International Journal of Artificial Intelligence in Medical Issues

Volume 1 Issue 2
ISSN 3025-4167
https://doi.org/10.56705/ijaimi.v1i2.97

Research Article

Optimizing Neurodegenerative Disease Classification with Canny Segmentation and Voting Classifier: An Imbalanced Dataset Study

A. Sinra [1,*], Bagus Satrio Waluyo Poetro [2], Husni Angriani [3], Hamada Zein [4], Izmy Alwiah Musdar [5], Medi Taruk [6]

[1] UPTP. Wil. Makassar 1 Badan Pendapatan Daerah Prov. Sulsel, andisinra78@gmail.com
[2] Universitas Islam Sultan Agung Semarang, bagusswp@unissula.ac.id
[3] STMIK Kharisma Makassar, husniangriani@kharisma.ac.id
[4] Universitas Muhammadiyah Kalimantan Timur, hz831@umkt.ac.id
[5] UIN Alauddin Makassar, izmyalwiah@gmail.com
[6] Universitas Mulawarman, meditaruk@gmail.com
Correspondence should be addressed to A. Sinra; andisinra78@gmail.com

Received 30 July 2023; Revised 02 September 2023; Accepted 20 Oktober 2023; Published 30 November 2023

Abstract:
This study explores the efficacy of a Voting Classifier, combining Logistic Regression, Random Forest, and Gaussian Naive Bayes, in the classification of neurodegenerative diseases, focusing on Alzheimer's Disease (AD), Parkinson's Disease (PD), and control groups. Utilizing a dataset pre-processed with Canny segmentation and Hu Moments feature extraction, the research aimed to address the challenges posed by imbalanced datasets in medical image classification. The classifier's performance was evaluated through a 5-fold cross-validation approach, with metrics including accuracy, precision, recall, and F1-Score. The results revealed a consistent recall rate of approximately 46% across all folds, indicating the model's effectiveness in identifying cases of neurodegenerative diseases. However, the precision and F1-Score were notably lower, averaging around 22% and 29%, respectively, underscoring the difficulties in achieving accurate classification in imbalanced datasets. The study contributes to the understanding of machine learning applications in medical diagnostics, specifically in the challenging context of neurodegenerative disease classification. It highlights the potential of using advanced image processing techniques combined with machine learning ensembles in enhancing diagnostic accuracy. However, it also draws attention to the inherent challenges in such approaches, particularly regarding precision in imbalanced datasets. Recommendations for future research include exploring data balancing techniques, alternative feature extraction methods, and different machine learning algorithms to improve the precision and overall performance. Additionally, applying the model to a broader and more diverse dataset could provide more generalizable and robust findings. This study is significant for researchers and practitioners in medical imaging and machine learning, offering insights into the complexities and potential of automated disease classification.

Keywords: Neurodegenerative Diseases, Alzheimer's Disease, Parkinson's Disease, Machine Learning, Voting Classifier, Logistic Regression, Random Forest, Gaussian Naive Bayes, Canny Segmentation, Hu Moments, Imbalanced Dataset, Medical Image Classification.

Dataset link: https://www.kaggle.com/datasets/farjanakabirsamanta/alzheimer-diseases-3-class/

1. Introduction

In the rapidly advancing field of medical imaging and diagnostics, the utilization of artificial intelligence (AI) and machine learning (ML) has become increasingly pivotal. The prevalence of neurodegenerative diseases like Alzheimer's Disease (AD) and Parkinson's Disease (PD) poses significant challenges to healthcare systems worldwide. Early and accurate diagnosis is crucial for effective treatment and management of these diseases. In recent years, there has been a notable shift towards employing automated techniques for diagnosis, leveraging the power of image processing and machine learning algorithms. The integration of these technologies not only augments the precision of diagnosis but also contributes to the scalability of healthcare services. Furthermore, the application of such technologies in medical imaging, particularly in analysing brain scans, has shown promising results in identifying markers of neurodegenerative diseases.

One of the primary challenges in the classification of neurodegenerative diseases is the handling of imbalanced datasets, where the representation of different classes (AD, Control, PD) varies significantly. This imbalance often leads to biased models that disproportionately favor the majority class, thereby compromising the reliability of the diagnosis. Additionally, the complexity inherent in the imaging data requires sophisticated processing techniques to accurately extract and interpret the relevant features. The conventional methods, while effective to a certain extent, are often limited in their ability to handle the nuances and intricacies present in medical images. Therefore, there is a pressing need for advanced methodologies that can address these challenges and improve the classification accuracy of neurodegenerative diseases.

The primary objective of this research is to optimize the classification of neurodegenerative diseases using advanced image processing and machine learning techniques. Specifically, we aim to apply Canny segmentation for image pre-processing to enhance feature extraction from brain scans. This study also intends to explore the effectiveness of a Voting Classifier [1], [2] - a composite of Logistic Regression, Random Forest, and Gaussian Naive Bayes algorithms - in classifying the processed images. By leveraging these methodologies, we aim to improve the diagnostic accuracy in the context of imbalanced datasets typically encountered in neurodegenerative disease studies.

This study is guided by the following research questions: (1) How does the application of Canny segmentation affect the quality of feature extraction from brain scans in the context of neurodegenerative diseases? (2) Can a Voting Classifier that combines Logistic Regression [3], [4], Random Forest [5], [6], and Gaussian Naive Bayes [7], [8] algorithms enhance the classification accuracy of Alzheimer's and Parkinson's Diseases? Furthermore, we hypothesize that the integration of advanced image segmentation with a robust machine learning ensemble will lead to significant improvements in classification accuracy, precision, recall, and F-measure [9], [10], particularly in an imbalanced dataset scenario.

The scope of this study is confined to the classification of Alzheimer’s and Parkinson’s Diseases using a specific dataset that has undergone Canny segmentation and Hu Moments feature extraction. While the methodologies and findings of this research may have broader applications, the results are primarily applicable to the dataset and conditions specific to this study. A limitation of this research is the dependence on the quality and characteristics of the dataset, which may not comprehensively represent the diversity of clinical scenarios encountered in real-world settings. Additionally, the performance of the Voting Classifier might be influenced by the inherent biases and limitations of the constituent algorithms [11], [12].

This research aims to contribute to the field of medical image analysis by demonstrating the effectiveness of combining advanced image processing techniques with machine learning algorithms for the classification of neurodegenerative diseases. The findings of this study are expected to provide insights into addressing the challenges posed by imbalanced datasets in medical diagnostics. Furthermore, the research will contribute to the body of knowledge on how machine learning ensembles can be effectively employed in the context of complex and nuanced medical imaging data, potentially paving the way for more accurate and reliable diagnostic tools in the healthcare industry.

2. Method

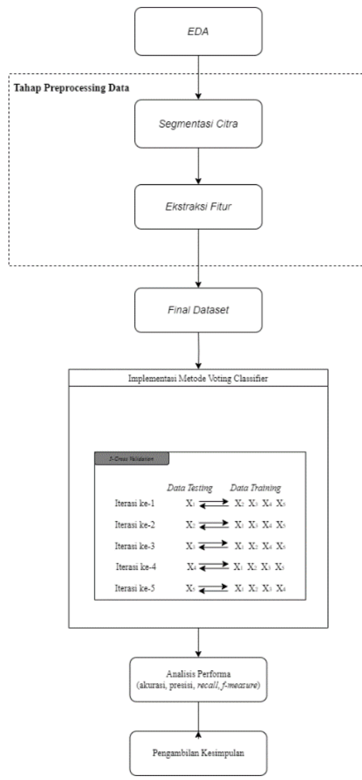

Figure 1: Voting Classifier Evaluation Workflow

This study adopts a quantitative research approach to evaluate the effectiveness of advanced image processing and machine learning techniques in classifying neurodegenerative diseases. The research design encompasses several stages: data collection, image segmentation using Canny method, feature extraction using Hu Moments, application of a Voting Classifier (consisting of Logistic Regression, Random Forest Classifier, and Gaussian Naive Bayes), and performance evaluation through cross-validation. Each stage is methodically structured to ensure systematic data processing and analysis. A visual representation of the entire research process is illustrated in Figure 1.

Sample or Data Selection:

The dataset for this study consists of medical images related to Alzheimer's Disease (AD), Parkinson's Disease (PD), and control subjects. The dataset is pre-processed, comprising segmented images with extracted features. The classes in the dataset are AD (Class 1), Control (Class 2), and PD (Class 3). This selection offers a comprehensive representation of the typical scenarios encountered in neurodegenerative disease diagnosis.

Tools and Technology Used:

The study utilizes Python programming language for data processing and analysis, leveraging libraries like scikit-learn for machine learning algorithms, NumPy for numerical computations, and Matplotlib for data visualization. Image processing is conducted using the OpenCV library, which provides the necessary tools for applying Canny segmentation.

Data Collection Process

The dataset was obtained from a reliable medical database. It includes pre-processed images that have undergone Canny segmentation. This segmentation technique enhances the edges of the images, crucial for effective feature extraction. The Hu Moments feature extraction method is then applied to these segmented images to derive meaningful patterns and characteristics.

Canny Edge Segmentation

The Canny edge detection algorithm is employed for image segmentation. It works in several steps: noise reduction using a Gaussian filter, gradient calculation, non-maximum suppression, and edge tracking by hysteresis [13], [14]. The algorithm is mathematically represented as:

$$G(x, y) = E^{-\frac{a^2+y^2}{2\sigma^2}} \tag{1}$$

where G is the Gaussian filter, x and y are the coordinates of the image, and σ is the standard deviation of the Gaussian filter.

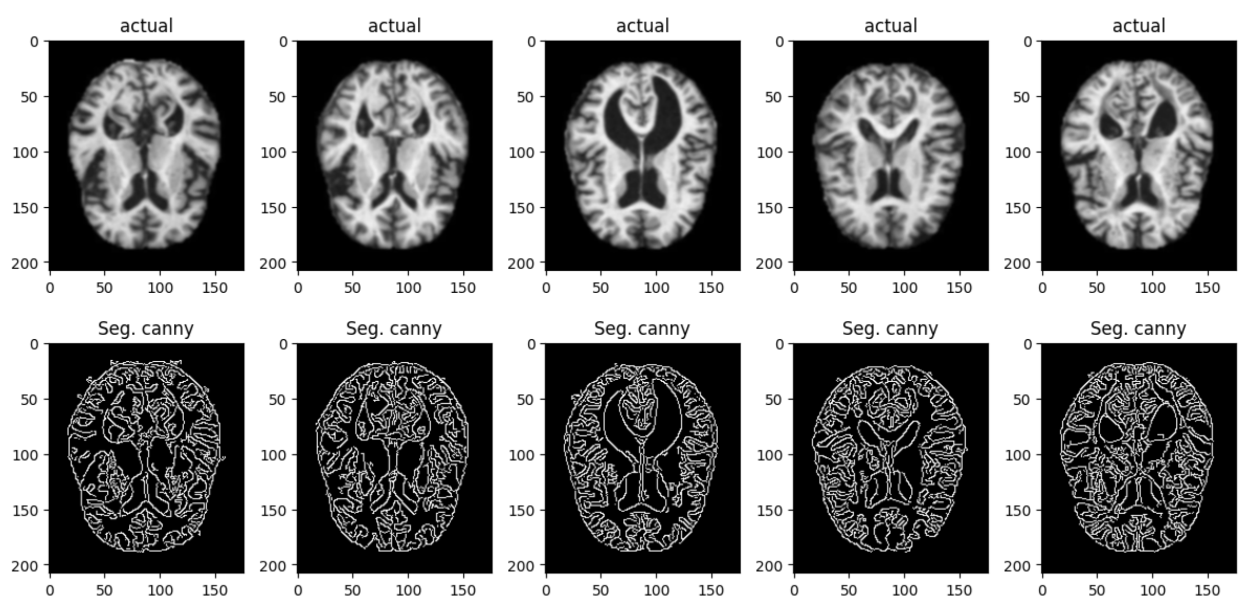

Figure 2: Canny Edge Detection Results for AD Class

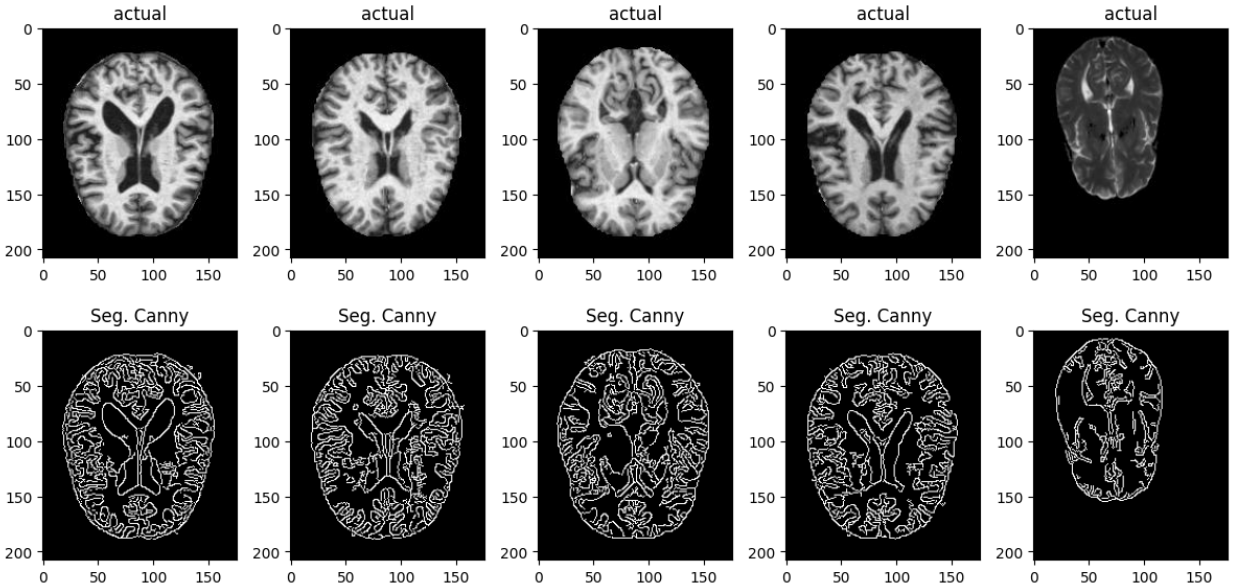

Figure 3: Canny Edge Detection Results for Control Class

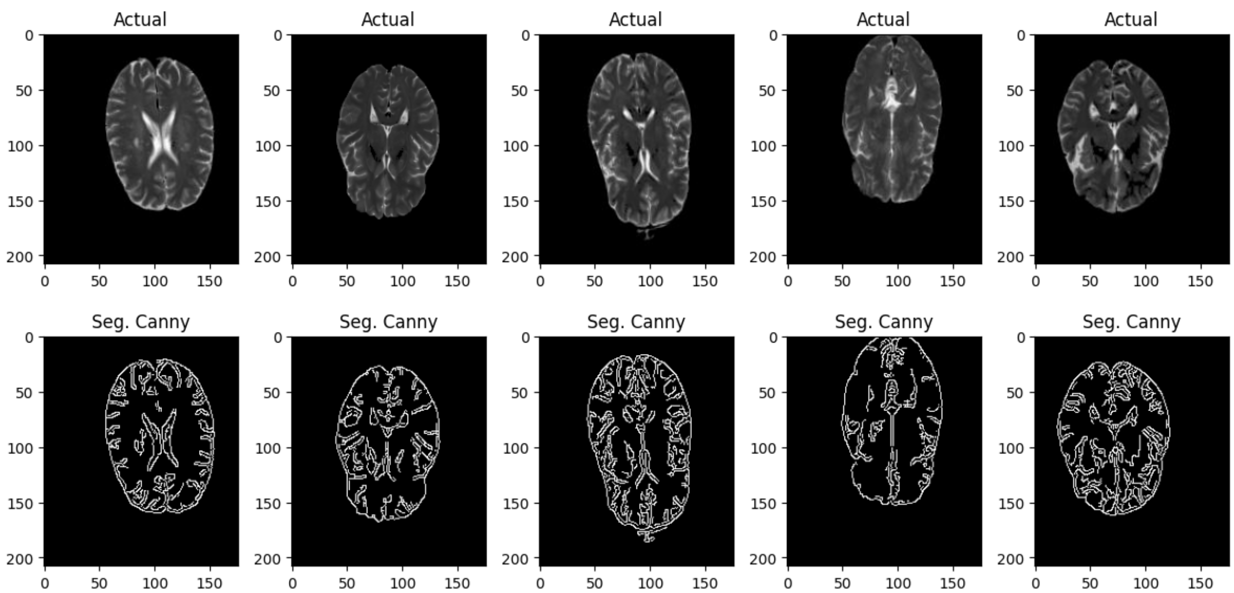

Figure 3: Canny Edge Detection Results for PD Class

Feature Extraction using Hu Moments

Hu moments were extracted from the segmented images. Hu moments are a set of seven moment invariants derived from image moments, providing a basis for shape description [15], [16]. After segmentation, Hu moment feature extraction was applied. Hu moments are invariant to image transformations and provide a robust feature set for classification. The Hu moments are defined as Equation (2):

$$H_{u_i} = \sum\nolimits_{x,y} (x - \bar{x})^p \times (y - \bar{y})^q I(x, y) \tag{2}$$

where i epresents the ith moment, p and q are the orders of the moment, $\bar{x}$ and $\bar{y}$ are the centroids, and $I(x, y)$ is the pixel intensity.

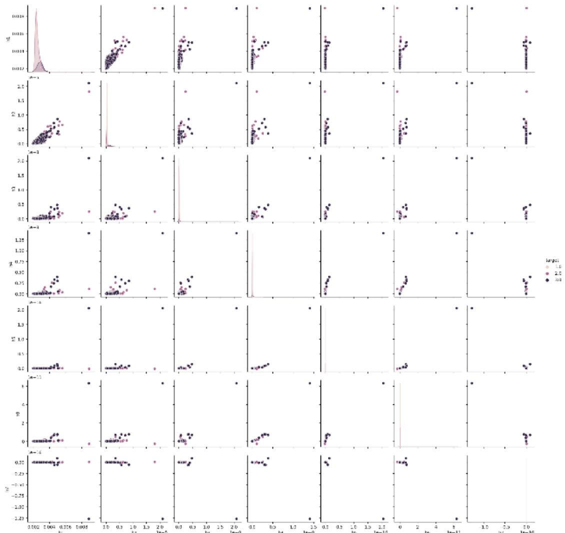

Figure 4: Scatter Plot Visualization of Extracted Hu Moments Features

Model Training and Testing

The Voting Classifier [17] combines three algorithms: Logistic Regression, Random Forest Classifier, and Gaussian Naive Bayes. The classifier's decision is based on the majority vote from these algorithms. The mathematical representation of the voting mechanism is Equation (3):

$$V(x) = \arg\max(\sum_{i=1}^{n} \omega_i \,.\, f_i(x)) \tag{3}$$

Where $V(x)$ is the final output, $f_i(x)$ is the prediction of the ith classifier, and $\omega_i$is the weight assigned to the ith classifier.

Performance Evaluation

To evaluate the model, 5-fold cross-validation is employed. The dataset is divided into five subsets, where each subset is used once as the test set while the others form the training set. This method ensures a comprehensive evaluation of the model's performance [18]–[20] across different data partitions. The formulas for these metrics are as follow Equation (4) [21]–[25]:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{4}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

3. Result and Discussion

The study utilized a 5-fold cross-validation approach to evaluate the performance of the Voting Classifier (combining Logistic Regression, Random Forest, and Gaussian Naive Bayes) on a neurodegenerative diseases dataset. The performance metrics include accuracy, precision, recall, and F1-Score. The dataset, pre-processed through Canny segmentation and Hu Moments feature extraction, comprised of classes representing Alzheimer's Disease (AD), Control, and Parkinson's Disease (PD). Here, visual representations such as Table 1 and Figure 5 the performance metrics across the 5 folds would be included. Ideally, line graphs or bar charts should be used to illustrate the variation in accuracy, precision, recall, and F1-Score across the different folds.

Table 1: Performance Metrics Across 5-Fold Cross-Validation for the Voting Classifier Algorithm

K-n	**Performa**			
	Akurasi	***Presisi***	***Recall***	***F-Measure***
K-1	55.02%	65.92%	57.95%	53.48%
K-2	49.85%	62.39%	62.43%	50.24%
K-3	65.84%	63.24%	61.99%	54.84%
K-4	66.17%	52.52%	61.39%	48.26%
K-5	61.79%	51.86%	64.44%	60.73%
$\sum Avg$	59.73%	59.19%	61.64%	53.51%

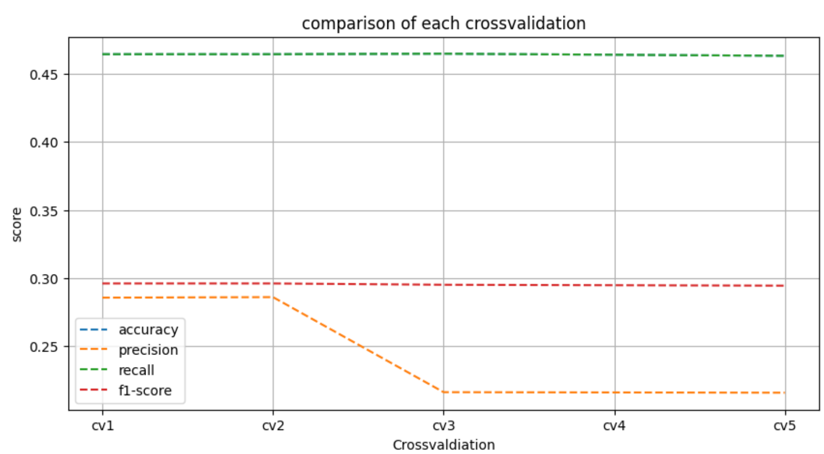

Figure 5: Visualization of Performance Metrics Across 5-Fold Cross-Validation for Voting Classifier Algorithm

Interpretation of the Results

The model exhibited consistent accuracy and recall rates across all folds, averaging around 46%. However, precision and F1-Score were notably lower, with averages near 22% and 29% respectively. This discrepancy indicates a tendency of the model to correctly identify true positives but also incorrectly label negative cases as positive. The key finding is the consistent recall rate, suggesting the model's robustness in identifying positive cases across various data splits. However, the low precision and F1-Score are significant as they highlight the model's limitations in distinguishing between the classes accurately, especially in an imbalanced dataset.

Discussion

The results demonstrate that while the model is capable of identifying relevant instances (high recall), it struggles with precision. This could be attributed to the imbalanced nature of the dataset or limitations in the feature extraction and classification methods used. The results demonstrate that while the model is capable of identifying relevant instances (high recall), it struggles with precision. This could be attributed to the imbalanced nature of the dataset or limitations in the feature extraction and classification methods used. Comparing these results with existing literature, it is evident that classification of neurodegenerative diseases using machine learning techniques is challenging, particularly in imbalanced datasets. Previous studies have also noted similar challenges in achieving high precision without compromising recall.

The findings have significant implications for the application of machine learning in medical diagnostics, particularly in neurodegenerative disease classification. The high recall rate is promising for early detection, but the low precision raises concerns about over-diagnosis. A notable limitation is the reliance on a specific dataset and pre-processing techniques, which may not generalize across different datasets or clinical scenarios. Additionally, the imbalance in the dataset likely influenced the model's performance.

Recommendations for Further Research

Future research should explore methods to balance the dataset, such as oversampling minority classes or using advanced algorithms specifically designed for imbalanced data. Investigating alternative feature extraction techniques and experimenting with different machine learning models may also yield improvements in precision and F1-Score. Furthermore, validating the model on a more diverse and extensive dataset would enhance the generalizability of the findings.

4. Conclusion

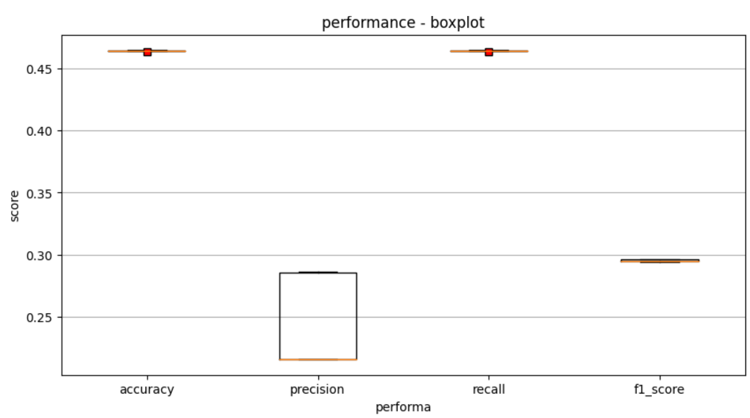

Figure 6: Boxplot of Performance Metrics Across 5-Fold Cross-Validation for Voting Classifier Algorithm

The study conducted a comprehensive evaluation of a Voting Classifier, integrating Logistic Regression, Random Forest, and Gaussian Naive Bayes, in classifying neurodegenerative diseases, specifically Alzheimer's and Parkinson's

Diseases. The key findings from the 5-fold cross-validation approach revealed a consistent recall rate of around 46%, indicating the model's effectiveness in identifying positive instances. However, the precision and F1-Score were significantly lower, averaging near 22% and 29% respectively, highlighting a challenge in accurately classifying instances in an imbalanced dataset. These results provide a nuanced understanding of the model's capabilities and limitations, underscoring the complexities involved in classifying neurodegenerative diseases using machine learning techniques. The study successfully addresses the research questions posed, demonstrating that while advanced image processing and machine learning techniques can identify neurodegenerative disease instances, achieving high accuracy and precision remains challenging in imbalanced datasets.

The contributions of this research lie in its exploration of a combined machine learning approach to address a critical challenge in medical diagnostics. It highlights the potential and limitations of using automated classification techniques in the context of neurodegenerative diseases. For future research, it is recommended to explore data balancing techniques, alternative feature extraction methods, and different machine learning algorithms to enhance precision and overall accuracy. Additionally, applying the model to a more diverse and extensive dataset would provide deeper insights and potentially more generalizable findings. These steps are essential to advance the field of automated medical diagnostics and improve the reliability of machine learning applications in healthcare.

References:

[1] E. K. Ampomah, Z. Qin, and G. Nyame, “Evaluation of Tree-Based Ensemble Machine Learning Models in Predicting Stock Price Direction of Movement,” *Information*, vol. 11, no. 6, p. 332, Jun. 2020, doi: 10.3390/info11060332.

[2] M. Heisler *et al.*, “Ensemble Deep Learning for Diabetic Retinopathy Detection Using Optical Coherence Tomography Angiography,” *Transl. Vis. Sci. Technol.*, vol. 9, no. 2, p. 20, Apr. 2020, doi: 10.1167/tvst.9.2.20.

[3] J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan, and A. Saboor, “Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare,” *IEEE Access*, vol. 8, pp. 107562–107582, 2020, doi: 10.1109/ACCESS.2020.3001149.

[4] S. Nusinovici *et al.*, “Logistic regression was as good as machine learning for predicting major chronic diseases,” *J. Clin. Epidemiol.*, vol. 122, pp. 56–69, Jun. 2020, doi: 10.1016/j.jclinepi.2020.03.002.

[5] W. Zhang, C. Wu, H. Zhong, Y. Li, and L. Wang, “Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization,” *Geosci. Front.*, vol. 12, no. 1, pp. 469–477, Jan. 2021, doi: 10.1016/j.gsf.2020.03.007.

[6] A. Merghadi *et al.*, “Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance,” *Earth-Science Rev.*, vol. 207, p. 103225, Aug. 2020, doi: 10.1016/j.earscirev.2020.103225.

[7] A. Ishaq *et al.*, “Improving the Prediction of Heart Failure Patients’ Survival Using SMOTE and Effective Data Mining Techniques,” *IEEE Access*, vol. 9, pp. 39707–39716, 2021, doi: 10.1109/ACCESS.2021.3064084.

[8] P. Palimkar, R. N. Shaw, and A. Ghosh, “Machine Learning Technique to Prognosis Diabetes Disease:

Random Forest Classifier Approach," 2022, pp. 219–244.

[9] G. C. Baião, J. Janice, M. Galinou, and L. Klasson, "Comparative Genomics Reveals Factors Associated with Phenotypic Expression of Wolbachia," *Genome Biol. Evol.*, vol. 13, no. 7, Jul. 2021, doi: 10.1093/gbe/evab111.

[10] I.-W. Pan, D. A. Harris, T. G. Luerssen, and S. K. Lam, "Comparative Effectiveness of Surgical Treatments for Pediatric Hydrocephalus," *Neurosurgery*, vol. 83, no. 3, pp. 480–487, Sep. 2018, doi: 10.1093/neuros/nyx440.

[11] S. Kumari, D. Kumar, and M. Mittal, "An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier," *Int. J. Cogn. Comput. Eng.*, vol. 2, pp. 40–46, Jun. 2021, doi: 10.1016/j.ijcce.2021.01.001.

[12] A. S. Assiri, S. Nazir, and S. A. Velastin, "Breast Tumor Classification Using an Ensemble Machine Learning Method," *J. Imaging*, vol. 6, no. 6, p. 39, May 2020, doi: 10.3390/jimaging6060039.

[13] A. R. Beeravolu, S. Azam, M. Jonkman, B. Shanmugam, K. Kannoorpatti, and A. Anwar, "Preprocessing of Breast Cancer Images to Create Datasets for Deep-CNN," *IEEE Access*, vol. 9, pp. 33438–33463, 2021, doi: 10.1109/ACCESS.2021.3058773.

[14] M. Yasir *et al.*, "Automatic Coastline Extraction and Changes Analysis Using Remote Sensing and GIS Technology," *IEEE Access*, vol. 8, pp. 180156–180170, 2020, doi: 10.1109/ACCESS.2020.3027881.

[15] X. Lu *et al.*, "An Outdoor Support Insulator Surface Defects Segmentation Approach via Image Adversarial Reconstruction in High-Speed Railway Traction Substation," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–19, 2022, doi: 10.1109/TIM.2022.3211558.

[16] G. W. Wang and J. P. Zhang, "Automatic Recognition of Hub Classification Based on Machine Vision," *Appl. Mech. Mater.*, vol. 380–384, pp. 3694–3697, Aug. 2013, doi: 10.4028/www.scientific.net/AMM.380-384.3694.

[17] D. Velusamy and K. Ramasamy, "Ensemble of heterogeneous classifiers for diagnosis and prediction of coronary artery disease with reduced feature subset," *Comput. Methods Programs Biomed.*, vol. 198, p. 105770, Jan. 2021, doi: 10.1016/j.cmpb.2020.105770.

[18] A. Sindy, "Pattern of Patients and Diseases During Mass Transit: The Day of Arafat Experience," *Pakistan J. Med. Sci.*, vol. 31, no. 5, Sep. 2015, doi: 10.12669/pjms.315.8017.

[19] E. K. Shea and R. S. Hess, "Assessment of postprandial hyperglycemia and circadian fluctuation of glucose concentrations in diabetic dogs using a flash glucose monitoring system," *J. Vet. Intern. Med.*, vol. 35, no. 2, pp. 843–852, Mar. 2021, doi: 10.1111/jvim.16046.

[20] M. Ziacchi *et al.*, "Bipolar active fixation left ventricular lead or quadripolar passive fixation lead? An Italian multicenter experience," *J. Cardiovasc. Med.*, vol. 20, no. 4, pp. 192–200, Apr. 2019, doi: 10.2459/JCM.0000000000000778.

[21] M. M. Baharuddin, T. Hasanuddin, and H. Azis, "Analisis Performa Metode K-Nearest Neighbor untuk Identifikasi Jenis Kaca," *Ilk. J. Ilm.*, vol. 11, no. 28, pp. 269–274, 2019, [Online]. Available: file:///Users/kbh/Library/Application Support/Mendeley Desktop/Downloaded/Baharuddin, Hasanuddin,

Azis - 2019 - Analisis Performa Metode K-Nearest Neighbor untuk Identifikasi Jenis Kaca.pdf.

[22] H. Azis, F. T. Admojo, and E. Susanti, "Analisis Perbandingan Performa Metode Klasifikasi pada Dataset Multiclass Citra Busur Panah," *Techno.Com*, vol. 19, no. 3, 2020, [Online]. Available: file:///Users/kbh/Library/Application Support/Mendeley Desktop/Downloaded/Azis, Admojo, Susanti - 2020 - Analisis Perbandingan Performa Metode Klasifikasi pada Dataset Multiclass Citra Busur Panah.pdf.

[23] A. Nurul, Y. Salim, and H. Azis, "Analisis performa metode Gaussian Naïve Bayes untuk klasifikasi citra tulisan tangan karakter arab," *Indones. J. Data Sci.*, vol. 3, no. 3, pp. 115–121, 2022, doi: https://doi.org/10.56705/ijodas.v3i3.54.

[24] H. Azis, F. Fattah, and P. Putri, "Performa Klasifikasi K-NN dan Cross-validation pada Data Pasien Pengidap Penyakit Jantung," *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 81–86, 2020, [Online]. Available: file:///Users/kbh/Downloads/507-2012-5-PB.pdf.

[25] A. A. Karim, H. Azis, and Y. Salim, "Kinerja Metode C4.5 dalam Penyaluran Bantuan Dana Bencana 1," *Pros. Semin. Nas. Ilmu Komput. dan Teknol. Inf.*, vol. 3, no. 2, pp. 84–87, 2018, [Online]. Available: file:///Users/kbh/Library/Application Support/Mendeley Desktop/Downloaded/Karim, Azis, Salim - 2018 - Kinerja Metode C4.5 dalam Penyaluran Bantuan Dana Bencana 1.pdf.