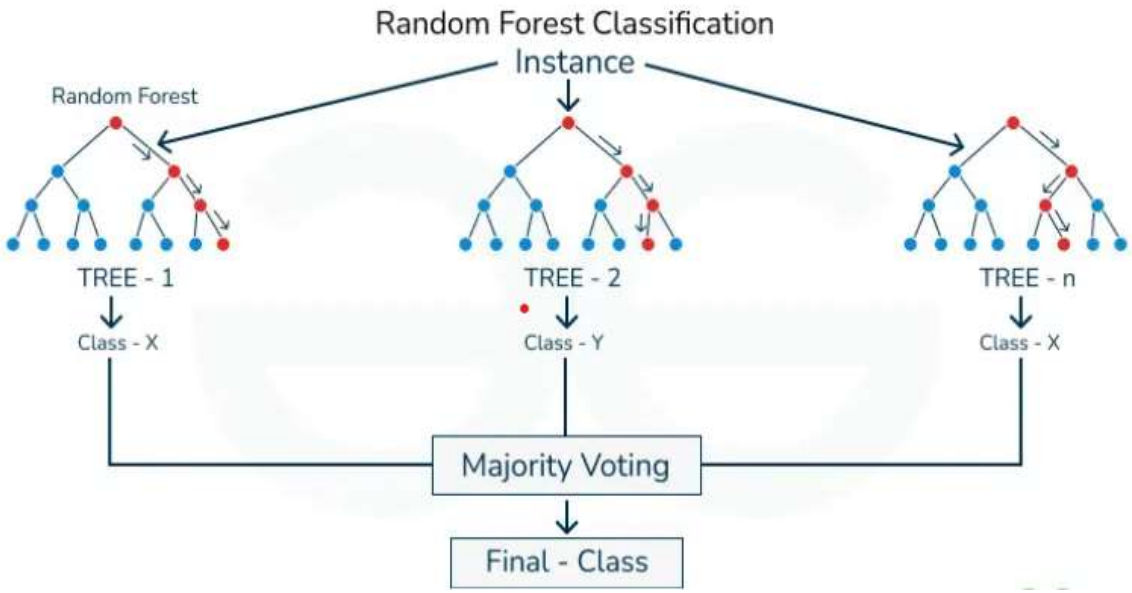

Figure 4: Random Forest model (Source: geekforgeeks.org)

Mathematically a random forest model can be broken down into the following key steps:

1. Bootstrap Sampling: A bootstrap sample is extracted from the training data for every B tree. N randomly chosen examples with replacement are taken from the training set to form a bootstrap sample.
2. Tree Construction: Each tree T_b where $(b = 1,2,3 \ldots . B)$ is grown using the bootstrap sample. During the construction of each tree:

   - At each node, a random subset of m features is chosen from the total p features$(m \leq p)$.
   - The best split is found among the selected mmm features.

3. Aggregation of Predictions: Once all the trees are constructed, the Random Forest makes predictions by aggregating the predictions of the individual trees.

For a classification task:

Each tree T_b gives a class prediction $C_b(x)$ for a given input x.

The final prediction $C(x)$ is determined by majority vote:

$$C(x) = mode\{C_1(x), C_2(x), \ldots \ldots \ldots \ldots . C_{1B}(x)\} \quad (1)$$

For a regression task:

Each tree T_b provides a predicted value $y_b(x)$ for a given input x.
The final prediction $y(x)$ is the average of all the tree predictions:

$$y(x) = \frac{1}{B}\sum_{b=1}^{B} y_b(x) \quad (2)$$

3. RESULTS AND DISCUSSION

3.1 The machine Learning Authentication Component

The machine learning authentication component is the backbone of the proposed e-voting system. The idea is to apply machine learning algorithms to enforce voter authentication and ensure that only legitimate voters can access the voting platform. This ML component will provide a high level of security and accuracy by employ multimodal biometric authentication that is, facial recognition and fingerprint recognition.

The random forest classifier model was build using python sklearn library after splitting the dataset in the ratio of 70:30 for training and testing purposes respectively. The results of evaluation using the accuracy score, precision, recall and f1-score shows 98% accuracy.

```
Accuracy : 0.978021978021978
Classification Report
              precision    recall  f1-score   support

           0       0.98      1.00      0.99       162
           1       1.00      0.80      0.89        20

    accuracy                           0.98       182
   macro avg       0.99      0.90      0.94       182
weighted avg       0.98      0.98      0.98       182

Confusion Matrix
[[162   0]
 [  4  16]]
Precision  1.0
Recall  0.8
R2 Score  0.7753086419753087
```

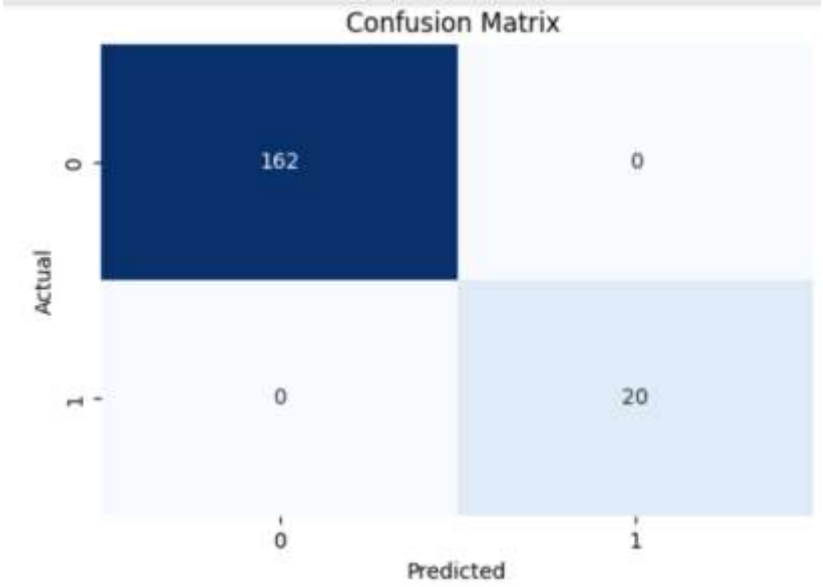

Figure 5: Evaluation results of the machine learning model

3.2 Web-Based Component

The web-based component of the e-voting system is designed to provide a easy to use and secure interface for voter registration, authentication, and voting. This component is implemented using Java Server Pages (JSP) and a MySQL database, ensuring robustness and scalability. Figure 6 shows the system flowchart illustrating the key aspects of the design and implementation.

The voter registration module is the first part of the application. It allows new users to sign up by providing necessary details and biometric data. This information especially the biometrics is then saved to database for future reference. After a successful registration, a voter can return to login in other to vote.

The login process requires the user to submit his biometrics again. This is then authenticated by the machine learning module that classifies the user as genuine or fake. If a user is successful at the authentication, he/she can then proceed to the voting dashboard, where he/she can view the list of candidates and their party affiliations. Figures 7,8 and 9 shows the registration, login and voters' dashboard.

The work presented in this research bears semblance to previous works that have proposed the use of biometrics for voter authentication in online voting systems. However this work presents a case scenario that is implemented as a web application that incorporates the three modules, which are; the voter registration module, the return voter authentication module (This incorporates the machine learning predictive model), and finally the online voting module. The presentation of this work therefore provides a complete system which is implemented as a web application for a given case study which allows for simulation and scenario analysis. When compared to previous works identified in literature, the machine learning prediction metrics show similar accuracy results (accuracy of over 98%), however this work presents a complete web based application, which demonstrates the practicability the system.

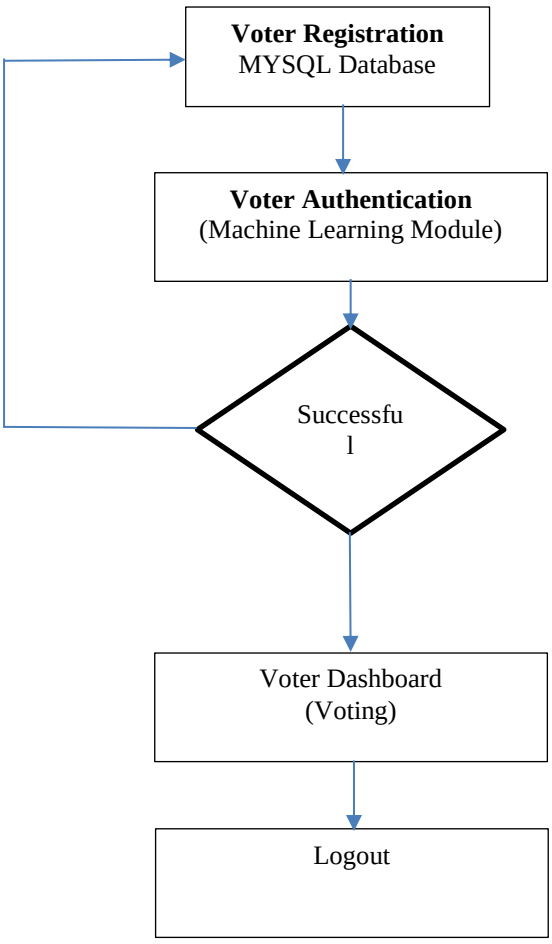

Figure 6: System Flowchart

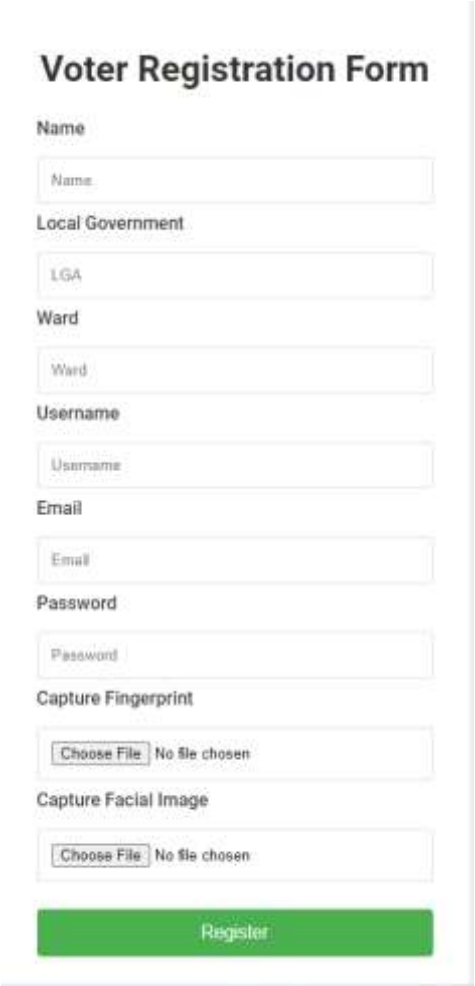

Figure 7: Registration Page

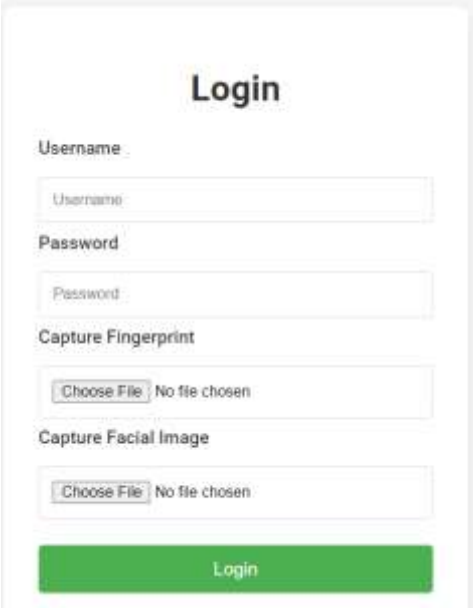

Figure 8: Login page

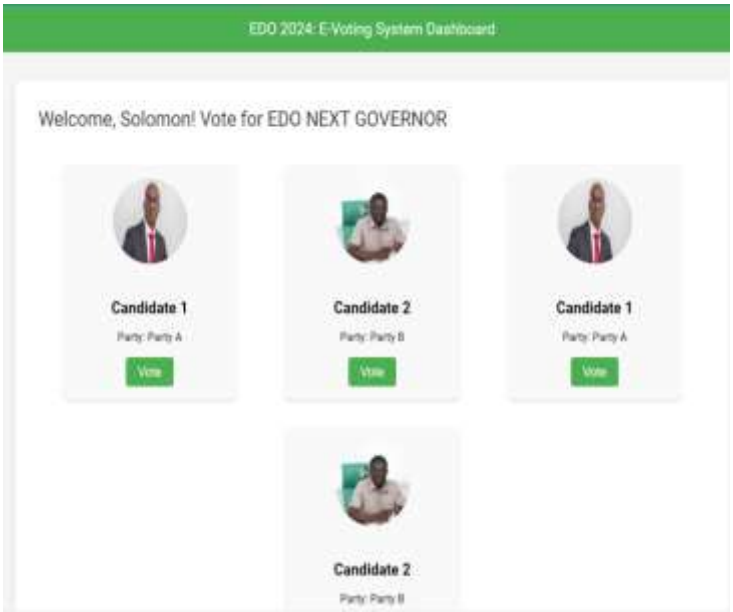

Figure 9: Voter Dashboard (Source: researchers' implementation)

4. CONCLUSION

The e-voting system presented in this work utilized a combination of web technologies and machine learning-based multimodal biometric authentication to assure a secure, efficient, and user-friendly platform for online voting. The project consists of two major components: the machine learning-based authentication component and the web-based voting platform. The ML component allows facial recognition and fingerprint recognition to ensure secure and reliable user authentication, incorporates machine learning algorithms to enhance the accuracy and robustness of biometric verification, and Addresses security concerns by employing multiple biometric modalities, thereby reducing the risk of fraud and unauthorized access. While the web based component; was built using Java Server Pages (JSP) for dynamic content generation and MySQL for database management, included key pages such as the homepage, registration page, login page, and voting dashboard, provides a seamless user experience with a responsive and accessible design, ensuring ease of use across various devices, ensures secure data handling through encryption and secure file upload mechanisms, and allows users to register, authenticate using their biometric data, and cast their votes in a straightforward and secure manner.

Future work may include further refinement of the machine learning algorithms to improve recognition accuracy, scalability testing to ensure the system can handle large numbers of users, and the incorporation of additional security measures to guard against emerging threats. This project has presented a significant step forward in the development of secure and efficient e-voting systems, contributing to the broader effort of modernizing the electoral process.

REFERENCES

[1] Okonta, D., 2019. *Exploring the Secure Online-Voting Strategies Information Assurance Specialists Need to Increase the Millennial Generation Voting Turnouts in a General Election* (Doctoral dissertation, Colorado Technical University).

[2] Dargan, S. & Kumar, M., 2020. A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities. *Expert Systems with Applications*, *143*, 113114.

[3] Katawazy, A. (2023). Identifying Challenges and Advantages of Internet Voting and Assessing the Impact on Voters Turnout in Municipal Elections.

[4] Grimaila, M. R., & Pfleeger, S. L. (2005). Security of electronic voting systems: Design requirements and security issues. IEEE Security & Privacy, 3(5), 26-33.

[5] Cetinkaya, O., Bicakci, K., & Savas, E. (2020). Security analysis of internet voting systems: A case study. Computers & Security, 94, 101873.

[6] Shaikh, T., Ranadhir, H., Gugale, S., Patil, V., & Kulkarni, O. (2022). Biometrics Based Secured Online Voting System Using Machine Learning Method. *EasyChair*.

[7] Singh, S. P., & Tiwari, S. (2023). A Dual Multimodal Biometric Authentication System Based on WOA-ANN and SSA-DBN Techniques. *Sci*, *5*(1), 10.

[8] Srivastava, R., Tomar, R., Sharma, A., Dhiman, G., Chilamkurti, N., & Kim, B. G. (2021). Real-Time Multimodal Biometric Authentication of Human Using Face Feature Analysis. *Computers, Materials & Continua*, *69*(1).

[9] Zhang, X., Cheng, D., Jia, P., Dai, Y., & Xu, X. (2020). An efficient android-based multimodal biometric authentication system with face and voice. *IEEE Access*, *8*, 102757-102772.

[10] Acheme, I. D., Nwankwo, W., Olayinka, A. S., Makinde, A. S., & Nwankwo, C. P. (2023). Petroleum Drilling Monitoring and Optimization: Ranking the Rate of Penetration Using Machine Learning Algorithms. In The International Conference on Artificial Intelligence and Logistics Engineering (152-164). Cham: Springer Nature Switzerland.

[11] Acheme, I. D., & Vincent, O. R. (2021). Machine-learning models for predicting survivability in COVID-19 patients. In Data Science for COVID-19, Academic Press, (317-336).

[12] Acheme, I. D., & Enoyoze, E. (2024). Customer personality analysis and clustering for targeted marketing. International Journal of Science and Research Archive, 2024, 12(01), 3048–3057 https://doi.org/10.30574/ijsra.2024.12.1.1003.

[13] Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology, 14(1), 4-20.

[14] Ross, A., & Jain, A. K. (2011). Multimodal biometrics: An overview. In Multimodal biometrics, Springer, New York, NY (pp. 3-22).

[15] Hermawan, D. R., Fatihah, M. F. G., Kurniawati, L., & Helen, A. (2021, October). Comparative study of J48 decision tree classification algorithm, random tree, and random forest on in-vehicle CouponRecommendation data. In *2021 International conference on artificial intelligence and big data analytics*, IEEE, 1-6.

[16] Khan, Z., Gul, A., Perperoglou, A., Miftahuddin, M., Mahmoud, O., Adler, W. and Lausen, B., 2020. Ensemble of optimal trees, random forest and random projection ensemble classification. *Advances in Data Analysis and Classification*, *14*, 97-116.

[17] GeeksforGeeks (2024) Image of random forest ML algorithm Available at: https://www.geeksforgeeks.org (Accessed: 23 October 2024).