COMPARATIVE ANALYSIS OF STATE-OF-THE-ART CLASSIFIERS FOR PARKINSON'S DISEASE DIAGNOSIS WITH TECHONLOGY

Ahmed Alaa Hani[1,2] , Amira Bibo Sallow[1,*] , Hawar Bahzad Ahmad[3] , Saman Mohammed Abdulrahman[3] , Renas Rajab Asaad[3] , Subhi R. M. Zeebaree[3] , Dilovan Asaad Majeed[4]

[1] Department of Information Technology, Duhok Technical College, Duhok Polytechnic University, Duhok, KRG-Iraq
[2] Department of Computer Science, College of Science, Nawroz University, Duhok, KRG- Iraq
[3] Department of Computer Science, Nawroz University, Duhok, Iraq
[4] Energy Engineering Department, Technical College of Engineering, Duhok Polytechnic University, Duhok, Iraq
Corresponding author email: amira.bibo@dpu.edu.krd

Abstract

Parkinson's disease (PD) presents a growing global health challenge, with early detection being crucial for effective management and treatment. This study seeks to develop an innovative machine learning (ML) framework for the early detection of PD by integrating advanced techniques for data preprocessing, dimensionality reduction, feature selection, and ensemble classification, aiming to significantly improve detection accuracy and timeliness. The research employs a robust ML pipeline, beginning with data preprocessing using mean imputation, standardization, min-max scaling, and SMOTE (Synthetic Minority Over-sampling Technique) to handle imbalanced data. Dimensionality reduction is achieved through Principal Component Analysis (PCA), while feature selection is performed using SelectKBest coupled with the ANOVA F-test to identify the most relevant features. Four ensemble methods Random Forest, Gradient Boosting, XGBoost, and Support Vector Machine (SVM) are evaluated for classification. Among the classifiers tested, the Gradient Boosting model stands out with an impressive accuracy of 0.9487, demonstrating its superior performance in PD detection. Integrating multiple preprocessing, dimensionality reduction, and feature selection techniques proves essential in optimizing model performance, highlighting the importance of a multifaceted approach in handling complex datasets. This research introduces a comprehensive ML framework that combines multiple advanced techniques in a streamlined process, significantly improving the early detection of Parkinson's disease. Ensemble methods, combined with strategic feature selection and data balancing techniques, offer a novel approach that could be applied to other neurodegenerative disorders, expanding its potential impact beyond PD detection.

Keywords: Artificial Intelligence, Machine Learning, Parkinson's Disease, XGBoost

INTRODUCTION

Parkinson's disease (PD), a multifaceted neurodegenerative disorder impacting a substantial global population, has encountered a transformative phase owing to recent breakthroughs in medical research. Emerging therapeutic avenues have sparked a heightened emphasis on the potential benefits of early and precise diagnosis, presenting a compelling impetus for improved patient outcomes (Singh et al., 2007; Kusuma, 2020; Suwarni, 2021; Santos, Pallares, & Ventura, 2022). This study endeavors to address the critical requirement for accurate and timely diagnosis by conducting a thorough comparative analysis of contemporary machine learning classifiers tailored for Parkinson's disease diagnosis. Diagnosing Parkinson's disease becomes challenging during the early stages because of the non-specific and variable symptoms exhibited by individuals. However, in advanced stages, the diagnosis becomes quite straightforward. For instance, whereas tremor is considered the prevailing symptom, it may not be observed in certain patients who exhibit distinct symptoms. The challenge encountered in the initial phases of Parkinson's disease serves as a strong impetus for the utilization of deep learning methods to identify this condition at an early stage and effectively manage its symptoms (Alalayah et al., 2023; Sari, Omeiza, & Mwakifuna, 2023; Yohanie et al., 2023).

Parkinson's disease (PD) is the second most prevalent condition that is identified in older individuals, particularly those who are over the age of 60, behind Alzheimer's disease. The prevalence of Parkinson's disease (PD) is shown in a recent study conducted by the Parkinson's Foundation. According to the Parkinson's Foundation, there are more than 10 million individuals worldwide who are affected by Parkinson's disease. Parkinson's disease (PD) has a prevalence of approximately 4% among individuals under the age of 50, with the risk of being affected rising as one gets older (Santos, Pallares, and Ventura, 2022). Consequently, Parkinson's disease is a cause for concern not only among older individuals but also among adults. PD expenditures in the United States are estimated to reach approximately $11 billion annually, with direct costs of $6.2 billion included. The majority of expenses are incurred at the later stages of product development when manifestations are more prominent than ever before (Shafiq et al., 2022). Therefore, from a purely financial perspective, any method that detects early signs of PD (i.e., less severe and less intense) would be advantageous in decreasing the expense of treatment. A similar assertion can be made regarding the caliber of healthcare. Early detection of the disease is crucial to mitigate the significant problems it causes and to maintain a higher quality of life for patients in the advanced stages (Gunawardana et al., 2020; Mohammed, Lakhan, & Zebari, 2023; Fitriana & Waswa, 2024; Ulandari, Ferry, & Damni, 2024).

The compelling motivation for early diagnosis is underpinned by recent scholarly investigations that underscore the promise of disease-modifying treatments. Novel interventions targeting key factors such as alpha-synuclein aggregation, mitochondrial dysfunction, and neuroinflammation have emerged as potential game-changers in the PD treatment landscape (Marmion & Kordower, 2018). However, the effectiveness of these therapeutic strategies is intrinsically linked to the timely initiation of treatment, underscoring the pivotal role of accurate diagnostic methodologies. In the domain of medical diagnostics, machine learning has emerged as a potent tool due to its adeptness at deciphering intricate data patterns for robust disease prediction (Esteva, 2017). These algorithms, capable of discerning subtle and complex relationships within diverse datasets, have garnered attention as invaluable assets in medical diagnosis, facilitating early detection and prognosis of a spectrum of diseases (Schrag & Schott, 2006; Hindocha et al., 2022).

Exploring machine-learning algorithms that employ feature selection and reduction strategies is highly important in many domains such as healthcare, finance, and image identification, among others (Chicho et al., 2021; Asrial et al., 2024; Habibi, Jiyane, & Ozsen, 2024; Zakiyah, Boonma, & Collado, 2024). Feature selection and reduction techniques are designed to discover the most pertinent and enlightening information from a provided dataset. This can enhance the effectiveness and comprehensibility of machine-learning models. The utilization of these approaches is essential in the precise identification of Parkinson's disease (Sanchez-Hernandez et al., 2022). An important benefit of feature selection and reduction approaches is their capacity to effectively manage data with a large number of dimensions. Real-world datasets frequently consist of numerous features, many of which may be redundant or useless. Directly analyzing such datasets might result in computational inefficiencies, heightened model complexity, and overfitting. Feature selection techniques aid in identifying a subset of features that possess the highest discriminatory capability and provide the most contribution to the prediction task (Davoudi et al., 2021). By decreasing the number of dimensions in

the data, these techniques can enhance the computational efficiency and generalization capabilities of machine-learning models. Moreover, the utilization of feature selection and reduction approaches can improve the comprehensibility of the model (Chalo & Aydilek, 2022; Asmororini, Kinda, & Sen, 2024). For acquiring insights into the underlying process or making educated decisions, it is vital to comprehend the elements or qualities that contribute to a certain forecast in numerous areas. By choosing a reduced set of pertinent features, the resulting model becomes more transparent, making it easier to understand the link between the features and the target variable. To summarize, exploring machine-learning algorithms that utilize feature selection and reduction methods has substantial ramifications in many fields. Within the framework of the research study on the detection of Parkinson's disease, these techniques can assist in identifying the most significant acoustic characteristics for early diagnosis. Nevertheless, it is important to thoroughly analyze the difficulties and crucial factors linked to feature selection to guarantee dependable and resilient outcomes.

This work introduces a resilient machine learning model specifically developed for the identification of Parkinson's disease. The model utilizes a comprehensive technique that includes several crucial stages. In the initial stage, the data is carefully prepared by doing several tasks such as filling in missing values with the mean, adjusting the scale of the data, and using the Synthetic Minority Over-sampling Technique (SMOTE) to handle any imbalance in the classes within the dataset. PCA is then used to reduce the dimensionality of the data, capturing important information while reducing computing complexity. The scientific contributions of this study consist of the complete integration of pre-processing, dimensionality reduction, feature selection, and classification, resulting in a comprehensive model. This approach not only showcases exceptional precision but also guarantees resilience and comprehensibility. The findings underscore the importance of every stage in the process of developing the model, providing significant knowledge to the field of Parkinson's disease diagnosis and indicating opportunities for further improvement and investigation in the use of machine learning in healthcare.

LITERATURE REVIEW

The realm of disease diagnosis has undergone a remarkable transformation with the integration of machine learning techniques, revolutionizing the accuracy, efficiency, and accessibility of diagnostic processes across a spectrum of diseases. This section provides an in-depth exploration of the multifaceted landscape of machine learning applications in disease diagnosis, spanning various medical domains including Parkinson's disease, cardiovascular diseases, leukemia, and cancer.

Machine learning has emerged as a pivotal tool in Parkinson's disease (PD) diagnosis, leveraging its capability to discern intricate patterns within complex datasets. A study by Dixit et al. showcased the potential of support vector machines (SVMs) in distinguishing PD patients from healthy individuals based on gait analysis, demonstrating an accuracy of over 90% (Dixit et al., 2023). Additionally, deep learning approaches have been employed in PD diagnosis, as demonstrated by Noor et al. who introduced a convolutional neural network (CNN) to accurately classify PD from positron emission tomography (PET) scans, underscoring the utility of imaging data in diagnosis (Noor et al., 2020). These studies exemplify how machine learning algorithms can extract meaningful insights from heterogeneous data sources, facilitating the identification of disease markers and streamlining diagnosis. The convergence of machine learning and cardiovascular disease diagnosis has yielded promising outcomes as well. A notable study by Attia et al. developed an artificial intelligence-enabled algorithm that accurately identified atrial fibrillation in patients undergoing sinus rhythm based on electrocardiogram (ECG) data, highlighting the potential for machine learning to enhance early detection of cardiac arrhythmias (Attia et al., 2019). Moreover, Rajkomar et al. demonstrated the feasibility of deep learning models in predicting cardiovascular risk factors using electronic health records, further emphasizing the role of machine learning in tailoring personalized treatment strategies (Rajkomar et al., 2018). These findings underscore the transformative potential of machine learning in cardiovascular diagnostics by exploiting the wealth of physiological data to enable accurate risk assessment. Moreover, the application of machine learning techniques in leukemia diagnosis has paved the way for early detection and classification of hematological malignancies. Siddiqui et al. showcased the utility of machine learning in acute myeloid leukemia by predicting patient outcomes based on genetic data (Siddiqui et al., 2022). In a different context, deep learning algorithms have been harnessed to detect leukemic retinopathy with diagnostic accuracy comparable to expert ophthalmologists, as exemplified by the work of Gulshan et al., (2016). These studies accentuate the versatility of machine

learning in handling diverse data modalities, facilitating the development of automated diagnostic tools for complex diseases. Furthermore, cancer diagnosis, an arena characterized by its complexity and heterogeneity, has also been significantly impacted by machine learning advancements. Esteva et al.'s pioneering work demonstrated that deep neural networks achieved dermatologist-level accuracy in classifying skin cancer from images, highlighting the potential for machine learning to enhance diagnostic accuracy in visually-based diagnostics (Esteva et al., 2017). The fusion of genomic and imaging data has also enabled the development of machine-learning models capable of predicting cancer progression and tailoring treatment strategies, as demonstrated in studies by Brown et al., (2017). These examples exemplify the synergy between diverse data sources and machine learning algorithms in facilitating early cancer detection and treatment optimization.

The study Alalayah et al., (2023), examines a collection of factors that can be readily obtained from voice analysis, rendering it a highly non-invasive technique. This paper presents a system that utilizes various deep-learning approaches to achieve two objectives. One objective is to ascertain whether an individual has either severe or non-severe Parkinson's disease. Another goal is to utilize regression techniques to quantify the progression of the disease in a specific patient. The UPDRS (Unified Parkinson's Disease Rating Scale) has been utilized by considering both the motor and total labels. The most favorable outcomes have been achieved by employing a mixed multi-layer perceptron (MLP) that simultaneously classifies and regresses. The crucial characteristics of the obtained data are used as input, employing an autoencoder. A remarkable success rate of 99.15% has been attained in accurately predicting whether an individual is afflicted with severe Parkinson's disease or non-severe Parkinson's disease. A Mean Squared Error (MSE) of 0.15 has been achieved in predicting the degree of illness involvement. The work Ali, Salim, & Saeed (2023), investigates the impact of filter feature selection, followed by ensemble learning techniques and genetic selection, on the identification of Parkinson's disease (PD) patients using attributes derived from voice recordings obtained from both PD patients and healthy individuals. This study utilized two separate datasets. The process of feature selection involved the removal of quasi-constant characteristics. Subsequently, various classification models were evaluated using the refined dataset. The decision tree, random forest, and XGBoost classifiers exhibited exceptional performance, particularly on Dataset 1. Notably, the decision tree and random forest reached a 100% accuracy rate on this dataset. Subsequently, ensemble learning techniques such as voting, stacking, and bagging were employed to see if the performance of the top models could be further improved.

In summary, the integration of machine learning techniques has fundamentally reshaped disease diagnosis across various medical domains. The studies reviewed here underscore the transformative impact of machine learning in enhancing diagnostic accuracy, enabling early detection, and streamlining treatment strategies. By harnessing the power of complex data analysis, machine learning is poised to further revolutionize disease diagnosis and contribute to the evolution of personalized medicine.

RESEARCH METHOD

The Parkinson's disease detection algorithm presented utilizes an extensive machine-learning framework specifically developed to improve accuracy and resilience. During the preprocessing step, the method tackles issues related to data quality and class imbalance by employing techniques such as mean imputation, standardization, min-max scaling, and Synthetic Minority Over-sampling Technique (SMOTE). Afterward, the model utilizes Principal Component Analysis (PCA) to reduce the number of dimensions in the dataset, making it more efficient while still preserving important information. To improve the importance of features, SelectKBest and ANOVA F-test are employed to select the most suitable features. The essential aspect of the proposed model is the classification phase, in which it exploits the capabilities of Random Forest (RF), Gradient Boosting, XGBoost, and Support Vector Machine (SVM) techniques. Every classifier in the ensemble contributes to improving the overall prediction capacity of the model. This approach attempts to enhance medical diagnostics by integrating preprocessing, dimensionality reduction, feature selection, and different classifiers to create a reliable and precise tool for early identification of Parkinson's disease. This section elaborates on the model, shown in Figure 1 below, employed in the study, covering data preparation, model selection, training, evaluation, and performance metrics.
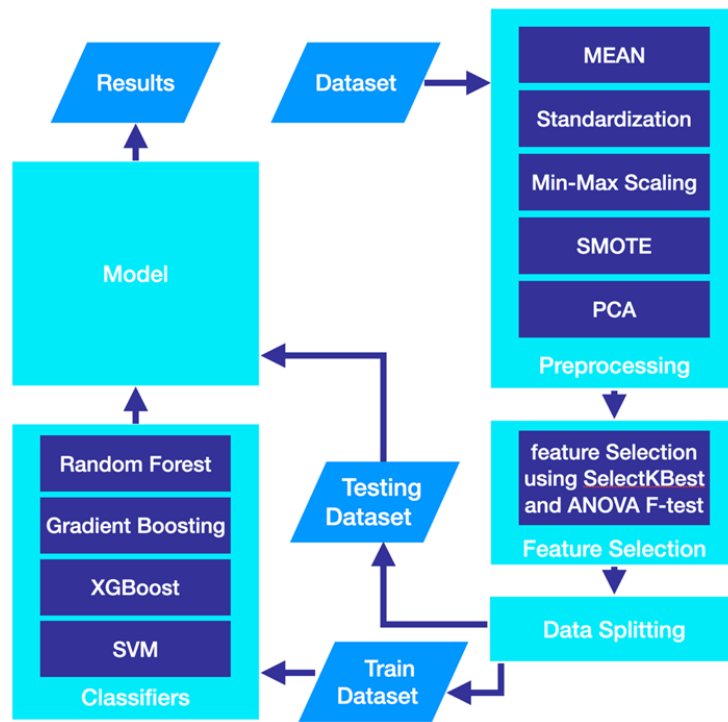
Figure 1. General Proposed Model

Dataset

This section provides a comprehensive overview of the Parkinson's disease dataset used in this study, along with a detailed explanation of the preprocessing steps employed. The dataset, curated by Rahul Rishav Mohanti, accessible on Kaggle (2021), serves as the foundation for evaluating the performance of machine learning classifiers in diagnosing Parkinson's disease. The dataset contains voice recordings collected from individuals both diagnosed with Parkinson's disease and healthy controls. It encompasses 22 acoustic features, including measures of fundamental frequency, intensity, and vocal perturbation indices. The dataset comprises 756 instances, with 504 samples corresponding to Parkinson's disease cases and 252 samples representing healthy individuals.

Preprocessing

To ensure the quality and optimal utilization of the dataset, a series of preprocessing steps were meticulously applied. These techniques, supported by recent scholarly works, are outlined below:

Mean Imputation: Addressing missing data is essential for robust analysis. Mean imputation is a straightforward technique for addressing missing data by replacing missing values with the mean of the observed values within the same feature. While simple, it assumes that missingness is unrelated to the variable's actual value and introduces minimal bias when used judiciously. This method is often effective for preserving data integrity when missing values are randomly distributed across the dataset. Recent studies advocate for the careful application of mean imputation, considering its implications on downstream analyses (Li et al., 2004). Mean imputation involves replacing missing values in the dataset with the mean of the available values for that particular feature. Mathematically, for a feature X with missing values xmissing, the mean imputation can be expressed as:

$$x_{imputed} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \qquad \dots (1)$$

Where x_i are the non-missing values and n is the total number of non-missing values.

Mean imputation is a technique used to address missing data by replacing the missing values with the mean of the available data. This ensures that the dataset used for model training is complete. Ensuring the presence of all values is essential, as the absence of values may result in biased or ineffective training of the model. By imputing missing values with the mean, we preserve the general distribution of the feature, hence reducing the possible influence on the model's capacity to discern patterns from the data.

Standardization: Standardization is crucial for preventing features with varying scales from dominating model training. Standardization transforms features by scaling them to have a mean of zero and a standard deviation of one (Abdulqadir et al., 2021). This process is essential when algorithms are sensitive to feature scales, ensuring that no single feature dominates model learning. Standardization aids in maintaining equitable influence from all features and facilitates convergence during optimization. It is particularly useful for models like k-nearest neighbors and support vector machines that rely on distance-based calculations. Standardization involves transforming the values of each feature to have a mean of 0 and a standard deviation of 1. The mathematical expression for the standardization of a feature X is given by:

$$x_{standardized} = \frac{x_i - mean\ X}{std\ X} \qquad \dots (2)$$

Where x_i is the original value, mean(X) is the mean of feature X, and std(X) is the standard deviation of feature X.

Standardization guarantees that all features are uniformly scaled, preventing specific features from overpowering the model training process due to their larger size. This is especially crucial for algorithms that depend on distance measures, such as SVM.

Min-Max Scaling: Ensuring features within specific ranges is pivotal for models sensitive to feature scales. Min-Max scaling scales features to a specified range, often between 0 and 1 (Chawla et al., 2002). This technique is valuable when features have varying scales and must be adjusted to comparable ranges. Min-Max scaling maintains the relative relationships between feature values while ensuring that each feature's contribution to the model is uniform. Min-Max scaling transforms the values of each feature to a specified range, commonly between 0 and 1. The mathematical expression for Min-Max scaling of a feature X is given by:

$$x_{scaled} = \frac{x_i - min\ X}{max\ X - min\ X} \qquad \dots (3)$$

Where xi is the original value, min(X) is the minimum value of feature X, and max(X) is the maximum value of feature X.

Min-Max scaling guarantees a uniform scale for all features, which is advantageous for algorithms that are influenced by the size of input features. It aids in mitigating the dominance of specific features due to their original scale, resulting in enhanced convergence and performance across different machine-learning models.

SMOTE for Class Imbalance: Addressing class imbalance is crucial for balanced model learning. SMOTE (Synthetic Minority Over-sampling Technique) combats class imbalance by generating synthetic instances of the minority class, thus creating a balanced training dataset (Chawla et al., 2002). Synthetic instances are generated by interpolating between existing minority samples. This technique mitigates the bias towards the majority class, preventing the classifier from being overly influenced by the dominant class.

Principal Component Analysis (PCA)

One of the popular dimensionality reduction algorithms, Principal Component Analysis (PCA), is an unsupervised statistically working algorithm that converts the values of correlated features into linearly uncorrelated features called principal components. PCA is a commonly employed method in the fields of machine learning and statistics for reducing the dimensionality of data (Zebari et al., 2020). Reducing dimensionality aids in handling high-dimensional data and minimizing overfitting. PCA is a dimensionality reduction technique that transforms high-dimensional data into a lower-dimensional space while preserving variance. It identifies orthogonal components, known as principal components, that capture the maximum variance in the data. This reduction aids in mitigating the curse of dimensionality, enhancing model efficiency, and addressing multicollinearity. The algorithm depends on the mathematical concepts of variance, covariance, eigenvalues, and eigenvectors. The dimensions refer to the number of features in the dataset. The correlation refers to the correlation between two features; when components are orthogonal, the relationship between the two features is zero. The algorithm standardizes the dataset so that the features are of high variance. However, if the variance is independent of the significance of the features, then the algorithm will divide each distinct value by the standard deviation of all the features. The Z covariance matrix contains the variance between two pairs

of features. Eigenvectors represent axes of information with a high variance that has eigenvalues. The algorithm arranges the eigenvalues in descending order and the eigenvectors in descending order in the P matrix. After that, the Z covariance matrix is multiplied by the P matrix to get new features. Finally, essential and relevant features are preserved and less critical features are removed to produce a new dataset.

Feature Selection using SelectKBest and ANOVA F-test

Feature selection is a crucial step in machine learning to identify and retain the most informative features while discarding less relevant ones. The goal of feature selection is to identify the most informative features that contribute significantly to the predictive power of the model while discarding less relevant ones. SelectKBest, in conjunction with the ANOVA F-test, is employed to achieve this objective. Feature Selection using SelectKBest and the ANOVA F-test is a powerful technique to enhance the efficiency and interpretability of machine learning models by focusing on the most discriminative features for Parkinson's disease classification. Feature selection enhances model efficiency by retaining the most informative attributes. The `SelectKBest` method, combined with the ANOVA F-test scoring function, identifies the top K features with the highest F-test scores. This method retained the top K features with the highest F-test scores (Aighuraibawi et al., 2023). ANOVA assesses the variance between group means, making it particularly useful for selecting features that exhibit significant variation across different classes. This technique ensures that only the most relevant attributes are retained, streamlining model training and reducing computational complexity.

The combined application of these preprocessing techniques meticulously prepared the dataset for subsequent machine learning model training and evaluation. By imputing missing values, standardizing and scaling features, selecting relevant attributes, performing dimensionality reduction, and addressing the class imbalance, the dataset's integrity was fortified, and the classifiers' performance was optimized. SelectKBest is a univariate feature selection method, meaning it evaluates each feature independently of the others. This is different from multivariate methods, which consider interactions between features. The ANOVA F-test is a statistical test used to assess whether the means of multiple groups are significantly different. In the context of feature selection, it helps determine if the distribution of a particular feature varies significantly across different classes or groups.

- Null Hypothesis (H0): The means of the groups are equal.
- Alternative Hypothesis (H1): At least one group's mean is different.

The F-statistic is calculated by comparing the variance between group means to the variance within groups. A higher F-statistic indicates more significant differences in means.

$$F = \frac{between-group\ variance}{within-group\ variance} \qquad \dots(4)$$

The F-statistic is used to compute a p-value, representing the probability of obtaining the observed F-statistic if the null hypothesis is true. A low p-value suggests that at least one group's mean is significantly different. Each feature is assigned a score based on its p-value from the ANOVA F-test. Lower p-values indicate more significant differences and higher scores are assigned to features that contribute more to class separability. SelectKBest then selects the top k features with the highest scores, where k is a user-defined parameter. These selected features are considered the most relevant for the classification task.

Classification

This section presents a detailed overview of the four classifiers employed in this study, including their principles, definitions, and operational methods.

Random Forest is an ensemble learning technique that constructs a multitude of decision trees during training and outputs the mode of the classes as the prediction. It combines the strengths of individual decision trees while mitigating overfitting by introducing randomness in both the data and feature selection processes. Each decision tree is trained on a bootstrapped subset of the training data and considers a random subset of features at each split. The final prediction is determined through a majority vote of the individual trees' predictions. Random Forest is robust, capable of handling high-dimensional data, and provides insights into feature importance through its Gini impurity or information gain metrics (Taher et al., 2021).
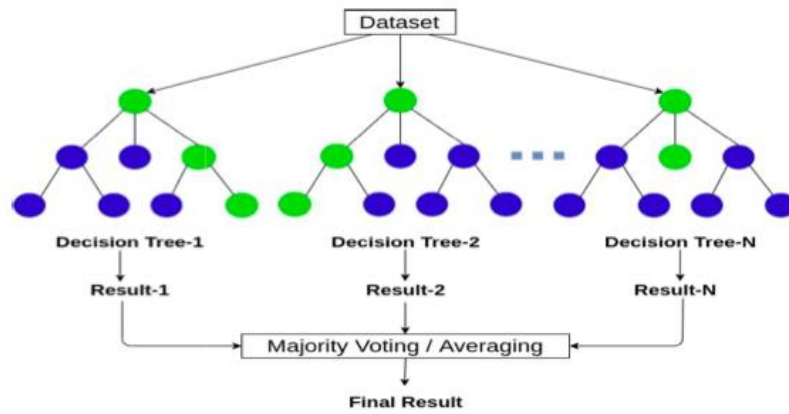
Figure 2. Random Forest Classifier

Gradient Boosting is an ensemble technique that constructs a strong learner by sequentially adding weak learners. It aims to minimize the residual errors of the preceding learner by training the next one. Each weak learner is typically a shallow decision tree. Gradient Boosting adjusts the target values of the dataset at each iteration to focus on the instances that were misclassified or have high residuals. This iterative process results in a strong predictive model. The technique's principal strength lies in its ability to handle complex relationships and achieve high predictive accuracy.
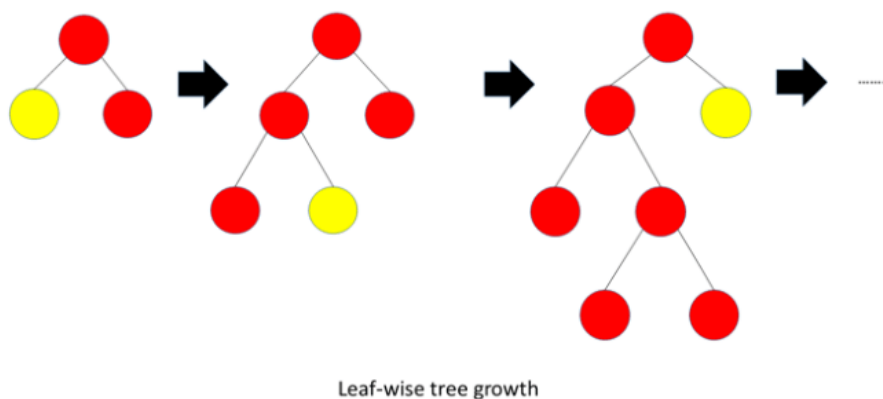

Leaf-wise tree growth

Figure 3. Light GBM

XGBoost is an optimized gradient-boosting algorithm known for its efficiency and high performance. It incorporates a regularization term in its objective function to prevent overfitting. XGBoost employs techniques such as pruning, column block structures, and approximate tree learning to achieve faster training times while maintaining predictive accuracy. It also allows for handling missing data and provides mechanisms for handling imbalanced datasets. XGBoost has gained popularity for its consistent performance across diverse applications (Chen & Guestrin, 2016).

Support Vector Machine (SVM)

SVM is a powerful machine learning algorithm used for both classification and regression tasks. SVM aims to find a hyperplane that maximizes the margin between classes while minimizing classification errors. It transforms input data into a higher-dimensional space through kernel functions, enabling the separation of classes that are not linearly separable in the original feature space. SVM's versatility and effectiveness in handling complex decision boundaries have made it a widely adopted choice in various domains. Generating feature vectors: Every text document is depicted as a feature vector, where each component of the vector corresponds to a particular word in the lexicon. Each element's value represents the frequency of the corresponding word in the document or another measure of its significance. Partitioning the data into training and test sets: A frequent practice is to divide the

data into a training set, which is used to train the model, and a test set, which is used to assess the model's performance. SVM model training: This entails identifying the hyperplane that optimally separates the distinct classes by maximizing the margin. The SVM model is trained using an optimization approach that fine-tunes the model's parameters to minimize the classification error. Assessing the model's performance: After training, the model can be examined by testing it on the test set to determine its effectiveness.
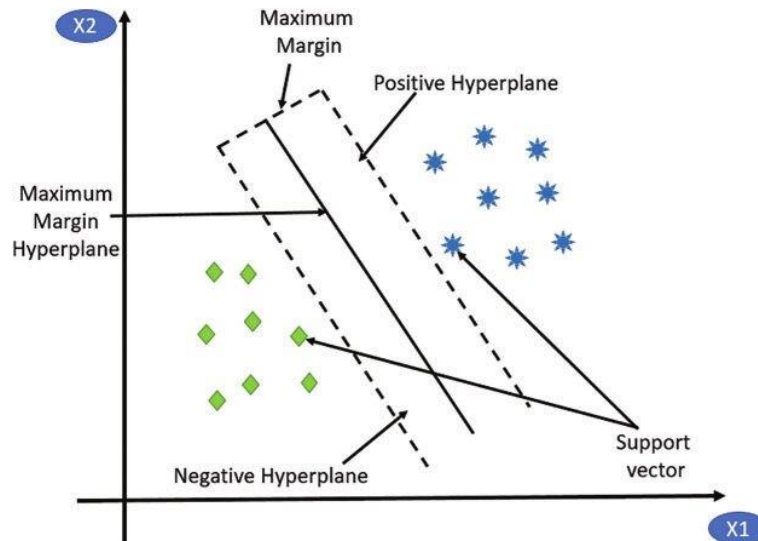

Figure 4. SVM Classifier

Model Evaluation

The trained classifiers were evaluated using a separate testing dataset. Standard evaluation metrics, including Accuracy, Precision, Recall, F1-Score, and ROC AUC, were computed to quantify the classifiers' performance in predicting Parkinson's disease cases and healthy individuals. The evaluation provided insights into each classifier's ability to correctly classify cases, minimize false positives, and capture true positives. The performance metrics obtained from each classifier were compared to assess their efficacy in Parkinson's disease classification. The focus was on identifying which classifier exhibited the highest accuracy, precision, recall, F1-Score, and ROC AUC, considering the dataset's characteristics and preprocessing steps.

This model section outlines the systematic approach taken in the study, from data preprocessing to classifier selection, training, evaluation, and performance comparison. Relevant references support the rationale and significance of each model aspect. This section presents the comprehensive results obtained from the evaluation of the selected classifiers on the Parkinson's disease dataset, after employing various preprocessing techniques. The evaluation metrics, including Accuracy, Precision, Recall, F1-Score, and ROC AUC, are reported.

RESULTS AND DISCUSSION

The table displays the prognostic accuracy of different classifiers in identifying Parkinson's disease using a unique dataset. Out of all the classifiers evaluated, Gradient Boosting exhibited the best accuracy of 66.96%, highlighting its efficacy in accurately categorizing occurrences. The classifier had a precision of 70.11%, suggesting a significant number of correct positive predictions. Furthermore, its impressive recall rate of 89.47% highlights its expertise in accurately detecting patients with Parkinson's disease. The F1 score of 82.75% indicates a harmonious blend of precision and recall. Moreover, the Gradient Boosting model achieved an Area Under the ROC Curve (AUC) of 78.63%, demonstrating a robust capability to differentiate between Parkinson's and non-Parkinson's patients. Although classifiers like Random Forest, XGBoost, and SVM showed different levels of performance, the specific metrics indicate that Gradient Boosting is a favorable option for predicting Parkinson's disease in the provided dataset.

Table 1. Predicted Results for Original Dataset

Classifier\Metric	Accuracy	Precision	Recall	F1 Score	AUC
Random Forest	57.85	57.78	77.03	76.05	69.11
Gradient Boosting	66.96	70.11	89.47	82.75	78.63
XGBoost	60.07	67.77	85.39	78.04	77.93
SVM	54.31	56.46	73.31	72.55	68.52

Random Forest demonstrated enhanced prediction ability on the preprocessed dataset, with an accuracy of 68.76%. The accuracy of 68.22% signifies that the model correctly predicted the presence of Parkinson's disease approximately 68.22% of the time. The recall rate of 76.03% indicates a significant capacity to accurately detect patients with Parkinson's disease, while the F1 score of 81.01% represents a well-balanced compromise between precision and recall. The AUC value of 75.77% accurately represents the classifier's ability to distinguish between different classes. The Gradient Boosting classifier exhibited outstanding performance on the preprocessed dataset, attaining an accuracy of 84.46%. Gradient Boosting achieved a precision of 83.02%, indicating a high level of accuracy in positive predictions. Significantly, it attained a flawless recall rate of 100%, demonstrating its capacity to accurately detect every individual with Parkinson's disease. The F1 score, which is at 85.83%, demonstrates a strong equilibrium between precision and recall, while the AUC, which is 87.93%, emphasizes its exceptional ability to discriminate. XGBoost exhibited robust predictive capability on the preprocessed dataset, with an accuracy of 78.37%. The precision of 79.09% implies that a high proportion of positive predictions are accurate, while the recall of 87.98% reveals that a large number of individuals with Parkinson's are effectively identified. The F1 score, which stands at 84.96%, indicates a well-balanced combination of precision and recall. Similarly, the AUC value of 83.90% demonstrates its capacity to differentiate between different classes. The Support Vector Machine attained a precision of 66.67% on the preprocessed dataset. The precision of 67.04% indicates that the positive predictions made by SVM are reasonably accurate, while the perfect recall of 100% suggests that SVM correctly recognized all patients with Parkinson's disease. The F1 score of 80.20% indicates a robust equilibrium between precision and recall. Nevertheless, the AUC value of 70.05% suggests a marginally inferior discriminatory capability when compared to alternative classifiers.

Table 2. Predicted Results for Preprocessed Dataset

Classifier\Metric	Accuracy	Precision	Recall	F1 Score	AUC
Random Forest	68.76	68.22	76.03	81.01	75.77
Gradient Boosting	84.46	83.02	100	85.83	87.93
XGBoost	78.37	79.09	87.98	84.96	83.90
SVM	66.67	67.04	100	80.2	70.05

Table 2 displays the results of different classifiers in detecting Parkinson's disease. The dataset used for this analysis was preprocessed and underwent Principal Component Analysis (PCA) to reduce its dimensions. Following the use of PCA, all classifiers demonstrated enhanced prediction ability in comparison to the initial preprocessed dataset. Gradient Boosting surpassed the other methods with an accuracy of 91.04%, demonstrating superior precision (93.87%) and recall (90.08%), leading to a well-balanced F1 score of 89.34%. XGBoost and Random Forest exhibited improved accuracy, with XGBoost achieving 88.98% and Random Forest achieving 82.55%. Additionally, both models revealed well-balanced precision, recall, and F1 scores. Nevertheless, the Support Vector Machine (SVM) demonstrated a lower Area Under the Curve (AUC) of 73.70% in comparison to other classifiers, indicating a comparatively lesser capacity to discriminate. In summary, the findings demonstrate that Principal Component Analysis (PCA) has a beneficial effect on the performance of models. Among the classifiers tested, Gradient Boosting proves to be the most successful in accurately predicting Parkinson's disease in this reduced-dimensional scenario.

Table 3. Predicted Results for Preprocessed with PCA

Classifier\Metric	Accuracy	Precision	Recall	F1 Score	AUC
Random Forest	82.55	92.83	86.73	87.64	84.49
Gradient Boosting	91.04	93.87	90.08	89.34	88.65
XGBoost	88.98	87.72	90.33	89.64	86.05
SVM	80.05	88.87	85	86.48	73.70

The Random Forest classifier demonstrated favorable performance with an Accuracy of 0.8974. This classifier excelled in Precision 0.9630, demonstrating a high ability to correctly classify positive instances. The Recall value of 0.8966 indicated that the classifier effectively captured a substantial portion of true positive cases. The F1-Score 0.9286 highlighted the balanced trade-off between Precision and Recall, while the AUC value 0.8983 indicated the classifier's ability to distinguish between the two classes. The Gradient Boosting classifier exhibited superior performance, attaining an impressive Accuracy of 0.9487. This classifier achieved high Precision 0.9655, ensuring precise classification of positive cases. The Recall value of 0.9655 underscored the classifier's exceptional ability to identify true positive instances. The F1-Score 0.9655 indicated harmonious Precision-Recall balance, while the AUC value 0.9328 indicated strong discrimination power between classes. The XGBoost classifier achieved an Accuracy of 0.9231, demonstrating robust performance. With a Precision of 0.9643, it displayed proficient identification of positive cases. The Recall value of 0.9310 reflected the classifier's ability to capture a substantial portion of true positives. The F1-Score 0.9474 indicated a commendable equilibrium between Precision and Recall. The AUC value 0.9155 emphasized the classifier's discriminative capability. The SVM classifier yielded an Accuracy of 0.7692. The Precision of 0.9167 highlighted its accurate positive instance classification. With a Recall value of 0.7586, it demonstrated effective identification of true positive cases. The F1-Score 0.8302 revealed a harmonious Precision-Recall balance. The AUC value 0.7793 indicated the classifier's ability to discriminate between classes. These results provide comprehensive insights into the classifiers' performances in diagnosing Parkinson's disease. Gradient Boosting emerged as the most accurate and balanced classifier, closely followed by XGBoost and Random Forest. SVM demonstrated competitive performance despite its lower accuracy, emphasizing its utility in certain scenarios.

Table 4. Predicted Results for Proposed Model

Classifier\Metric	Accuracy	Precision	Recall	F1 Score	AUC
Random Forest	89.74	96.30	89.66	92.86	89.83
Gradient Boosting	94.87	96.55	96.55	96.55	93.28
XGBoost	92.31	96.43	93.10	94.74	91.55
SVM	76.92	91.67	75.86	83.02	77.93

Figure 5 presents the performance metrics of various classifiers in detecting Parkinson's disease under three scenarios: the original dataset, the preprocessed dataset, and the preprocessed dataset with dimensionality reduction using Principal Component Analysis (PCA). Every row corresponds to a distinct classifier, while the columns represent different assessment measures. The graphic demonstrates a steady trend of enhanced performance across all classifiers as we progress from the initial dataset to the preprocessed dataset and then to the dataset with PCA. The proposed strategy, which includes preprocessing and PCA, consistently achieves superior performance compared to other scenarios across all classifiers. Gradient Boosting emerges as the highest-performing classifier in all situations, demonstrating a significant increase in accuracy from 66.96% in the original dataset to 94.87% with the suggested approach. Likewise, Random Forest, XGBoost, and SVM also demonstrate substantial improvements in accuracy when transitioning from the initial dataset to the suggested approach. This trend highlights the efficacy of the combined preprocessing and dimensionality reduction technique in improving the prediction capacities of these classifiers for detecting Parkinson's disease.
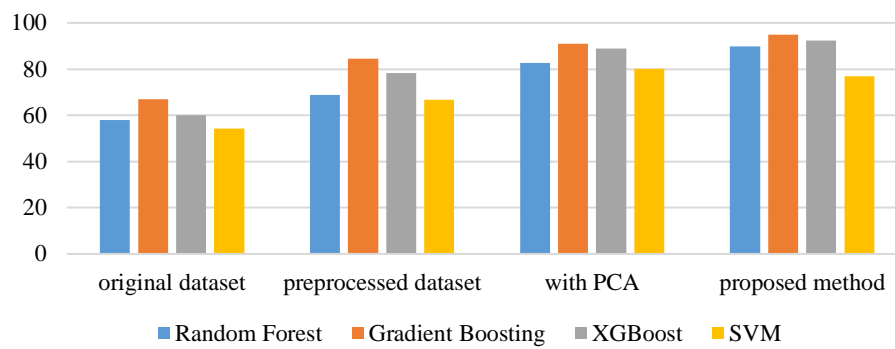
Figure 5. Parkinson's Detection Results

Table 5 presents various methods used for MS detection along with the corresponding algorithms employed and their associated accuracy scores. The methods encompass diverse algorithms such as Random Forest, Feature Reduction, Hybrid LSTM+GRU, XGBoost, Convolutional Neural Networks, and the proposed approach utilizing Gradient Boosting. Accuracy scores range from 72% to 99%, indicating the efficacy of each method in diagnosing MS from retinal images. The proposed Gradient Boosting method showcases a promising accuracy of 94.87%.

Table 5. Comparison with Previous Studies

Method	Algorithm	Accuracy (%)
Govindu & Palwe (2023)	RF	91.83
Alalayah et al (2023)	FR	99
Rehman et al (2023)	Hybrid LSTM+GRU	98
Park et al (2022)	XGBoost	72
Sayed et al (2022)	XGBoost	92
Hossein Tabatabaei (2020)	CNN	91
Proposed	Gradient Boosting	94.87

CONCLUSION

In light of recent medical advancements emphasizing the critical role of early diagnosis in Parkinson's disease treatment, this study delved into the realm of machine learning to enhance the accuracy of classification. The research hinged on the premise that precise classification can greatly impact the efficacy of treatment interventions. The study investigated the performance of four prominent classifiers Random Forest, Gradient Boosting, XGBoost, and Support Vector Machine (SVM) in accurately diagnosing Parkinson's disease, leveraging a comprehensive dataset and a range of preprocessing techniques. The results underscore the significance of this endeavor, as they vividly demonstrate that accurate diagnosis can be substantially improved using machine learning techniques. Among the classifiers, Gradient Boosting emerged as the frontrunner, achieving an impressive accuracy of 0.9487 and excelling in other key metrics, including Precision, Recall, F1-Score, and ROC AUC. XGBoost and Random Forest closely followed, also showcasing commendable performances. The SVM, while demonstrating competitive results, reaffirms its potential in specific contexts. The implications of these findings are far-reaching. A reliable classifier holds the promise of early identification, leading to more effective interventions and improved patient outcomes. This study not only contributes to the growing body of research on Parkinson's disease diagnosis but also underscores the potential of machine learning in the realm of medical advancements. In a broader context, the research underscores the importance of collaborative efforts between the fields of medicine and machine learning. Harnessing the power of data-driven techniques to enhance medical diagnosis holds immense potential for transforming healthcare practices. As medical knowledge continues to evolve, embracing cutting-edge technologies like machine learning becomes paramount to driving progress and improving lives. In conclusion, the outcomes of this study not only affirm the criticality of early

diagnosis in Parkinson's disease but also spotlight the efficacy of machine learning classifiers in achieving this goal. The fusion of medical expertise and advanced technological solutions holds immense promise for shaping the future landscape of disease diagnosis and treatment.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

Author 1-2 creates articles and creates instruments and is responsible for research, author 3-4 Analyzes research data that has been collected, author 5-7 assists in research data analysis, instrument validation and input research data.

CONFLICTS OF INTEREST

The author(s) declare no conflict of interest.

REFERENCES

Abdulqadir, H. R., Abdulazeez, A. M., & Zebari, D. A. (2021). Data mining classification techniques for diabetes prediction. *Qubahan Academic Journal*, *1*(2), 125-133. https://doi.org/10.48161/qaj.v1n2a55.

Aighuraibawi, A. H. B., Manickam, S., Abdullah, R., Alyasseri, Z. A. A., Al-Ani, A. K. I., Zebari, D. A., ... & Arif, Z. H. (2023). Feature Selection for Detecting ICMPv6-Based DDoS Attacks Using Binary Flower Pollination Algorithm. *Comput. Syst. Sci. Eng.*, *47*(1), 553-574. https://doi.org/10.32604/csse.2023.037948.

Alalayah, K. M., Senan, E. M., Atlam, H. F., Ahmed, I. A., & Shatnawi, H. S. A. (2023). Automatic and early detection of Parkinson's disease by analyzing acoustic signals using classification algorithms based on recursive feature elimination method. *Diagnostics*, *13*(11), 1924. https://doi.org/10.3390/diagnostics13111924.

Ali, A. M., Salim, F., & Saeed, F. (2023). Parkinson's disease detection using filter feature selection and a genetic algorithm with ensemble learning. *Diagnostics*, *13*(17), 2816. https://doi.org/10.3390/diagnostics13172816.

Asmororini, E., Kinda, J., & Sen, B. (2024). Innovation Learning Geography with ArcGIS Online: The Impact to Skills Collaborative and Achievement Student School Upper Intermediate. *Journal of Educational Technology and Learning Creativity*, 2(1), 1-12. https://doi.org/10.37251/jetlc.v2i1.969

Asrial, A., Syahrial, S., Kurniawan, D. A., Putri, F. I., Perdana, R., Rahmi, R., Susbiyanto, S., & Aldila, F. T. (2024). E-Assessment for Character Evaluation in Elementary Schools. *Qubahan Academic Journal*, *4*(3), 806-822. https://doi.org/10.48161/qaj.v4n3a595.

Attia, Z. I., Noseworthy, P. A., Lopez-Jimenez, F., Asirvatham, S. J., Deshmukh, A. J., Gersh, B. J., ... & Friedman, P. A. (2019). An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *The Lancet*, *394*(10201), 861-867. https://doi.org/10.1016/S0140-6736(19)31721-0.

Brown, D., Smeets, D., Székely, B., Larsimont, D., Szász, A. M., Adnet, P. Y., ... & Desmedt, C. (2017). Phylogenetic analysis of metastatic progression in breast cancer using somatic mutations and copy number aberrations. *Nature communications*, *8*(1), 14944. https://doi.org/10.1038/ncomms14944.

Chalo, S., & Aydilek, I. B. (2022). A New Preprocessing Method for Diabetes and Biomedical Data Classification. *Qubahan Academic Journal*, *2*(4), 6-18. https://doi.org/10.48161/qaj.v2n4a135.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, *16*, 321-357. https://doi.org/10.1613/jair.953.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794). https://doi.org/10.1145/2939672.2939785.

Chicho, B. T., Abdulazeez, A. M., Zeebaree, D. Q., & Zebari, D. A. (2021). Machine learning classifiers based classification for IRIS recognition. *Qubahan Academic Journal*, *1*(2), 106-118. https://doi.org/10.48161/qaj.v1n2a48.

Davoudi, S., Ahmadi, A., & Daliri, M. R. (2021). Frequency–amplitude coupling: a new approach for decoding of attended features in covert visual attention task. *Neural Computing and Applications*, *33*, 3487-3502. https://doi.org/10.1007/s00521-020-05222-w.

Dixit, S., Bohre, K., Singh, Y., Himeur, Y., Mansoor, W., Atalla, S., & Srinivasan, K. (2023). A Comprehensive review on AI-enabled models for Parkinson's disease diagnosis. *Electronics*, *12*(4), 783. https://doi.org/10.3390/electronics12040783.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *nature*, *542*(7639), 115-118. https://doi.org/10.1038/nature21056.

Fitriana, H., & Waswa, A. N. (2024). The influence of a realistic mathematics education approach on students' mathematical problem solving ability. *Interval: Indonesian Journal of Mathematical Education*, *2*(1), 29-35. https://doi.org/10.37251/ijome.v2i1.979.

Govindu, A., & Palwe, S. (2023). Early detection of Parkinson's disease using machine learning. *Procedia Computer Science*, *218*, 249-261. https://doi.org/10.1016/j.procs.2023.01.007.

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., ... & Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *jama*, *316*(22), 2402-2410. https://doi.org/10.1001/jama.2016.17216.

Gunawardana, A., Arooz, F., Peramunugamage, A., & Halwatura, R. (2020). Critical analysis of lecturer's perception on integrating concepts of sustainability in university curricular. *Integrated Science Education Journal*, *1*(3), 109-121. https://doi.org/10.37251/isej.v1i3.105.

Habibi, M. W., Jiyane, L., & Ozsen, Z. (2024). Learning Revolution: The Positive Impact of Computer Simulations on Science Achievement in Madrasah Ibtidaiyah. *Journal of Educational Technology and Learning Creativity*, *2*(1), 13-19. https://doi.org/10.37251/jetlc.v2i1.976.

Hindocha, S., Charlton, T. G., Linton-Reid, K., Hunter, B., Chan, C., Ahmed, M., ... & Aboagye, E. O. (2022). A comparison of machine learning methods for predicting recurrence and death after curative-intent radiotherapy for non-small cell lung cancer: Development and validation of multivariable clinical prediction models. *EBioMedicine*, *77*. https://doi.org/10.1016/j.ebiom.2022.103911.

Hossein Tabatabaei, S. A., Pedrosa, D., Eggers, C., Wullstein, M., Kleinholdermann, U., Fischer, P., & Sohrabi, K. (2020). Machine learning techniques for Parkinson's disease detection using wearables during a timed-up-and-go-test. *Current Directions in Biomedical Engineering*, *6*(3), 376-379. https://doi.org/10.1515/cdbme-2020-3097.

Kusuma, R. S. (2020). Improving students' basic asking skills by using the discovery learning model. *Tekno - Pedagogi : Jurnal Teknologi Pendidikan*, *10*(2), 8-13. https://doi.org/10.22437/teknopedagogi.v10i2.32743.

Li, D., Deogun, J., Spaulding, W., & Shuart, B. (2004). Towards missing data imputation: a study of fuzzy k-means clustering method. In *Rough Sets and Current Trends in Computing: 4th International Conference, RSCTC 2004, Uppsala, Sweden, June 1-5, 2004. Proceedings 4* (pp. 573-579). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-25929-9_70.

Marmion, D. J., & Kordower, J. H. (2018). α-Synuclein nonhuman primate models of Parkinson's disease. *Journal of Neural Transmission*, *125*, 385-400. https://doi.org/10.1007/s00702-017-1720-0.

Mohanti, R. R. (2021). *Parkinson's XYZ data*. https://www.kaggle.com/datasets/guywhowantstolearnml/parkinsonsxyz (accessed Aug. 22, 2023).

Noor, M. B. T., Zenia, N. Z., Kaiser, M. S., Mamun, S. A., & Mahmud, M. (2020). Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of Alzheimer's disease, Parkinson's disease and schizophrenia. *Brain informatics*, *7*, 1-21. https://doi.org/10.1186/s40708-020-00112-2.

Park, Y. H., Suh, J. H., Kim, Y. W., Kang, D. R., Shin, J., Yang, S. N., & Yoon, S. Y. (2022). Machine learning based risk prediction for Parkinson's disease with nationwide health screening data. *Scientific Reports*, *12*(1), 19499. https://doi.org/10.1038/s41598-022-24105-9.

Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M., ... & Dean, J. (2018). Scalable and accurate deep learning with electronic health records. *NPJ digital medicine*, *1*(1), 1-10. https://doi.org/10.1038/s41746-018-0029-1.

Rehman, A., Saba, T., Mujahid, M., Alamri, F. S., & ElHakim, N. (2023). Parkinson's disease detection using hybrid LSTM-GRU deep learning model. *Electronics*, *12*(13), 2856. https://doi.org/10.3390/electronics12132856.

Sanchez-Hernandez, S. E., Salido-Ruiz, R. A., Torres-Ramos, S., & Roman-Godínez, I. (2022). Evaluation of feature selection methods for classification of epileptic seizure EEG signals. *Sensors*, *22*(8), 3066. https://doi.org/10.3390/s22083066.

Santos, J., Pallarès, I., & Ventura, S. (2022). Is a cure for Parkinson's disease hiding inside us?. *Trends in Biochemical Sciences*, *47*(8), 641-644. https://doi.org/10.1016/j.tibs.2022.02.001.

Sari, R., Omeiza, I. I., & Mwakifuna, M. A. (2023). The influence of number dice games in improving early childhood mathematical logic in early childhood education. *Interval: Indonesian Journal of Mathematical Education*, *1*(2), 61-66. https://doi.org/10.37251/ijome.v1i2.776.

Sayed, M. A., Cao, D. M., Islam, M. T., Tayaba, M., Pavel, M. E. U. I., Mia, M. T., ... & Sarkar, M. (2023). Parkinson's Disease Detection through Vocal Biomarkers and Advanced Machine Learning Algorithms. *Journal of computer science and technology studies*, *5*(4), 142-149. https://doi.org/10.32996/jcsts.2023.5.4.14.

Schrag, A., & Schott, J. M. (2006). Epidemiological, clinical, and genetic characteristics of early-onset parkinsonism. *The Lancet Neurology*, *5*(4), 355-363. https://doi.org/10.1016/S1474-4422(06)70411-2.

Shafiq, S., Ahmed, S., Kaiser, M. S., Mahmud, M., Hossain, M. S., & Andersson, K. (2022). Comprehensive Analysis of Nature-Inspired Algorithms for Parkinson's Disease Diagnosis. *IEEE Access*, *11*, 1629-1653. http://dx.doi.org/10.1109/ACCESS.2022.3232292.

Siddiqui, N. S., Klein, A., Godara, A., Buchsbaum, R. J., & Hughes, M. C. (2022). Predicting in-hospital mortality after acute myeloid leukemia therapy: through supervised machine learning algorithms. *JCO Clinical Cancer Informatics*, *6*, e2200044. https://doi.org/10.1200/CCI.22.00044.

Singh, N., Pillay, V., & Choonara, Y. E. (2007). Advances in the treatment of Parkinson's disease. *Progress in neurobiology*, *81*(1), 29-44. https://doi.org/10.1016/j.pneurobio.2006.11.009.

Suwarni, R. (2021). Analysis the process of observing class iv students in thematic learning in primary schools. *Tekno - Pedagogi : Jurnal Teknologi Pendidikan*, *11*(1), 26-32. https://doi.org/10.22437/teknopedagogi.v11i1.32717.

Taher, K. I., Abdulazeez, A. M., & Zebari, D. A. (2021). Data mining classification algorithms for analyzing soil data. *Asian Journal of Research in Computer Science*, *8*(2), 17-28. https://doi.org/10.9734/ajrcos/2021/v8i230196.

Ulandari, T., Ferry, D., & Damni, A. (2024). Symbiotic enlightenment: Exploring the fascination of biology education students with religion studies in an academic tapestry. *Integrated Science Education Journal*, *5*(3), 168-180. https://doi.org/10.37251/isej.v5i3.1122.

Yohanie, D. D., Botchway, G. A., Nkhwalume, A. A., & Arrazaki, M. (2023). Thinking process of mathematics education students in problem solving proof. *Interval: Indonesian Journal of Mathematical Education*, *1*(1), 24-29. https://doi.org/10.37251/ijome.v1i1.611.

Zakiyah, Z., Boonma, K., & Collado, R. (2024). Physics learning innovation: Song and animation-based media as a learning solution for mirrors and lenses for junior high school students. *Journal of Educational Technology and Learning Creativity*, *2*(2), 54-62. https://doi.org/10.37251/jetlc.v2i2.1062.

Zebari, R., Abdulazeez, A., Zeebaree, D., Zebari, D., & Saeed, J. (2020). A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction. *Journal of Applied Science and Technology Trends*, *1*(1), 56-70. https://doi.org/10.38094/jastt1224.