Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871

Vol. 6, No. 5, October 2025, Page. 3495-3508
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4726

Performance Comparison of AdaBoost, LightGBM, and CatBoost for Parkinson's Disease Classification Using ADASYN Balancing

Muhammad Ridha Anshari[*1], Triando Hamonangan Saragih[2], Muliadi Muliadi[3], Dwi Kartini[4], Fatma Indriani[5], Hasri Akbar Awal Rozaq[6], Oktay Yıldız[7]

[1,2,3,4,5]Faculty of Mathematics and Natural Science, Department of Computer Science, Lambung Mangkurat University, Kalimantan, Indonesia
[6]Graduate School of Informatics, Department of Computer Science, Gazi University, Ankara, Türkiye
[7]Faculty of Engineering, Department of Computer Engineering, Gazi University, Ankara, Türkiye

Email: [1] ridhoanshari.7b24@gmail.com

Received : May 13, 2025; Revised : Jul 4, 2025; Accepted : Jul 6, 2025; Published : Oct 16, 2025

Abstract

Parkinson's disease is a neurodegenerative condition identified by the decline of neurons that produce dopamine, causing motor symptoms such as tremors and muscle stiffness. Early diagnosis is challenging as there is no definitive laboratory test. This study aims to improve the accuracy of Parkinson's diagnosis using voice recordings with machine learning algorithms, such as AdaBoost, LightGBM, and CatBoost. The dataset used is Parkinson's Disease Detection from Kaggle, consisting of 195 records with 22 attributes. The data was normalized with Min-Max normalization, and class imbalance was resolved with ADASYN. Results show that ADASYN-LightGBM and ADASYN-CatBoost have the best performance with 96.92% accuracy, 97.10% precision, 96.92% recall, and 96.92% F1 score. This improvement suggests that combining boosting methods and data balancing techniques can improve the accuracy of Parkinson's diagnosis. These results demonstrate the effectiveness of ADASYN in addressing data imbalance and improving the performance of boosting algorithms for medical classification problems. The findings contribute to the development of intelligent diagnostic systems in the field of medical informatics and computer science. These findings are essential for developing more accurate and efficient diagnostic tools, supporting early diagnosis and better management of Parkinson's disease.

Keywords : *ADASYN, AdaBoost, CatBoost, LightGBM, Parkinson's Disease.*

1. INTRODUCTION

Parkinson's disease stands as a predominant neurodegenerative condition, distinguished by the progressive loss of dopamine-producing neurons, which are crucial for regulating body movements [1], [2]. This decline leads to the hallmark motor symptoms associated with the disease, such as bradykinesia, muscle stiffness, tremors, and postural instability [3], [4]. Beyond these, individuals may also encounter non-motor symptoms, including sleep disturbances, dementia, and sensory disruptions [5]. Given its prevalence, especially among older adults, and the challenges it presents in diagnosis and management, Parkinson's disease represents a significant area of concern within the medical community [6].

The diagnosis of Parkinson's traditionally relies on clinical evaluations, focusing on neurological history and motor function assessments. The absence of definitive laboratory tests makes early diagnosis challenging, underscoring the necessity for innovative diagnostic approaches [7]. However, the limited availability of large and diverse datasets for Parkinson's diagnosis poses challenges in developing robust and generalizable models. In this context, using voice recordings is a promising non-invasive diagnostic tool. When coupled with machine learning algorithms, specifically boosting methods, these recordings

Vol. 6, No. 5, October 2025, Page. 3495-3508
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4726

offer a practical preliminary screening option [8], [9]. The effectiveness of boosting methods in complex classification tasks has been well documented, illustrating their potential to enhance diagnostic accuracy [10].

Boosting methods, by design, aim to facilitate the accuracy of learning algorithms prone to overfitting. They achieve this by integrating multiple weak learners to form a robust algorithm capable of effective classification and prediction across various domains, including medical diagnostics [11], [12]. Among the numerous variants, Adaboost, LightGBM, and CatBoost have been identified as particularly potent, demonstrating superior performance in medical applications and remote sensing data classification [13]. Recent studies have shown promising results in applying these individual boosting methods to various medical diagnostic tasks [14], [15]. Nevertheless, the reliance on small datasets, such as the one used in this study with only 195 records, limits the broader applicability and generalizability of these methods. In a comparative analysis of six Ensemble Learning methods, XGBoost and LightGBM outperformed others, particularly with hyperspectral datasets. At the same time, XGBoost and RF (Random Forest) showed remarkable accuracy in PolSAR data classification, achieving 84.62% and 81.94% [8], respectively. This highlights the advanced capability of boosting methods like XGBoost and LightGBM to handle complex data scenarios. Further exploration of boosting methods on larger and more representative datasets is critical to validate their scalability and effectiveness.

However, the performance of these methods can be significantly influenced by the balance of the dataset used. In the realm of machine learning, imbalanced datasets commonly skew predictions towards the majority class, detrimentally affecting the accuracy for minority classes [16]–[18]. This challenge necessitates applying data balancing techniques, such as Adaptive Synthetic Sampling (ADASYN), which aims to equilibrate class distribution, enhancing the model's learning and predictive capabilities. While ADASYN has been successfully applied to various classification problems [19], [20], limited research has explored the comprehensive comparison of AdaBoost, LightGBM, and CatBoost combined with ADASYN specifically for Parkinson's disease classification using voice data [21]. Given this backdrop, the research aims to scrutinize and compare the efficacy of Adaboost, LightGBM, and CatBoost in classifying Parkinson's disease data, utilizing ADASYN for data balancing [22]. This study aims to evaluate the performance of AdaBoost, LightGBM, and CatBoost in classifying Parkinson's disease using voice data, particularly when balanced using ADASYN, and to determine the most effective combination for accurate diagnosis. This endeavor seeks not only to contribute new insights into the application of machine learning in early disease detection and diagnosis but also to pave the way for developing more precise and efficient diagnostic tools [23], [24]. The urgency for such advancements is underscored by the ongoing challenges in Parkinson's disease management and the potential impact of early, accurate diagnosis on improving patient outcomes [25].

2. METHOD

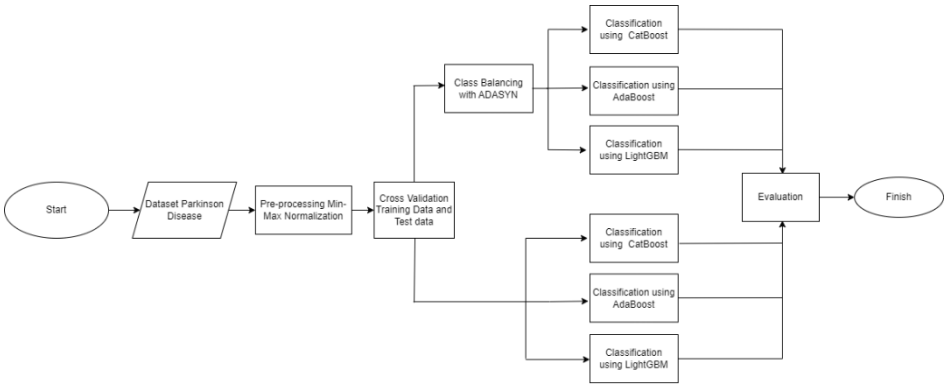

Figure 1. Flow of Research

Vol. 6, No. 5, October 2025, Page. 3495-3508
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4726

In this research, the data to be used will first go through a pre-processing stage, such as normalizing the data so that the distribution is the same for each column, separating the information into two (training data and test data), and utilizing 10-fold validation because it has the power to handle overfitting a model [26]. The last stage is to implement the data against the model to be used and evaluate it. Good research needs systematic stages. The stages of this study are shown in Figure 1.

2.1. Data Collection

This research uses the Parkinson's Disease Detection dataset downloaded from Kaggle. This dataset includes one target class to be predicted. It is available at the following link: https://www.kaggle.com/datasets/debasisdotcom/parkinson-disease-detection/data [27]. This Parkinson's Disease Detection dataset has 195 data records with 22 independent and one dependent attribute. Details concerning every characteristic for every attribute are displayed in Table 1.

Table 1. Dataset's Information

Features	Descriptions
name	The ASCII subject identifier and recording number
MDVP: Fo(Hz)	voice's mean fundamental frequency
MDVP: Fhi(Hz)	Voice's maximum fundamental frequency
MDVP: Flo(Hz)	Voice's minimum fundamental frequency
MDVP: Jitter(%)	Different metrics for basic frequency variability
MDVP:Jitter(Abs)	
MDVP: RAP	
MDVP: PPQ	
Jitter: DDP	
MDVP: Shimmer	Different amplitude variability metrics
MDVP: Shimmer(dB)	
Shimmer: APQ3	
Shimmer: APQ5	
MDVP: APQ	
Shimmer: DDA	
NHR	Noise to tonal components ratio in speech
HNR	
status	Subject's health status: (1) Parkinson's, (0) healthy
RPDE	Nonlinear dynamical complexity metrics
DFA	Exponent of fractal scaling for the signal
spread1	Nonlinear indicators of underlying frequency fluctuation
spread2	
D2	Nonlinear dynamical complexity metrics
PPE	Measures of fundamental frequency variations that are nonlinear

In determining subject status, the dependent variable in this study, it is essential to consider all available independent variables or features. This is due to the complexity and variety of symptoms associated with Parkinson's Disease. The features in question provide varying information about an individual's voice characteristics that may be associated with Parkinson's condition.

For example, some features such as MDVP: Fo(Hz) (Average vocal fundamental frequency), MDVP: Jitter(%) (Various measures of variation in fundamental frequency), and NHR (Ratio of noise to tonal components in the voice) can provide insight into changes in vocal frequency and voice stability

Vol. 6, No. 5, October 2025, Page. 3495-3508
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4726

that may have a relationship with the progression of the disease. On the other hand, features such as RPDE (Nonlinear dynamic complexity measurement) and DFA (Signal fractal expression) can reveal the complexity of the voice signal, which may be related to the characteristics of Parkinson's disease.

By incorporating all these features into the model, we allow the model to explore and understand the complex patterns and subtypes of Parkinson's Disease that may be present in the data. We need to take all these features into account to ensure we get all the essential information required to classify the status of the subjects accurately. Thus, the comprehensive use of all independent variables allows for the construction of models that are more effective in distinguishing the health status of subjects who have Parkinson's disease or are healthy, as can be seen in Figure 2.

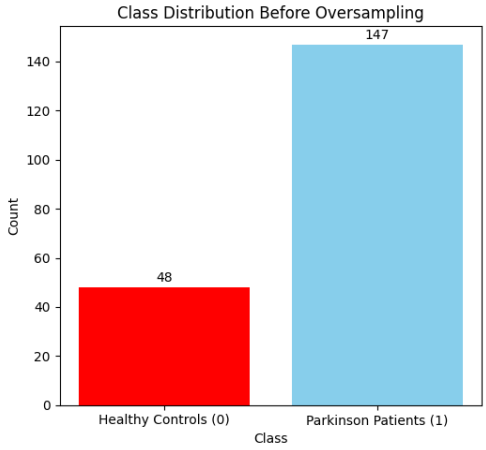

Figure 2. Class of Datasets

The independent variable “status” used in this study refers to the category or condition of the research subject. In this context, two statuses were observed: “Healthy Controls” (0) and “Parkinson Patients” (1). “Healthy Controls” refers to subjects who do not have Parkinson's disease, while “Parkinson Patients” refers to subjects diagnosed with Parkinson's disease. In this case, 48 subjects fell into the “Healthy Controls” category, and 147 subjects fell into the “Parkinson's Patients” category. The number of subjects in each category gives an idea of the population distribution observed in the study and allows for comparative analysis between the two groups.

2.2. Min-Max Normalization

Each data has different value characteristics. This is the case with the dataset used, which has variables with values in various ranges. This will impact the AI model because AI models work better if the data range is the same [28]. One way to equalize the range is to normalize the data. Min-Max Normalization was chosen for this study due to its simplicity and effectiveness in rescaling data to a range between 0 and 1, making it compatible with machine learning algorithms that are sensitive to scale differences. The normalization process also aims to minimize the memory used. Normalized data will have the same range between 0 and 1 [29]. However, the specific impact of Min-Max Normalization on model performance has not been evaluated against alternative methods, such as Z-score normalization, which could be explored in future studies. The following is the formula for the data normalization process, which uses the Min-Max normalization calculation [30].

$$Z = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

Where Z is the Min-Max normalization result, x is the original value, min(x) is the minimum value of the calculated attribute, and max(x) is the maximum value of the computed attribute.

Vol. 6, No. 5, October 2025, Page. 3495-3508
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4726

2.3. Cross Validation

The normalization process can improve the performance of an AI model. Still, another method can also enhance the performance of an AI model, which is choosing the best proportion of data for training and test data [31]. Previous studies have argued that proportions such as 90:10 [32], 80:20 [33], and 70:30 [34] are the best proportions for each model used in the study. It would be very time-consuming to compare them one by one. Therefore, we must use a method to decide which proportion is best for the AI model. One method that we can use is cross-validation.

Cross-validation is a method in which the initial dataset is divided into two parts: a subset for training and a subset for validation. The model is then trained using the training subset and tested on the validation subset [7]. The K-fold Cross-Validation method splits the dataset into K similar subsets or folds [8]. Here is an overview of this process.

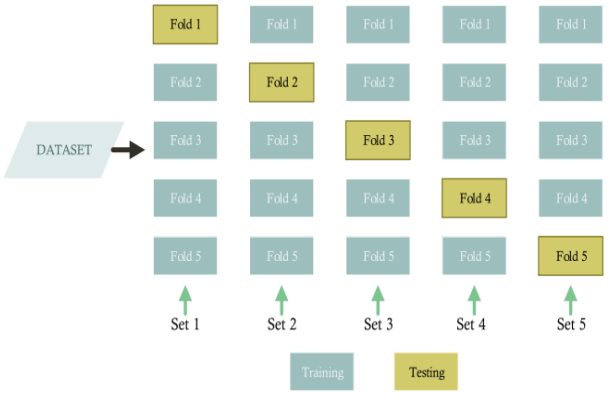

Figure 3. Cross-Validation Process

Figure 3 shows that each fold is partitioned for training and testing in K iterations. At each iteration, one fold is used for testing, while the other K-1 folds are used for model training. The most common K values are 5 and 10, but the results are sensitive to too-small K values. However, the computational process will increase when the K value is too significant. Model accuracy is calculated by taking the average accuracy achieved at each iteration.

2.4. Adaptive Synthetic Sampling

AI models will work optimally from several aspects, as mentioned earlier. However, one process also determines the model's performance: data with unbalanced labels [35]. It affects the AI model, which will predict the data inaccurately because it is covered by more labeled data [36]. The data visualization from this study shows that the labels between those affected by Parkinson's disease and those not affected are more representative of those affected by Parkinson's disease. This may assume that the model will always predict new data to be affected by Parkinson's disease. This answer is not desirable because it could be that the patient is not affected by the disease but is diagnosed with it. Therefore, we can solve the problem of unbalanced data by using the oversampling method.

There are two ways to overcome this, namely, oversampling and undersampling. The oversampling method here is chosen to balance the data that is not affected by the disease with at least the same amount as the data affected by the disease. The oversampling method has also been proven to improve the accuracy of an AI model. Adaptive synthetic sampling (ADASYN) is a well-known method for oversampling [37]. ADASYN was developed to adaptively generate synthetic samples in minority classes with high classification difficulty. As such, ADASYN focuses on improving the representation of minority classes by generating synthetic samples in regions that are underrepresented by those classes. This approach allows the classification model to pay more attention to cases that are difficult to predict

Vol. 6, No. 5, October 2025, Page. 3495-3508
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4726

[9]. ADASYN can be executed using two methods: reducing the bias caused by unbalanced classes or shifting the classification decision boundaries. The steps that illustrate how the ADASYN algorithm works have been presented in previous research [10]. The steps have six processes: calculating the level of data imbalance, calculating the amount of synthesized data, calculating the distance between the data, normalizing the distribution data, summing the synthesized data, and calculating the sample data.

2.5. AdaBoost

AdaBoost (Adaptive Boosting) is one boosting algorithm in the ensemble learning category that introduces a series of simple predictive models, such as weak decision trees, which are then adaptively combined to form a strong predictive model [12]. AdaBoost assigns a different weight to each training sample in each iteration based on the model's previous performance in predicting that sample, allowing for greater emphasis on complex samples. These adjusted weights train the next model, explicitly focusing on previously misclassified samples. The final weights of each model are used to combine the predictions of all the model components. AdaBoost and its variations have been successfully applied in various domains due to their strong theoretical basis, accurate prediction capabilities, and simplicity, making them easy to use and implement. Several classifiers (weak classifiers) are trained on the same training set, and these weak classifiers are then combined to generate a stronger final classifier (strong classifier), which is the central notion of the technique [11]. The steps of the AdaBoost algorithm have been described in previous research [13].

2.6. LightGBM

LightGBM (Light Gradient Boosting Machine) is a boosting algorithm developed by Microsoft Research. It implements gradient boosting that is specifically designed to provide extremely fast and efficient performance. LightGBM uses an ensemble learning approach that builds a robust set of predictive models by combining predictions from relatively simple decision tree models [17]. One of the advantages of LightGBM is its ability to handle massive datasets quickly and its efficiency in processing categorical data. LightGBM is also known for its ability to handle overfitting and exemplary performance in various machine-learning tasks [16].

The general formulas used in LightGBM may include those associated with gradient boosting, such as those used in the AdaBoost algorithm. However, since LightGBM uses a somewhat different approach to the traditional gradient boosting method, the specific formulas associated with this algorithm may differ [18]. Possible formulas associated with LightGBM include optimization steps, gradient calculations to minimize the loss function, and special techniques used in the model training process. In LightGBM, histogram-based algorithms and tree growth strategies with maximum depth constraints are applied to improve training efficiency and reduce memory usage. The histogram-based algorithm in the decision tree is shown in Figure 4 [22]. It can be seen that the continuous eigenvalues are converted into discrete values that are grouped into S small bins. After passing the initialization stage, the required statistics, such as the number of gradients and samples in each bin, are accumulated in the histogram. The optimal segment can be determined based on the histogram's discrete values.

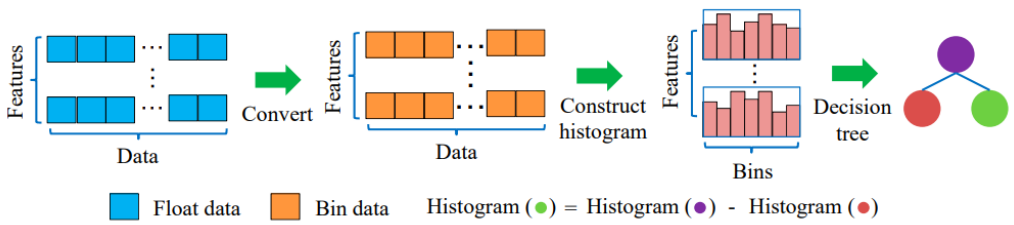

Figure 4.LightGBM Process

Vol. 6, No. 5, October 2025, Page. 3495-3508
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4726

LightGBM distinguishes itself from other tree-boosting algorithms with a leaf-wise tree growth approach that selects the most optimal split, unlike other tree-boosting algorithms that tend to split the tree in depth or alignment (level-wise tree growth), as shown in Figure 5. When growing on the same leaf in LightGBM, the leaf-wise approach can reduce losses and yield better accuracy than the level-wise approach and other boosting algorithms. However, it should be noted that the leaf-wise approach tends to be more prone to overfitting.

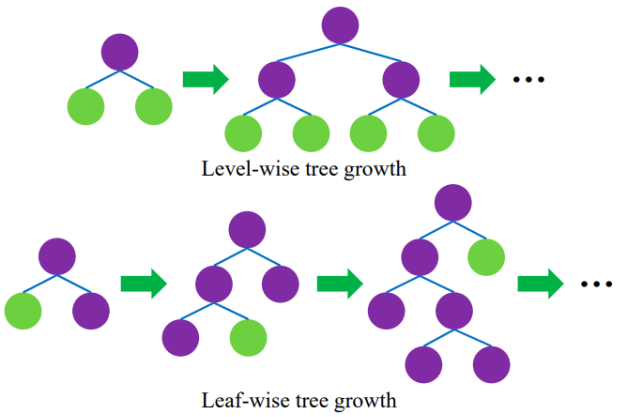

Figure 5. Leaf-Wise Tree Growth

2.7. CatBoost

A boosting technique is used by the open-source machine-learning package CatBoost. This algorithm performs better when dealing with categorical features. In addition, to lessen the chance of overfitting, CatBoost can handle classification features more logically and effectively [23]. CatBoost has better performance and shorter running time than XGBoost and LightGBM algorithms. CatBoost differs from other gradient boosting algorithms; it uses ordered boosting, an effective way to solve the target leakage issue with the gradient boosting algorithm [24]. CatBoost uses a symmetric tree commonly called an Oblivious Decision Tree (ODT) as a weak learner. Symmetric trees are balanced, not prone to overfitting, and can significantly speed up test execution time. CatBoost forms a group of ODTs with full binary trees. The resulting tree structure is always symmetric, and the same splitting criteria are used on the same tree layers. Symmetrical trees can predict about ten times faster than asymmetric trees [22].

2.8. Evaluation

The success of an AI model that has been built can predict new data well and has high accuracy, which is the primary goal of this research. So, we need to evaluate the AI model that has been built. Model evaluation depends on what problem is being solved and whether the data used is classification, regression, or clustering data. In this research, the problem being solved is the classification of data. So, the AI model evaluation process can use calculations such as accuracy, precision, and recall through confusion matrix calculations.

The confusion matrix is an evaluation method used in classification to evaluate the performance of a model by comparing the predicted class produced by the model with the actual class of the data [38], [39]. A confusion matrix provides a more detailed picture of the performance of a classification model than a single metric, such as accuracy, because it considers the model's ability to classify each class separately. In the context of Parkinson’s disease classification, the confusion matrix metrics have real-world implications. For instance, True Positive (TP) indicates correctly identified patients with Parkinson's, while False Negative (FN) represents missed diagnoses, which could delay necessary

Vol. 6, No. 5, October 2025, Page. 3495-3508
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4726

treatment. Similarly, False Positive (FP) refers to patients incorrectly diagnosed as having Parkinson's, potentially causing unnecessary stress and medical expenses. This makes the confusion matrix a useful evaluation tool for understanding the weaknesses and strengths of classification models [40]. The confusion matrix consists of four cells: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), whose assumptions can be seen in Table 2.

Table 2. Confusion Matrix

Actual Class	Predicted Class	
	Class = Yes	Class = No
Class = Yes	True Positive (TP)	False Negative (FN)
Class = No	False Positive (FP)	True Negative (TN)

TP represents the number of samples correctly classified as positive, TN is correctly classified as negative, FP is incorrectly classified as positive, and FN is incorrectly classified as negative. As explained earlier, we will calculate accuracy (Equation 2), precision (Equation 3), recall (Equation 4), and F1-Score (Equation 5) as a reference to decide whether our model is robust or not to the data we use.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$F1 - Score = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (5)$$

3. RESULT

3.1. Experimental Environment

The material used in this research is the Parkinson's Disease Detection dataset from Kaggle. This dataset has 195 data records with 22 independent and one dependent attribute. Not only that, this research also uses Google Colab and Python version 3.10 as the leading platform for data processing and analysis. Google Colab provides an integrated environment with the necessary Python libraries, easy collaboration, and access to higher computing resources, which significantly supports this research.

3.2. ADASYN Process

Before applying ADASYN, a check was made on the class balance of the dataset. The results showed a significant imbalance, where the minority class had a much smaller number of samples than the majority class. To address this issue, the ADASYN method was applied to add synthetic samples to the minority class to balance the number with the majority class. The application of ADASYN helps to ensure that the machine learning model has a better chance of learning fairly and effectively from both classes. The results before and after applying the data balancing method are presented in the figure below, which compares the sample quantity between the majority and minority classes before and after the application of ADASYN.

Vol. 6, No. 5, October 2025, Page. 3495-3508
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4726

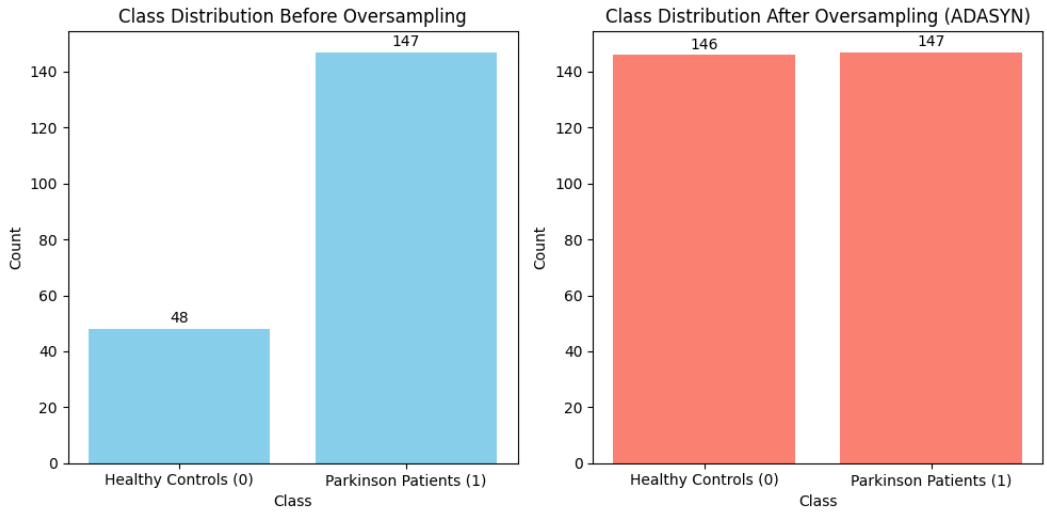

Figure 6. Class Distribution After Oversampling

Based on Figure 6, the data before balancing were 48 records for healthy status and 147 records for positive status. When balancing is done, it becomes 146 records for healthy and 147 for Parkinson's patient status.

3.3. Model Performances

In this study, we compared the performance of three different classification models, namely AdaBoost, LightGBM, and CatBoost, in classifying Parkinson's disease data. We analyze two scenarios: using raw data and using data that has been processed with the ADASYN technique to address class imbalance. The evaluation uses several metrics: accuracy, precision, recall, and F1 score. The following are the overall results of the trials carried out, which can be seen in Table 3.

Table 3. Results of Models

Model	Accuracy	Precision	Recall	F1 Score
AdaBoost	89.74%	89.53%	89.74%	89.59%
LightGBM	95.38%	95.34%	95.38%	95.33%
CatBoost	93.33%	93.30%	93.33%	93.15%
AdaBoost + ADASYN	94.52%	94.73%	94.52%	94.52%
LightGBM + ADASYN	96.92%	97.10%	96.92%	96.92%
CatBoost + ADASYN	96.92%	97.10%	96.92%	96.92%

In the scenario without ADASYN, LightGBM shows the best performance with 95.38% accuracy, 95.34% precision, 95.38% recall, and 95.33% F1 score. CatBoost ranked second with 93.33% accuracy, 93.30% precision, 93.33% recall, and 93.15% F1 score. Meanwhile, AdaBoost has the lowest performance with an accuracy of 89.74%, a precision of 89.53%, a recall of 89.74%, and an F1 score of 89.59%. When using ADASYN to handle class imbalance, performance improvement was observed in all models. LightGBM and CatBoost achieved the same performance with 96.92% accuracy, 97.10% precision, 96.92% recall, 96.92%, and F1-Score. Meanwhile, AdaBoost experienced a significant increase in accuracy of 94.52%, precision of 94.73%, recall of 94.52%, and F1-score of 94.52%. AdaBoost demonstrated the most critical performance improvement (4.78% increase in accuracy) when combined with ADASYN, suggesting that this ensemble method benefits significantly from balanced training data by reducing bias toward the majority class and improving its ability to classify minority samples correctly.

Vol. 6, No. 5, October 2025, Page. 3495-3508
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4726

The identical performance results between LightGBM+ADASYN and CatBoost+ADASYN (96.92% accuracy, 97.10% precision, 96.92% recall, and 96.92% F1-Score) can be attributed to several factors: both algorithms utilize similar gradient boosting frameworks with tree-based learners, the ADASYN balancing technique creates an optimal data distribution that allows both models to reach their performance ceiling on this dataset size, and the relatively small dataset (195 records) may limit the ability to distinguish between the fine-grained differences in these advanced boosting algorithms.

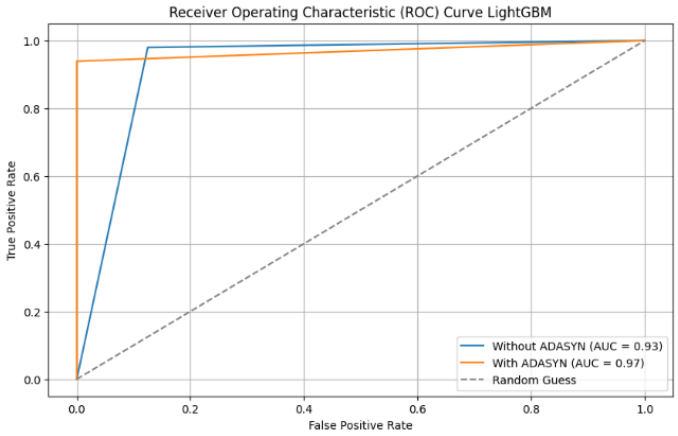

Figure 7. ROC Curve of LightGBM Algorithm

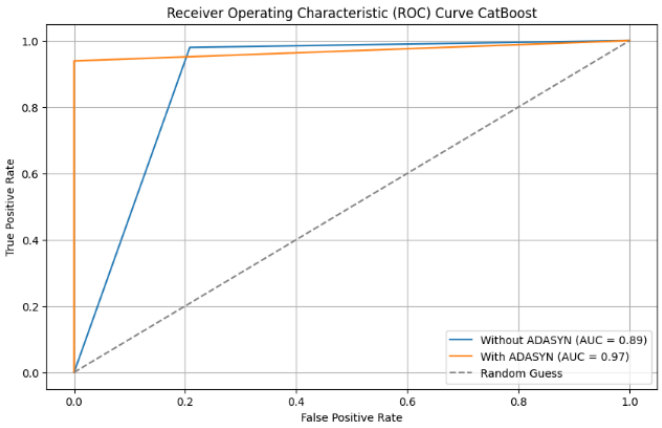

Figure 8. ROC Curve of Catboost Algorithm

Figures 7 and 8 show that using the ADASYN (Adaptive Synthetic Sampling) method significantly improves the AUC results on both machine learning algorithm model performances, namely LightGBM and CatBoost. As illustrated in Figure 7**,** in the model using LightGBM, there is an improvement from an AUC of 0.927 without ADASYN to 0.969 after using ADASYN. This indicates that the addition of synthetic data generated by ADASYN effectively assists the model in classifying unbalanced data, thus improving the model's ability to separate different classes. Similarly, Figure 8 demonstrates that the results of the model using CatBoost also showed significant improvement. The AUC without ADASYN was 0.886, while after using ADASYN, the AUC increased to 0.969. This confirms that the use of ADASYN effectively improved the ability of the CatBoost model to handle class imbalance, thereby improving the model's overall performance in performing classification. The ROC curves in both Figure 7 and Figure 8 clearly show the superior performance of ADASYN-enhanced models, with curves positioned closer to the top-left corner, indicating better true positive rates at lower false positive rates. Thus, the results show that using ADASYN as a resampling method makes an

Vol. 6, No. 5, October 2025, Page. 3495-3508
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4726

essential contribution to improving model performance on classification problems with imbalanced data, using both LightGBM and CatBoost.

4. DISCUSSIONS

In this discussion section, it is worth emphasizing that the results show that using the ADASYN (Adaptive Synthetic Sampling) technique combined with machine learning algorithms such as LightGBM and CatBoost has significantly improved classification performance. The test results in Table 3 show that the model reinforced with ADASYN exhibits higher accuracy than the non-reinforced model. In particular, the ADASYN-LightGBM and ADASYN-CatBoost models achieved an accuracy of 96.92%, demonstrating the ADASYN technique's effectiveness in handling class imbalance in the dataset. This study contributes to the field of computer science and medical informatics by demonstrating the effectiveness of ensemble learning combined with synthetic oversampling techniques in addressing critical challenges in automated disease classification systems.

Table 4. Comparison Models

References	Method	Results
Senturk [41]	SVM	90.76%
	CART	93.84%
	ANN	91.54%
Kadam and Jadhav [42]	DNN	92.19%
	FESA-DNN	93.84%
Al-Fatlawi et al. [43]	DBN	94.00%
Benba et al. [44]	LOSO-SVM	87.50%
Proposed Method	ADASYN-LightGBM	96.92%
	ADASYN-CatBoost	96.92%

As presented in Table 4, our proposed methods (ADASYN-LightGBM and ADASYN-CatBoost with 96.92% accuracy) outperformed several previous studies conducted on Parkinson's disease classification. In addition, the comparison to prior studies presented in Table 4 shows that the approach proposed in this article can compete with current methods in data classification. Specifically, Table 4 demonstrates that our approach surpassed Senturk's SVM (90.76%), CART (93.84%), and ANN (91.54%) methods, as well as Kadam and Jadhav's DNN (92.19%) and FESA-DNN (93.84%) approaches. Furthermore, our results exceeded Al-Fatlawi et al.'s DBN method (94.00%) and Benba et al.'s LOSO-SVM approach (87.50%). Compared with previous studies using various algorithms such as SVM, CART, ANN, DNN, FESA-DNN, DBN, and LOSO-SVM, the results of the proposed approach show significant performance improvement. However, it is essential to note that these comparisons in Table 4 involve different datasets and experimental conditions, which may limit the direct comparability of results. For more robust validation, future studies should evaluate these methods on standardized benchmark datasets to ensure fair comparison. This shows that the combination of ADASYN with machine learning algorithms such as LightGBM and CatBoost has the potential to be an attractive option in addressing the class imbalance problem on classification datasets.

From a computer science perspective, this research advances the field by providing empirical evidence for the synergistic effects of data balancing techniques and ensemble methods in medical classification tasks. The findings contribute to the development of intelligent diagnostic systems and provide practical insights for implementing machine learning solutions in healthcare informatics.

However, although the results obtained show a significant performance improvement, some aspects still need to be considered and improved in the future. For example, further analysis needs to be

Vol. 6, No. 5, October 2025, Page. 3495-3508
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4726

conducted regarding the stability and scalability of the proposed approach, as well as its potential applicability to larger and more diverse datasets. Thus, further research can focus on developing and refining this method to increase its usefulness in various real-world applications.

5. CONCLUSION

Analysis of the results shows that the application of ADASYN to the proposed boosting methods significantly improves classification performance. The findings contribute to computer science by providing empirical evidence that combining ensemble learning algorithms with synthetic oversampling techniques can effectively address class imbalance challenges in medical diagnostic systems. Applying the ADASYN technique improved the model's ability to handle class imbalance, reflected in improved accuracy, precision, recall, F1 score, and AUC. This demonstrates the effectiveness of ADASYN in mitigating the challenges posed by imbalanced datasets, particularly in the classification of Parkinson's disease. This study advances the field of medical informatics and computer science by demonstrating that the integration of ADASYN with gradient boosting methods (LightGBM and CatBoost) can achieve superior diagnostic accuracy (96.92%), providing practical insights for developing intelligent healthcare systems and automated diagnostic tools.

LightGBM and CatBoost achieved optimal performance when combined with ADASYN, while AdaBoost showed the most significant improvement through class balancing. These results establish a foundation for the future development of machine learning-based diagnostic systems in healthcare applications. Despite these promising results, the study's reliance on a small dataset (195 records) limits the generalizability of the findings. Future research should focus on validation with larger, more diverse datasets, exploration of advanced hyperparameter optimization techniques, and investigation of alternative data balancing methods to enhance diagnostic accuracy and system robustness further.

ACKNOWLEDGEMENT

This work was supported by Lambung Mangkurat University and Gazi University, which provided essential resources and assistance.

REFERENCES

[1] M. Hirano, M. Samukawa, C. Isono, S. Kusunoki, and Y. Nagai, “The effect of rasagiline on swallowing function in Parkinson’s disease,” *Heliyon*, vol. 10, no. 1, p. e23407, Jan. 2024, doi: 10.1016/j.heliyon.2023.e23407.

[2] R. K. Sharma and A. K. Gupta, “Voice Analysis for Telediagnosis of Parkinson Disease Using Artificial Neural Networks and Support Vector Machines,” *Int. J. Intell. Syst. Appl.*, vol. 7, no. 6, pp. 41–47, May 2015, doi: 10.5815/ijisa.2015.06.04.

[3] S. J. Chia, E.-K. Tan, and Y.-X. Chao, “Historical Perspective: Models of Parkinson’s Disease,” *Int. J. Mol. Sci.*, vol. 21, no. 7, p. 2464, Apr. 2020, doi: 10.3390/ijms21072464.

[4] C. Taleb, M. Khachab, C. Mokbel, and L. Likforman-Sulem, “A Reliable Method to Predict Parkinson’s Disease Stage and Progression based on Handwriting and Re-sampling Approaches,” in *2018 IEEE 2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR)*, Mar. 2018, pp. 7–12, doi: 10.1109/ASAR.2018.8480209.

[5] H. M. Qasim, O. Ata, M. A. Ansari, M. N. Alomary, S. Alghamdi, and M. Almehmadi, “Hybrid Feature Selection Framework for the Parkinson Imbalanced Dataset Prediction Problem,” *Medicina (B. Aires).*, vol. 57, no. 11, p. 1217, Nov. 2021, doi: 10.3390/medicina57111217.

[6] Q. Huang *et al.*, “Identification and validation of senescence-related genes in Parkinson’s disease,” *Hum. Gene*, vol. 39, p. 201258, Feb. 2024, doi: 10.1016/j.humgen.2024.201258.

[7] S. Rahman, M. Irfan, M. Raza, K. Moyeezullah Ghori, S. Yaqoob, and M. Awais, “Performance Analysis of Boosting Classifiers in Recognizing Activities of Daily Living,” *Int. J. Environ. Res. Public Health*, vol. 17, no. 3, p. 1082, Feb. 2020, doi: 10.3390/ijerph17031082.

[8] H. Jafarzadeh, M. Mahdianpari, E. Gill, F. Mohammadimanesh, and S. Homayouni, “Bagging

Vol. 6, No. 5, October 2025, Page. 3495-3508
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4726

and Boosting Ensemble Classifiers for Classification of Multispectral, Hyperspectral and PolSAR Data: A Comparative Evaluation," *Remote Sens.*, vol. 13, no. 21, p. 4405, Nov. 2021, doi: 10.3390/rs13214405.

[9] S. Basha, D. Rajput, and V. Vandhan, "Impact of Gradient Ascent and Boosting Algorithm in Classification," *Int. J. Intell. Eng. Syst.*, vol. 11, no. 1, pp. 41–49, Feb. 2018, doi: 10.22266/ijies2018.0228.05.

[10] X. Ying, "An Overview of Overfitting and its Solutions," *J. Phys. Conf. Ser.*, vol. 1168, p. 022022, Feb. 2019, doi: 10.1088/1742-6596/1168/2/022022.

[11] S. Wang and Q. Chen, "The Study of Multiple Classes Boosting Classification Method Based on Local Similarity," *Algorithms*, vol. 14, no. 2, p. 37, Jan. 2021, doi: 10.3390/a14020037.

[12] A. Y. Aravkin, G. Bottegal, and G. Pillonetto, "Boosting as a kernel-based method," *Mach. Learn.*, vol. 108, no. 11, pp. 1951–1974, Nov. 2019, doi: 10.1007/s10994-019-05797-z.

[13] H. EL Hamdaoui, S. Boujraf, N. E. H. Chaoui, B. Alami, and M. Maaroufi, "Improving Heart Disease Prediction Using Random Forest and AdaBoost Algorithms," *Int. J. Online Biomed. Eng.*, vol. 17, no. 11, p. 60, Nov. 2021, doi: 10.3991/ijoe.v17i11.24781.

[14] S. M. Ganie, P. K. D. Pramanik, M. Bashir Malik, S. Mallik, and H. Qin, "An ensemble learning approach for diabetes prediction using boosting techniques," *Front. Genet.*, vol. 14, Oct. 2023, doi: 10.3389/fgene.2023.1252159.

[15] S. Aymaz, "Boosting medical diagnostics with a novel gradient-based sample selection method," *Comput. Biol. Med.*, vol. 182, p. 109165, Nov. 2024, doi: 10.1016/j.compbiomed.2024.109165.

[16] D. D. Rufo, T. G. Debelee, A. Ibenthal, and W. G. Negera, "Diagnosis of Diabetes Mellitus Using Gradient Boosting Machine (LightGBM)," *Diagnostics*, vol. 11, no. 9, p. 1714, Sep. 2021, doi: 10.3390/diagnostics11091714.

[17] J. Yan *et al.*, "LightGBM: accelerated genomically designed crop breeding through ensemble learning," *Genome Biol.*, vol. 22, no. 1, p. 271, Dec. 2021, doi: 10.1186/s13059-021-02492-y.

[18] A. V. Dorogush, V. Ershov, and A. Gulin, "CatBoost: gradient boosting with categorical features support," *arXiv Prepr.*, pp. 1–7, 2018, [Online]. Available: http://arxiv.org/abs/1810.11363.

[19] T. M. Khan, S. Xu, Z. G. Khan, and M. Uzair chishti, "Implementing Multilabeling, ADASYN, and ReliefF Techniques for Classification of Breast Cancer Diagnostic through Machine Learning: Efficient Computer-Aided Diagnostic System," *J. Healthc. Eng.*, vol. 2021, pp. 1–15, Mar. 2021, doi: 10.1155/2021/5577636.

[20] M. Mujahid *et al.*, "Data oversampling and imbalanced datasets: an investigation of performance for machine learning and feature engineering," *J. Big Data*, vol. 11, no. 1, p. 87, Jun. 2024, doi: 10.1186/s40537-024-00943-4.

[21] B. A. Omodunbi *et al.*, "Stacked Ensemble Learning for Classification of Parkinson's Disease Using Telemonitoring Vocal Features," *Diagnostics*, vol. 15, no. 12, p. 1467, Jun. 2025, doi: 10.3390/diagnostics15121467.

[22] X. Zhang *et al.*, "An accurate diagnosis of coronary heart disease by Catboost, with easily accessible data," *J. Phys. Conf. Ser.*, vol. 1955, no. 1, p. 012027, Jun. 2021, doi: 10.1088/1742-6596/1955/1/012027.

[23] J. Tanha, Y. Abdi, N. Samadi, N. Razzaghi, and M. Asadpour, "Boosting methods for multi-class imbalanced data classification: an experimental review," *J. Big Data*, vol. 7, no. 1, p. 70, Dec. 2020, doi: 10.1186/s40537-020-00349-y.

[24] J. Liu, Y. Gao, and F. Hu, "A fast network intrusion detection system using adaptive synthetic oversampling and LightGBM," *Comput. Secur.*, vol. 106, p. 102289, Jul. 2021, doi: 10.1016/j.cose.2021.102289.

[25] Y. Cao, X. Zhao, Z. Zhou, Y. Chen, X. Liu, and Y. Lang, "MIAC: Mutual-Information Classifier with ADASYN for Imbalanced Classification," in *2018 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*, Dec. 2018, pp. 494–498, doi: 10.1109/SPAC46244.2018.8965597.

[26] M. Abdel-Basset, D. El-Shahat, I. El-henawy, V. H. C. de Albuquerque, and S. Mirjalili, "A New Fusion of Grey Wolf Optimizer Algorithm with a Two-Phase Mutation for Feature Selection," *Expert Syst. Appl.*, vol. 139, no. 112824, pp. 0957–4174, 2020, doi: 10.1016/j.eswa.2019.112824.

[27] D. Dotcom, "Parkinson Disease Dataset," *Kaggle*. https://www.kaggle.com/datasets/debasisdotcom/parkinson-disease-detection/data (accessed May 01, 2024).
[28] A. N. Ahmed *et al.*, "Machine learning methods for better water quality prediction," *J. Hydrol.*, vol. 578, p. 124084, Nov. 2019, doi: 10.1016/j.jhydrol.2019.124084.
[29] M. Mahmud *et al.*, "Implementation of C5.0 Algorithm using Chi-Square Feature Selection for Early Detection of Hepatitis C Disease," *J. Electron. Electromed. Eng. Med. Informatics*, vol. 6, no. 2, pp. 116–124, Mar. 2024, doi: 10.35882/jeeemi.v6i2.384.
[30] P. Chen, "Effects of normalization on the entropy-based TOPSIS method," *Expert Syst. Appl.*, vol. 136, pp. 33–41, Dec. 2019, doi: 10.1016/j.eswa.2019.06.035.
[31] I. Castiglioni *et al.*, "AI applications to medical images: From machine learning to deep learning," *Phys. Medica*, vol. 83, pp. 9–24, Mar. 2021, doi: 10.1016/j.ejmp.2021.02.006.
[32] F. Thabtah, S. Hammoud, F. Kamalov, and A. Gonsalves, "Data imbalance in classification: Experimental evaluation," *Inf. Sci. (Ny).*, vol. 513, pp. 429–441, Mar. 2020, doi: 10.1016/j.ins.2019.11.004.
[33] S. Dhanka and S. Maini, "Random Forest for Heart Disease Detection: A Classification Approach," in *2021 IEEE 2nd International Conference On Electrical Power and Energy Systems (ICEPES)*, Dec. 2021, pp. 1–3, doi: 10.1109/ICEPES52894.2021.9699506.
[34] H. Kok, M. S. Izgi, and A. M. Acilar, "Evaluation of the Artificial Neural Network and Naive Bayes Models Trained with Vertebra Ratios for Growth and Development Determination," *Turkish J. Orthod.*, vol. 34, no. 1, pp. 2–9, Mar. 2021, doi: 10.5152/TurkJOrthod.2020.20059.
[35] L. Wang, M. Han, X. Li, N. Zhang, and H. Cheng, "Review of Classification Methods on Unbalanced Data Sets," *IEEE Access*, vol. 9, pp. 64606–64628, 2021, doi: 10.1109/ACCESS.2021.3074243.
[36] C. Esposito, G. A. Landrum, N. Schneider, N. Stiefl, and S. Riniker, "GHOST: Adjusting the Decision Threshold to Handle Imbalanced Data in Machine Learning," *J. Chem. Inf. Model.*, vol. 61, no. 6, pp. 2623–2640, Jun. 2021, doi: 10.1021/acs.jcim.1c00160.
[37] M. Zakariah, S. A. AlQahtani, and M. S. Al-Rakhami, "Machine Learning-Based Adaptive Synthetic Sampling Technique for Intrusion Detection," *Appl. Sci.*, vol. 13, no. 11, p. 6504, May 2023, doi: 10.3390/app13116504.
[38] Y. F. Zamzam, T. H. Saragih, R. Herteno, Muliadi, D. T. Nugrahadi, and P.-H. Huynh, "Comparison of CatBoost and Random Forest Methods for Lung Cancer Classification using Hyperparameter Tuning Bayesian Optimization-based," *J. Electron. Electromed. Eng. Med. Informatics*, vol. 6, no. 2, pp. 125–136, Mar. 2024, doi: 10.35882/jeeemi.v6i2.382.
[39] I. Tahyudin, H. A. A. Rozaq, and H. Nambo, "Machine Learning Analysis for Temperature Classification using Bioelectric Potential of Plant," in *2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, Dec. 2022, pp. 465–470, doi: 10.1109/ICITISEE57756.2022.10057768.
[40] S. Napi'ah, T. H. Saragih, D. T. Nugrahadi, D. Kartini, and F. Abadi, "Implementation of Monarch Butterfly Optimization for Feature Selection in Coronary Artery Disease Classification Using Gradient Boosting Decision Tree," *J. Electron. Electromed. Eng. Med. Informatics*, vol. 5, no. 4, Oct. 2023, doi: 10.35882/jeeemi.v5i4.331.
[41] Z. Karapinar Senturk, "Early diagnosis of Parkinson's disease using machine learning algorithms," *Med. Hypotheses*, vol. 138, p. 109603, May 2020, doi: 10.1016/j.mehy.2020.109603.
[42] V. J. Kadam and S. M. Jadhav, "Feature Ensemble Learning Based on Sparse Autoencoders for Diagnosis of Parkinson's Disease," 2019, pp. 567–581.
[43] A. H. Al-Fatlawi, M. H. Jabardi, and S. H. Ling, "Efficient diagnosis system for Parkinson's disease using deep belief network," in *2016 IEEE Congress on Evolutionary Computation (CEC)*, Jul. 2016, pp. 1324–1330, doi: 10.1109/CEC.2016.7743941.
[44] A. Benba, A. Jilbab, and A. Hammouch, "Using Human Factor Cepstral Coefficient on Multiple Types of Voice Recordings for Detecting Patients with Parkinson's Disease," *IRBM*, vol. 38, no. 6, pp. 346–351, Nov. 2017, doi: 10.1016/j.irbm.2017.10.002.