LAUTECH Journal of Computing and Informatics (LAUJCI) – ISSN: 2714-4194
Volume 3 Issue 1, June 2023 – www.laujci.lautech.edu.ng

AN IMAGE BASED PARKINSON'S DISEASE DETECTION SYSTEM USING MACHINE LEARNING APPROACH

Olanloye D.O
College of Computing and Communication Studies
Bowen University, Iwo
odun.olanloye@unimed.edu.ng

Halleluyah O.A
College of Computing and Communication Studies
Bowen University, Iwo
aworinde.hallelyah@unimed.edu.ng

Ibitoye Ayodeji
College of Computing and Communication Studies
Bowen University, Iwo
ayodeji.ibitoye@bowen.edu.ng

Emuoyibofarhe O.N
College of Computing and Communication Studies
Bowen University, Iwo
ozichi.emuoyibofarhe@bowen.edu.ng

John Oloyede
College of Computing and Communication Studies
Bowen University, Iwo
john.oloyede@bowen.edu.ng

ABSTRACT

Quite several people have been sent to untimely grave and some with part of their body system paralyzed resulting from Parkinson Disease (PD). PD is a neurodegenerative disorder that affect neuron in the brain called substantial nigra with central movement. In this research work, an attempt is being made to design and implement a system for early detection of (PD) making use of relevant data (drawing of waves) from patient of PD and non-patient of the disease. Machine Learning models which include random forest (RF), Decision Tree (DT) and K- Nearest Neighbor (KNN), Dummy Classifier, Support Vector Machine (SVM), Convolutional Neural Network (CNN) were used to develop the model and the model in turn were used to develop the system. Machine Learning (ML) models helped in predicting whether the patient have PD or not using both spiral and wave drawing and hence, displays the result. Performance metrics such as Precision, recall and f1-score were used to determine the accuracy of the model result. Confusion matrix on the various algorithms were also displayed. Random Forest Classifier is discovered to be the most accurate machine learning algorithm for both the wave and spiral classification with an accuracy of 100% and a precision score for healthy (100%) and Parkinson's (100%) after the extracted features were trained for wave classification, and an accuracy of 80% with a precision score for healthy (75%) and Parkinson's (86%) after the extracted features were trained for spiral classification. Logistic Regression and Dummy Classifier had the worst accuracy with wave classification with an accuracy of 47%. Logistic Regression produced the worst accuracy (47%) in term of spiral classification. The best model (Random Forest) with 100% accuracy was used to develop the system and hence, the system was found to be highly efficient and far better than the existing methods or techniques.

Keywords: Machine Learning Algorithms, Parkinson Disease, Spiral and Wave Classifiers, Logistic Regression, Random Forest, Dummy Classifier

1. INTRODUCTION

According to [1], Human being by virtue of their nature can think naturally on their own therefore possesses natural intelligence. Hence, they can think and act appropriately. On the other hand, computer machines cannot think or act because they are not naturally intelligent. Artificial intelligence is created in computer machines making them think and act appropriately in an attempt to solve real-life problems in a human-like manner [2]–[4].
Machine learning is one of the major areas of AI where different models are trained using relevant learning algorithms to acquire enough knowledge or experience that enables the system to solve future problems without being explicitly programmed[2][5][6]. The model learns from experience and uses such experience to solve similar future problems. ML models have been found to be more accurate and faster at diagnosis of diseases than human doctors [7][8]
Parkinson's disease is a very deadly and dangerous disease. The major challenge of this disease is that it is very difficult to detect at its early stage. This made it difficult to establish a specific testing process for the disease. The major symptoms noticed in PD patients are tremors, stiffness, slow movement, and rigidity of some parts of the body system.

The patient may also have some sleepless nights. But these symptoms are misleading since they may also indicate some other diseases [9]–[11].
To solve this problem, series of research were embarked on with the central mission of finding the lasting solution to the problem. Literature reveals that most of the authors made use of ML models to identify or classified PD. But this research work takes a step further by not just using ML models to identify the disease but use such models to develop a system that can be used to identify the disease at an early stage. The use of spiral and wave data also adds more flavour to the novelty of our work[12]–[15].
In this research, the system aids in the early detection of Parkinson's disease using machine learning principles. The system is to help the doctors to reduce the cost and the time taken to monitor the patient. It assists the doctors in detecting the disease as early as possible and helps in avoiding the problem of a wrong diagnosis. This promotes the health sector and reinforces the workforce, which automatically improves the economy of the country. In the long run, the standard of living is drastically improved.
Various algorithms were used to train the models, and the best of the models of 100% accuracy produced by Random Forest was picked to develop the system [16], [17][12], [14], [18], [19]. Hence a system of almost 100% accuracy was developed. The system was implemented using Python programming language, tested, and found to be more efficient compared to the existing methods.

2. LITERATURE REVIEW

[10] Presented a research work which was on Parkinson's disease detection and classification using ML and deep learning algorithms. Different classification models were used to test the samples of voices of PD patients. It was finally concluded that the deep belief classifier proved to be more efficient compared to others, with an accuracy value of 94%

[20] Published a research article making use of ML approaches for the diagnosis of Parkinson disorder through speech data. The authors presented a comparative analysis of various feature extraction algorithm used in the classification and prediction of PD. It was reported that the range of prediction accuracy lies between 85% and 95%, which is good enough. This provided evidence that ML algorithms are capable enough to predict Parkinson disease by analyzing voice signals.

[21] Published a research article that can detect Parkinson's disease using is surface-based features and convolutional Neural Network. In this research work, relevant information used as input was obtained from CNN architecture. Hence a classification system with CNN architecture was implemented. The accuracy of the system was reported as 95.1% and the area under curve AVC as 97%, which appeared to compete with values obtained from the existing methods. It was concluded that computation on its surface reduces the complexity of input and keeps classification accuracy on the high side.

[22] Proposed a research work purposely on the classification of PD, making use of MRI (Magnetic Resonance Image) data and deep learning algorithm. The MRI data of PD patients were classified with an accuracy of 97.63% with batch normalization, but without batch normalization, an accuracy of 97.91% was reported. Hence it was concluded that the deep learning CNN provides future scope for medical image analysis.

[23] Published a research work titled algorithm for image-based biomarker detection for differential diagnosis of PD. The research methodology comprised of 3 steps. The preprocessed (MRI) was modeled using a self-organizing map (SOM), fiber discriminant Radio was used to reveal distinctive features, and SVM was used for patient classification. The research contributed a lot in the early-stage diagnosis of PD as well as determining its progression.

3. METHODOLOGY

`

The first step is to gather the PD drawing dataset and classify them into various groups. Physical features were extracted from two different drawings, one of the waves and the other of a spiral. The features extracted were trained using deep learning and classified machine learning algorithms such as Convolutional Neural Network (CNN), Naive Bayes Classifier (NBC), K-Nearest Neighbour (KNN), and support vector machine (SVM), Decision Tree, Random Forest, and Dummy Classifier[17], [24], [25].

The models were evaluated using a performance matrix called a confusion matrix. The mean absolute error was obtained, which is the difference between the original values and the predicted values. It gives us the measure of how far the predictions were from the actual output. The accuracy, Precision, recall, and f1-score were also obtained. The most accurate model was then chosen to develop the system.

3.1 Data Collection

A secondary dataset was used in the implementation of the research. The data (shown in fig 1) was obtained from the experiments conducted at PD outpatient clinic and deadening neurology, Melbourne, Australia. The experimental report protocol was approved by RMIT University, Human Research Ethics committee, and in accordance with the Declaration of Helsinki (revised 2004). The image data (spiral and wave) were uploaded into the system, preprocessed, and resized to the application specification. This is followed by feature extraction and segmentation to ensure that the data was standardized. This simplifies the image such that it becomes more meaningful to analyze. The data is then classified into training and testing data. After a set of data has been trained, the testing is done with the remaining set.

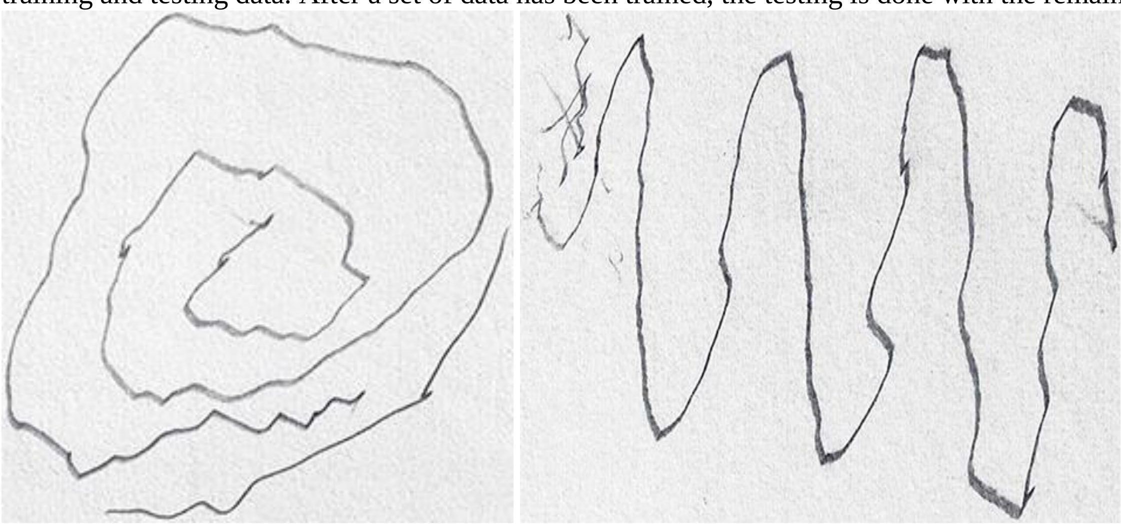


Fig 1 Spiral and Wave data

4. DESCRIPTION OF PROPOSED SYSTEM

This system assists the doctors in the automatic detection of Parkinson's disease, using composite index of sketching of spirals and waves. When the drawing is uploaded onto the software, the images are preprocessed and resized to the applications specification. The goal of feature extraction and image segmentation is to simplify the representation of the images into something that is more meaningful and easier to analyze.

The proposed system describes the nature of the data, explore the relation of data, and create a model to summarize the understanding, prove the validity of the model that helps in deciphering if the patient has Parkinson's disease, or he/she is healthy. Fig 2 below shows the architectural view of the proposed system:

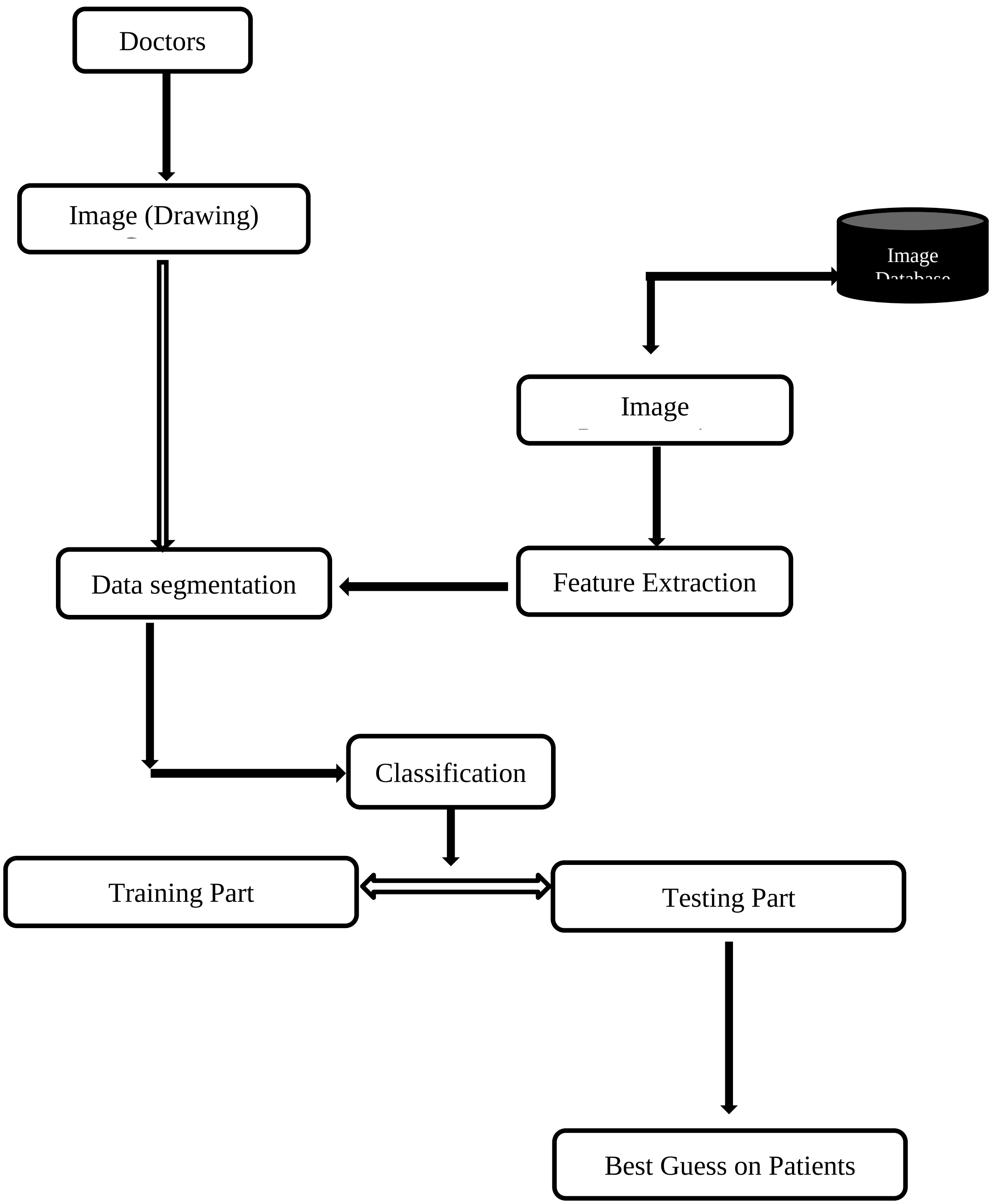
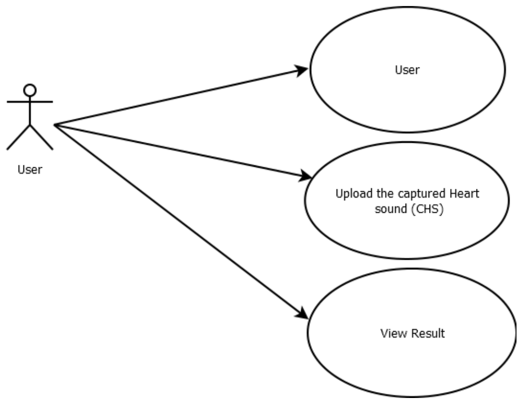
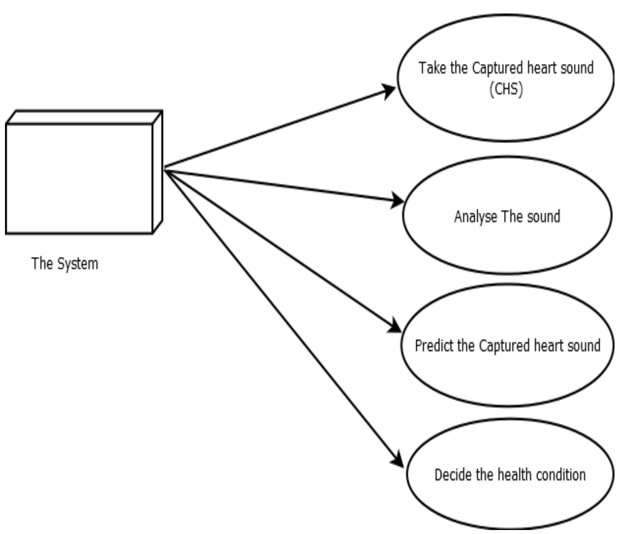


Fig 2 Architectural view of the proposed system

Fig 3: Use Case Diagram for the user





4.1 System Implementation and Evaluation

Multiple ML models were used to predict the presence or absence of Parkinson disease and to determine the extent of the disease. The model made use of Spiral and wave data shown in figure 1 above. The sample of results obtained at various implemental level are as shown Figures 4 and 5.

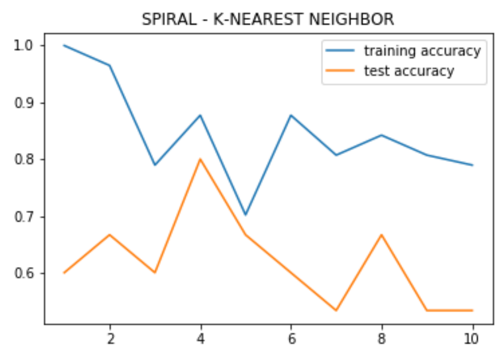


Figure 4

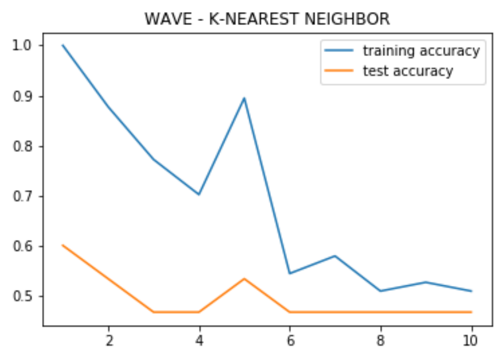


Figure 5

KNN algorithm accuracy plot for both spiral and wave are shown in figure 4 and 5 above

Table 1: Performances of Classical Machine Learning Algorithms

		Wave				Spiral			
ML Algorithm	**image category**	**precision**	**recall**	**f1-score**	**support**	**precision**	**recall**	**f1-score**	**support**
Dummy Classifier	**healthy**	0.47	1	0.64	7	0.47	1	0.64	7
	parkison	0	0	0	8	0	0	0	8
	accuracy			0.47	15			0.47	15
KNN(Neighbour =5)	**healthy**	0.58	1	0.67	7	0.75	0.43	0.55	7
	parkison	1	0.12	0.22	8	0.64	0.88	0.74	8
	accuracy			0.53	15			0.67	15
Logistics Regression	**healthy**	0.46	0.86	0.6	7	0.67	0.86	0.75	7
	parkison	0.5	0.12	0.2	8	0.83	0.62	0.71	8
	accuracy			0.47	15			0.73	15
Support Vector Machine	**healthy**	0.55	0.86	0.67	7	0.67	0.86	0.75	7
	parkison	0.75	0.38	0.5	8	0.83	0.62	0.71	8
	accuracy			0.6	15			0.73	15
Decision Tree	**healthy**	0.71	0.71	0.71	7	0.71	0.71	0.71	7
	parkison	0.75	0.75	0.75	8	0.75	0.75	0.75	8
	accuracy			0.73	15			0.73	15
Random Forest	**healthy**	1	1	1	7	0.75	0.86	0.8	7
	parkison	1	1	1	8	0.86	0.75	0.8	8
	accuracy			1	15			0.8	15

Table 1 below shows the different accuracy values and confusion matrix on the various algorithms used. Random Forest Classifier is the most accurate machine learning algorithm for both the wave classification and spiral classification with an accuracy of 100% with a precision score for healthy 100% and Parkinson's 100% after the extracted features were trained for wave classification, and an accuracy of 80% with a precision score for healthy 75% and Parkinson's 86% after the extracted features were trained for spiral classification. In contrast, both Logistic Regression and Dummy Classifier had the worst accuracy of 47% after the extracted features were trained for spiral classification.

Table 2: Wave Classification (Random Forest)

Classification Report for New Data(Wave) - Random Forest

	precision	recall	f1-score	support
healthy	0.73	0.73	0.73	15
parkison	0.73	0.73	0.73	15
accuracy			0.73	30
macro avg	0.73	0.73	0.73	30
weighted avg	0.73	0.73	0.73	30

Table 3: Spiral Classification (Random Forest)

Classification Report for New Data(Spiral) - Random Forest

	precision	recall	f1-score	support
healthy	0.67	0.80	0.73	15
parkison	0.75	0.60	0.67	15
accuracy			0.70	30
macro avg	0.71	0.70	0.70	30
weighted avg	0.71	0.70	0.70	30

Table 4: Wave Classification (Decision Tree)

Classification Report for New Data (Wave) - Decision Treee

	precision	recall	f1-score	support
healthy	0.64	0.60	0.62	15
parkison	0.62	0.67	0.65	15
accuracy			0.63	30
macro avg	0.63	0.63	0.63	30
weighted avg	0.63	0.63	0.63	30

Classification Report for New Data(Spiral) - Decision Tree

	precision	recall	f1-score	support
healthy	0.50	0.47	0.48	15
parkison	0.50	0.53	0.52	15
accuracy			0.50	30
macro avg	0.50	0.50	0.50	30
weighted avg	0.50	0.50	0.50	30

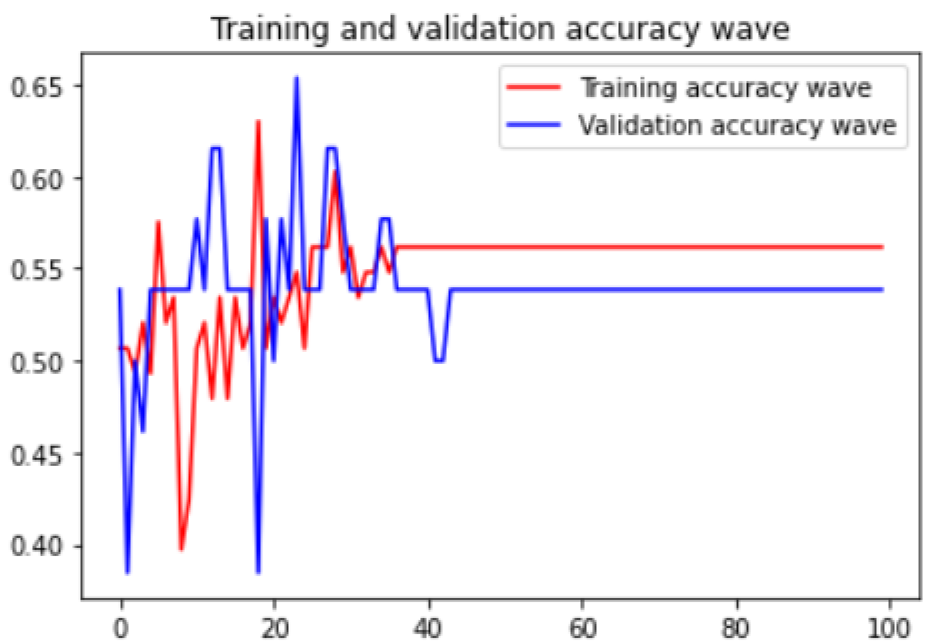


Fig 6: Training and validation accuracy wave in CNN model

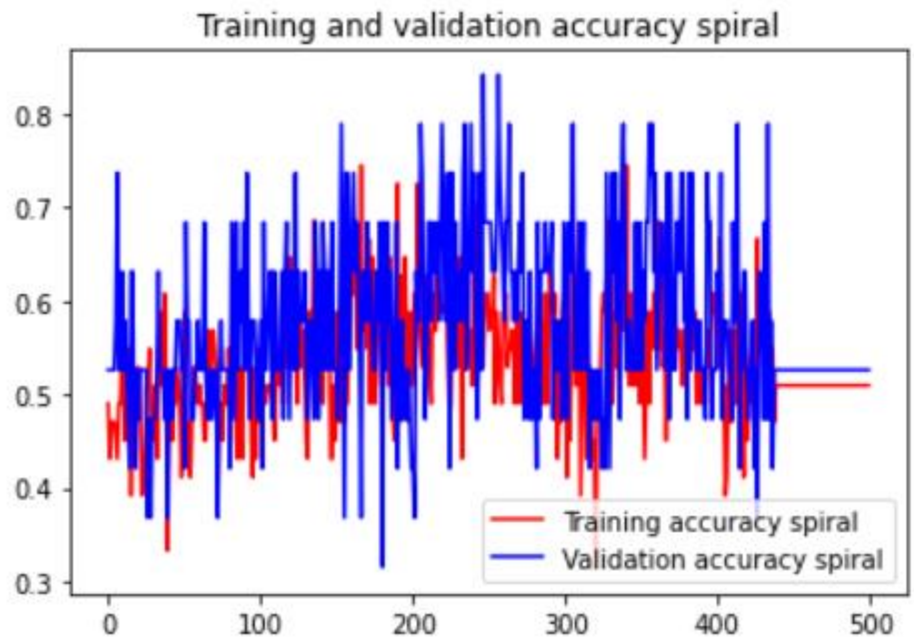


Fig 7: Training and Validation accuracy Spiral in CNN model

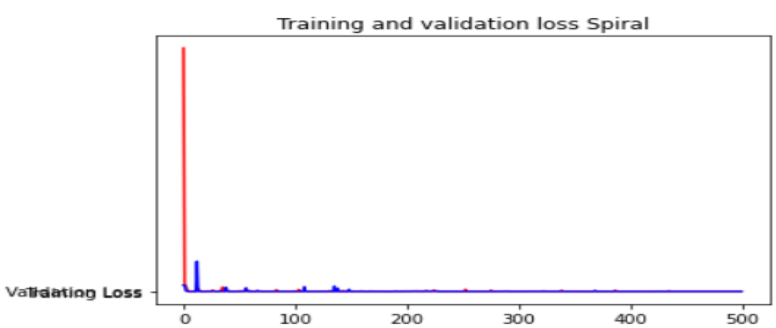


Figure 8: Training and validation loss spiral in CNN model

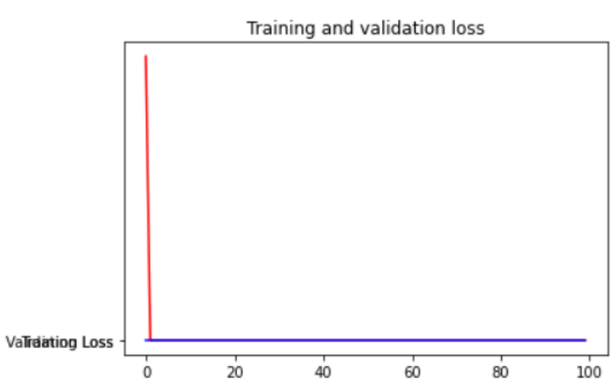


Fig 9: Training and validation loss wave in CNN model.

The Parkinson's Disease Detection System has been developed on a web application. The graphical user interface of the application was developed using Visual studio code and was programmed using HTML5, CSS, and JavaScript.
As shown in figure 10 below, the result of the model is displayed in a pop-up that shows the probability of patient having Parkinson's disease. The probability of having the disease is 24.40%, meaning that the chance or possibility of having PD is still very low.

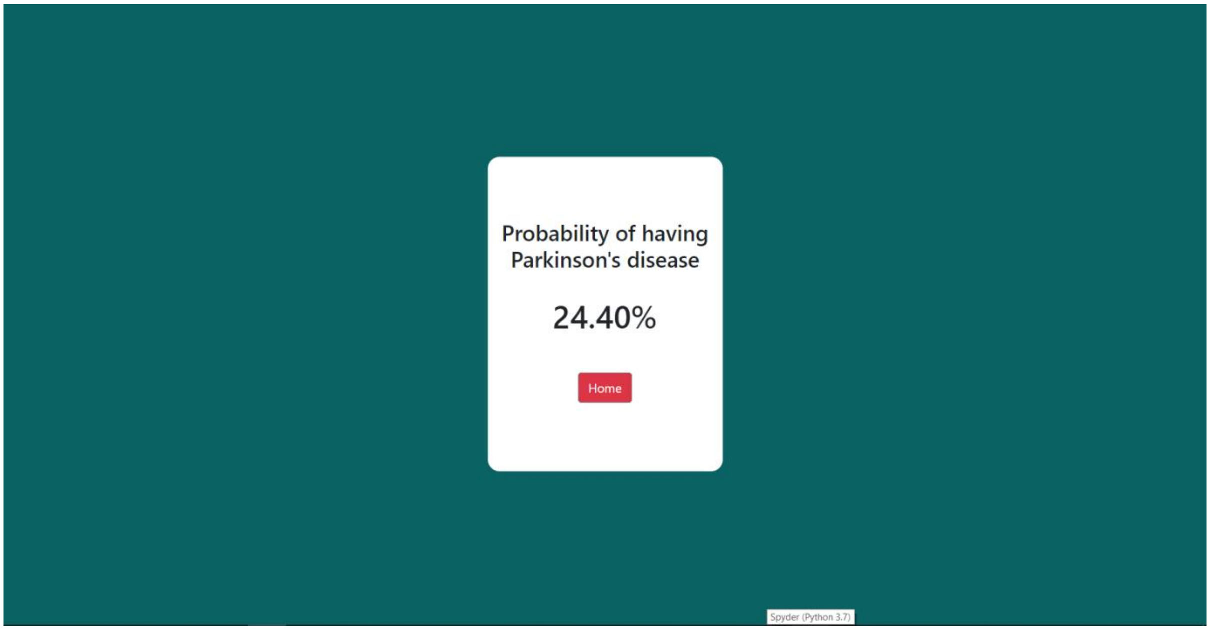


Figure 10: Prediction Result pop-up page

5. SUMMARY AND CONCLUSION

In this research work, various machine learning models were used to classify Parkinson's disease. It was established that FT is the best of all the five models employed in this research work.
Based on this, a system was developed and implemented based on the ML models. The system can detect Parkinson's disease symptoms at an early stage. It also displays the percentage level of the disease found in the patients. The system was tested and found to be efficient. The system will contribute immensely to the early detection and management of Parkinson's disease. Though there so many existing methods or models for detecting Parkinson disease in literature but this research work did not stop at modeling level but developed a system that can be used to detect Parkinson diseases.

6. ACKNOWLEDMENTS

Our thanks to Bowen University for providing us the support that contributed towards success of this work.

7. REFERENCES

[1] D. O. Olanloye, "Development of an Artificial Intelligence Geoinformatics System for Solid Minerals Prospecting.," *A Dissertation Submitted to the Department of Computer Science, Faculty of Physical Sciences, Nnamdi Azikiwe University, Awka.* -, 2017.

https://www.scirp.org/%28S%28lz5mqp453edsnp55rrgjct55%29%29/reference/referencespapers.aspx?referenceid=2974449 (accessed Jan. 11, 2022).

[2] H. Aworinde and O. F. W. Onifade, "A Soft Computing Model of Soft Biometric Traits for Gender and Ethnicity Classification," *international Journal of Engineering and Manufacturing*, vol. 9, no. 2, pp. 54–63, 2019, doi: 10.5815/ijem.2019.02.05.

[3] S. Adebayo, A. O. Akinwunmi, H. O. Aworinde, and E. O. Ogunti, "Increasing Agricultural Productivity in Nigeria Using Wireless Sensor Network (WSN)," *African Journal of Computing & ICT African Journal of Computing & ICT Reference Format: Afri J Comp & ICTs*, vol. 8, no. 2, pp. 121–128, 2015.

[4] E. O. Ogunti, F. K. Akingbade, S. Adebayo, and O. Oloyede, "Decision support system using mobile applications in the provision of day to day information about farm status to improve crop yield," *Periodicals of Engineering and Natural Sciences*, vol. 6, no. 2, pp. 89–99, 2018, doi: 10.21533/pen.v6i2.183.

[5] H. O. Aworinde, A. O. Afolabi, A. S. Falohun, and O. T. Adedeji, "Performance Evaluation of Feature Extraction Techniques in Multi-Layer Based Fingerprint Ethnicity Recognition System," *Asian Journal of Research in Computer Science*, vol. 3, no. 1, pp. 1–9, 2019, doi: 10.9734/AJRCOS/2019/v3i130084.

[6] B. Latinwo, A. Falohun, E. Omidiora, and B. Makinde, "Iris Texture Analysis for Ethnicity Classification Using Self-Organizing Feature Maps," *Journal of Advances in Mathematics and Computer Science*, vol. 25, no. 6, pp. 1–10, 2018, doi: 10.9734/jamcs/2017/29634.

[7] F. O. Alamu, H. O. Aworinde, and O. J. Oparah, "No Title," *Journal of Multidisciplinary Engineering Science and Technology (JMEST)*, vol. 1, no. 3, pp. 69–76, 2014.

[8] O. Friefeld, "Machine Learning in Computer Vision: Elements of Information Theory," *Computer Science, Ben-Gurion University*, 2018. http://webcache.googleusercontent.com/search?q=cache:https://www.cs.bgu.ac.il/~mlcv182/wiki.files/MLCV182_lec_IT.pdf (accessed Jan. 11, 2020).

[9] Z. Karapinar Senturk, "Early diagnosis of Parkinson's disease using machine learning algorithms," *Medical hypotheses*, vol. 138, May 2020, doi: 10.1016/J.MEHY.2020.109603.

[10] E. J. Daphne, "Parkinson's Disease Detection And Classification Using Machine Learning And Deep Learning Algorithms-A Survey," pp. 56–63, 2018.

[11] "Diagnosis of Parkinson's Disorder through Speech Data using Machine Learning Algorithms", doi: 10.35940/ijitee.C8060.019320.

[12] S. Aich, H. C. Kim, K. Younga, K. L. Hui, A. A. Al-Absi, and M. Sain, "A Supervised Machine Learning Approach using Different Feature Selection Techniques on Voice Datasets for Prediction of Parkinson's Disease," *International Conference on Advanced Communication Technology, ICACT*, vol. 2019-February, pp. 1116–1121, Apr. 2019, doi: 10.23919/ICACT.2019.8701961.

[13] S. Grover, S. Bhartia, Akshama, A. Yadav, and K. R. Seeja, "Predicting Severity Of Parkinson's Disease Using Deep Learning," *Procedia Computer Science*, vol. 132, pp. 1788–1794, Jan. 2018, doi: 10.1016/J.PROCS.2018.05.154.

[14] F. Javed Mehedi Shamrat, M. Asaduzzaman, A. Sazzadur Rahman, R. Tariqul Hasan Tusher, and Z. Tasnim, "A Comparative Analysis Of Parkinson Disease Prediction Using Machine Learning Approaches," *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, vol. 8, no. 11, 2019.

[15] J. Mei, C. Desrosiers, and J. Frasnelli, "Machine Learning for the Diagnosis of Parkinson's Disease: A Review of Literature," *Frontiers in aging neuroscience*, vol. 13, May 2021, doi: 10.3389/FNAGI.2021.633752.

[16] J. Brownlee, "Supervised and Unsupervised Machine Learning Algorithms - Machine Learning Mastery," *Machine Learning Algorithms*, 2016.

[17] P. Thanh Noi and M. Kappas, "Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery," *Sensors (Basel, Switzerland)*, vol. 18, no. 1, Dec. 2017, doi: 10.3390/s18010018.

[18] R. Saha and A. Uk, "Classification of Parkinson's Disease Using MRI Data and Deep Classification of Parkinson's Disease Using MRI Data and Deep Learning Convolution Neural Networks Learning Convolution Neural Networks CORE View metadata, citation and similar papers at core," 2019.

[19] M. S. Alzubaidi *et al.*, "The Role of Neural Network for the Detection of Parkinson's Disease: A Scoping Review," *Healthcare*, vol. 9, no. 6, Jun. 2021, doi: 10.3390/HEALTHCARE9060740.

[20] M. S. Abhishek, C. R. Chethan, C. R. Aditya, D. Divitha, and T. R. Nagaraju, "Diagnosis of Parkinson's Disorder through Speech Data using Machine Learning Algorithms," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 3, pp. 69–72, 2020, doi: 10.35940/ijitee.c8060.019320.

[21] J. Xu and M. Zhang, "Use of Magnetic Resonance Imaging and Artificial Intelligence in Studies of Diagnosis of Parkinson's Disease," *ACS chemical neuroscience*, vol. 10, no. 6, pp. 2658–2667, Jun. 2019, doi: 10.1021/ACSCHEMNEURO.9B00207.

[22] C. De Stefano *et al.*, "Feature Evaluation for Discriminating Handwriting Fragments," no. June, 2015.

[23] G. Singh and L. Samavedham, "Algorithm for image-based biomarker detection for differential diagnosis of Parkinson's disease," 2015.

[24] S. S. Tirumala, "Deep Learning Approaches," in *Deep Learning : Fundamentals, Methods and Application*, 2016.

[25] A. A. Amri, A. R. Ismail, and A. A. Zarir, "Comparative performance of deep learning and machine learning algorithms on imbalanced handwritten data," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 2, pp. 258–264, 2018, doi: 10.14569/IJACSA.2018.090236.