Journal of Resource Management and Decision Engineering

Volume 5, Issue 1, pp 1-11

Article history:
Received 18 June 2025
Revised 07 September 2025
Accepted 19 October 2025
Published online 01 January 2026

An Optimized and Robust Machine Learning Framework for Early Parkinson's Disease Prediction Using Speech Signals

Ghadeer. Aqil Ali[1], Leila. Sharifi[2], Parviz. Rashidi-Khazaee[3*], Hossein. Nahid-Titkanlue[4]

[1] Department of Electrical and Computer Engineering, Urmia University, Urmia, Iran.
[2] Assistant Professor, Department of Electrical and Computer Engineering, Urmia University, Urmia, Iran.
[3] Assistant Professor, Department of Information Technology and Computer Engineering, Urmia University of Technology, Urmia, Iran.
[4] Assistant Professor, Department of Industrial Engineering, Payame Noor University, Tehran, Iran.

*** Corresponding author email address**: p.rashidi@uut.ac.ir

Article Info

Article type:
Original Research

How to cite this article:
Aqil Ali, G., Sharifi, L., Rashidi-Khazaee, P. & Nahid-Titkanlue, H. (2026). An Optimized and Robust Machine Learning Framework for Early Parkinson's Disease Prediction Using Speech Signals. *Journal of Resource Management and Decision Engineering, 5*(1), 1-11.
https://doi.org/10.61838/kman.jrmde.5.1.187

ABSTRACT

With the rapid advancement of technologies in the present era, predicting Parkinson's disease (PD) early using non-invasive and low-cost methods, such as speech analysis with machine learning (ML) tools, remains a challenging task and lacks sufficient confidence for healthcare providers to use in daily practice. Therefore, this study presents an optimized early PD prediction tool and investigates its stability and robustness using a rigorous evaluation mechanism. For early PD prediction using speech signal data, the eXtreme Gradient Boosting (XGB) model is optimized using the Tree-structured Parzen Estimator (TPE) method and the Synthetic Minority Oversampling Technique (SMOTE) for solving the imbalanced dataset problem. Its performance was rigorously evaluated using an optimized strategy to ensure reliability and to earn the trust of clinicians for real-world operational use. To validate the model's trustworthiness and prediction capability, it was evaluated through 10 different runs of Stratified 10-Fold Cross Validation (SCV). The average measures of accuracy as 96.76%, precision as 97.70%, f1-score as 96.70%, recall as 95.91% and ROC-AUC 98.72% show great progress and performance in comparison with similar works. The model performance and stability were evaluated in many different situations and showed that the proposed model is stable and strong enough, and could be used as a practical tool in daily medical care. This tool brings the opportunity to be used easily as a decision support system through a website and detect PD early using patient voice signal with low cost in a non-invasive way that could be used remotely and easily.

Keywords: *XGBoost, Tree-structured Parzen Estimator, Data Augmentation, SMOTE, Decision Support System*

1. Introduction

In today's medical world, early diagnosis of diseases, before their clinical symptoms become acute, is very important (Almeida et al., 2019; Braga et al., 2019; Srinivasan et al., 2024). The main goal of such an approach is to prevent the rapid development of the disease, start early treatment, and, in the optimal case, treat the disease promptly. To achieve this goal, machine learning (ML) technologies have attracted the attention of experts and shown their capability to assist medical professionals in examination, diagnosis, treatment, and follow-up, and even go further, can help them in early diagnosis of the disease (Beriich et al., 2025; Islam et al., 2024; Mei et al., 2021; A. Reddy et al., 2024). Parkinson's disease (PD), in which 80% of the dopamine-producing brain cells were damaged (Schapira et al., 2017), is a neurodegenerative disease that is considered the second most common disease of the central nervous system (Poewe et al., 2017), and because of its rapid growth, it was described as Parkinson's pandemic (Dorsey et al., 2018).

There are several universal and non-universal issues in the diagnosis and treatment of PD: First, neurologists and movement specialists can diagnose the disease definitively only when they thoroughly check the patient's medical records and perform repeated scans and tests (Armstrong & Okun, 2020; Bloem et al., 2021). Which is challenging due to the time-consuming nature of tests/examinations and the subsequent costs. These challenges are exacerbated in developing countries due to the lack of knowledgeable specialists and up-to-date medical facilities (Ben-Shlomo et al., 2024). Second, although using standard clinical diagnostic tools can help determine the severity rate of PD and differentiate it from other neurological disorders (Zolin et al., 2025) But the lack of distinctive biomarkers and the similarity of symptoms with other diseases have made its diagnosis challenging, especially in the early stages (Almeida et al., 2019). Third, as the symptoms of the disease worsen and the costs of treatment and care increase exponentially (Lamba et al., 2022), many other organs (Aarsland et al., 2021; Kardan et al., 2024) than the brain is affected by PD. Therefore, early prediction/diagnosis of this disease is very cost-effective and is expected to provide greater life expectancy and well-being for older people (Cantürk & Günay, 2024; Kadhim et al., 2024; Patel[1] et al., 2025; Srinivasan et al., 2024; Yang et al., 2025).

Searching scientific databases shows that the application of ML in the early diagnosis of PD has increased significantly in the past few years. Researchers have used different ML-related approaches to achieve an optimal model. The main component of these approaches is using classifier algorithms like Artificial Neural Network (ANN) (Pahuja & Nagabhushan, 2021; Senturk, 2020), K-Nearest Neighbor (KNN) (Gupta et al., 2018; Lamba et al.; Pahuja & Nagabhushan, 2021; Sharma et al., 2019), Support Vector Machine (SVM) (Baruah et al., 2025; Pahuja & Nagabhushan, 2021; Senturk, 2020; Thirapanish et al., 2024), Deep Neural Network (DNN) (Jain et al., 2020; Kadam & Jadhav, 2018; Rahman et al., 2023), Random Forest (RF) (Baruah et al., 2025; Kadam & Jadhav, 2018; Sharma et al., 2019), eXtreme Gradient Boosting(XGB) (Balaha et al., 2025; H. Reddy et al., 2024), and Convolutional Neural Network (CNN) (Akila & Nayahi, 2024; Patel[1] et al., 2025; Saha & Nath, 2025).

To improve ML classifier performance, data augmentation techniques are enhanced ML classifier performance, especially in the case of low data size. Data augmentation is a collection of techniques for increasing the size of a dataset by generating new synthetic but similar data. The most widely used in PD early prediction is the Synthetic Minority Oversampling Technique (SMOTE) (Alshammri et al., 2023; Baruah et al., 2025; Jain et al., 2021; Lamba et al.; Saha & Nath, 2025), which leads to model performance improvement.

Tuning parameters and hyperparameters of ML models are a very important step in designing ML models. Different algorithms like particle swarm optimization (PSO) (Das et al., 2020), grid search cross validation(GSCV) (Alshammri et al., 2023; Baruah et al., 2025; Thirapanish et al., 2024), tree-structured parzen estimator(TPE) (Balaha et al., 2025) was successfully used in PD early prediction.

In this study, after evaluation of different data augmentation algorithms, the Synthetic Minority Oversampling Technique (SMOTE), for the first time, for solving the imbalanced PD dataset problem. Then, a new optimized XGB (Chen & Guestrin, 2016) using the TPE (Yang & Shami, 2020) was proposed for PD early prediction. These combinations bring the opportunity that the proposed model outperforms others and brings stability in PD prediction.

Another key distinction of our proposed model is the way of model evaluation. A rigid and rigorous model evaluation mechanism, section 2.6, was proposed for model performance evaluation in a way that clinicians can trust the model's prediction results and use it confidently in daily consultations.

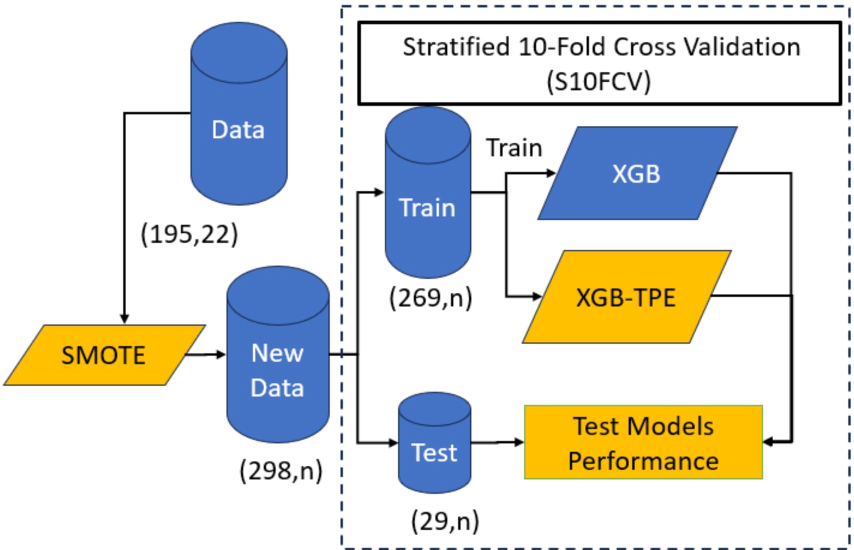
The remainder of this work is organized as methodology in section 2, results in section 3, discussion in section 4, and conclusions in section 5.

2. Methodology

In this study, to have a stable and strong PD early prediction model, an advanced practical model has been proposed. Figure 1 shows its structure. At first, the SMOTE (Alshammri et al., 2023; Baruah et al., 2025; Jain et al., 2021; Lamba et al.; Saha & Nath, 2025) was used to solve the PD dataset imbalance problem and generate new synthetic data. Second, stratified 10-Fold cross validation (S10FCV) was used to split the data into a training and testing set. Third, a new optimized XGB-TPE (Baqer & Rashidi-Khazaee, 2025) was proposed for PD early prediction. Forth, at each iteration, each model was evaluated on different unseen test data. Finally, the average of 10 different runs of S10FCV (A10S10FCV) evaluation metrics was considered as the model's overall performance measures. The proposed model is named SXT(SMOTE-XGB-TPE).

Figure 1

The structure of the proposed SXT model for PD prediction



2.1. Dataset Description

The collected voice speech signal data by the University of Oxford is publicly available in the UCI repository (Little et al., 2008) bring the opportunity to use non-invasive methods for the early detection of PD. The data was collected from 31 male and female individuals, aged between 46 and 85. Among the participants, 23 were diagnosed with PD, while the remaining eight served as the healthy control group. The total of 195 biomedical voice measurements is available in the dataset. Table 1 brings the detailed information about the dataset features and their meaning. Because of the fact that different feature has different data ranges, all input data was normalized using a standard scaler, and each feature has a mean of 0 and a standard deviation of 1.

Table 1

The Voice Speech PD Dataset Characteristics

#	Feature	Description	Category	Description
Input Variables				
1	MDVP:Fo(Hz)	Average vocal fundamental frequency	Fundamental Frequency	Measures of baseline vocal frequency
2	MDVP:Fhi(Hz)	Maximum vocal fundamental frequency		
3	MDVP:Flo(Hz)	Minimum vocal fundamental frequency		

4	Jitter (%)	Several measures of variation in fundamental frequency	Jitter Features	Measures of frequency variation
5	Jitter (Abs)	Measure of Jitter in absolute terms		
6	RAP	Measure of Rapid Jitter		
7	PPQ	Measure of Jitter using the PPQ method		
8	DDP	Measure of Jitter using the DDP method		
9	Shimmer	Several measures of variation in amplitude	Shimmer Features	Measures of amplitude variation
10	Shimmer(dB)	Shimmer in decibels		
11	APQ	Average perturbation quotient		
12	APQ3	Amplitude perturbation quotient measured in the first three instants		
13	APQ5	Amplitude perturbation quotient measured in the first five instants		
14	DDA	Shimmer—difference between the amplitudes of consecutive periods		
15	NHR	Noise-to-harmonics ratio	Noise-to-Harmonic Ratio	Measures of voice quality degradation
16	HNR	Harmonics-to-noise ratio	Harmonic-to-Noise Ratio (HNR)	Measures of voice quality degradation
17	RPDE	Recurrence period density entropy	Nonlinear Dynamical Complexity Measures	Features capturing fractal scaling and frequency variations
18	DFA	Signal fractal scaling exponent		
19	Spread1	Nonlinear measure of fundamental frequency variation		
20	Spread2	Another nonlinear measure of fundamental frequency variation		
21	D2	Correlation dimension		
22	PPE	Pitch period entropy		
Target Variable				
23	Status	Health status of the subject	1=Parkinson's (23 Patients), 0=healthy (8 Patients) Total Sample Count = 195 (147 Parkinson, 48 Healthy)	

*-The Marked feature was selected by the feature selection algorithm as the final model input.

2.2. *Data Augmentation Technique*

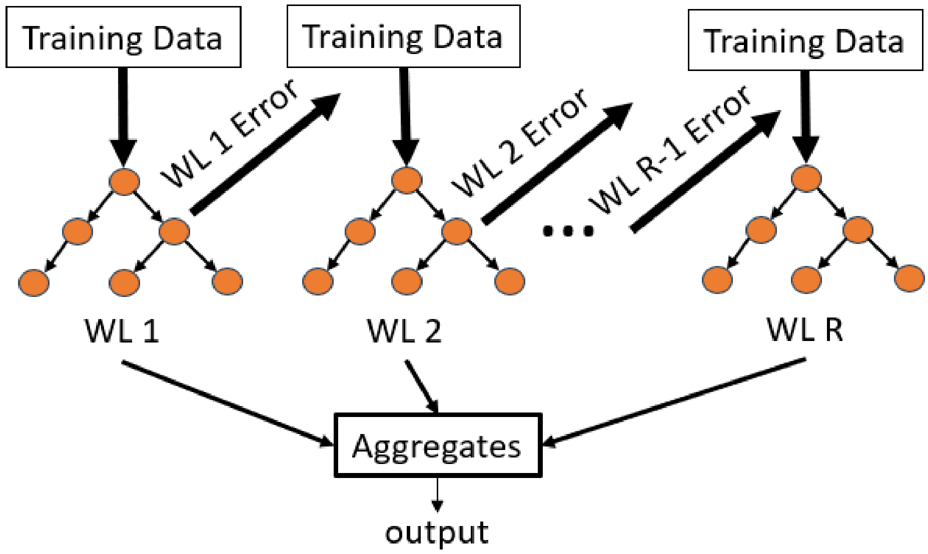
In this study, after evaluating different data augmentation techniques, the SMOTE method, with its distinctive ability to address the imbalanced dataset problem, was used for the first time to generate new synthetic data for PD. SMOTE (Alshammri et al., 2023; Baruah et al., 2025; Jain et al., 2021; Lamba et al.; Saha & Nath, 2025) is a well-known data augmentation approach designed to handle class imbalance in machine learning. Instead of merely duplicating minority class samples, SMOTE creates synthetic examples by interpolating between existing minority class instances and their nearest neighbors. This process increases the diversity of the minority class and leads to a more balanced dataset, ultimately enhancing the model's ability to learn from underrepresented cases. This approach is particularly valuable in medical diagnosis tasks, where rare but crucial cases could otherwise be neglected.

2.3. *XGBoost Model*

Figure 2 shows the general structure of the XGB Model (Chen & Guestrin, 2016). This model comprises a series of weak regression trees. The sequential structure of WL allows each learner to focus on correcting the error of the previous learner. By this, each WL tries to correct its previous WL error and make it possible for the whole result of the model to be strong. The optimal XGB tree structure, composed of weak learners (WLs), is constructed during training. After that, it predicts unseen samples.

Figure 2

The basic structure of the XGB model (Baqer & Rashidi-Khazaee, 2025)



XGB possesses unique features that enable it to outperform other models and establish itself as a state-of-the-art approach in various fields (Chen & Guestrin, 2016). XGB's key advantages include speed, high performance, scalability, customizability, the ability to handle missing values, and interpretability, all of which contribute to its success. However, XGB also has some drawbacks, such as computational complexity, overfitting, and a large number of hyperparameters requiring careful tuning. These limitations should be carefully considered when applying XGB.

2.4. Hyperparameter Tuning

Among different hyperparameter tuning methods like random search and grid search cross validation (GSCV), the TPE intelligently explore the search space, focusing on promising areas based on past evaluations. TPE outperforms other techniques in different areas, like energy prediction (Baqer & Rashidi-Khazaee, 2025). Bergstra et al. used the Bayesian theorem and proposed a TPE algorithm based on conditional probability distribution (Bergstra et al., 2013). The $p(x|y)$ is the conditional probability that calculates the probability of x occurrence in the condition of y. In this study, the goal is to find a suitable value for x^* as a prediction model hyperparameter, under the condition of low loss, which is denoted by y. At first, a threshold y^* for the loss was defined based on in-hand data, e.g., based on median data. The ultimate value of x^* will be the optimized value of the model hyperparameters. In this study, the original implementation of the TPE algorithm in the Hyperopt library was used (Bergstra et al., 2013).

2.5. Proposed Model Evaluation Strategy

To have a fair model evaluation, especially when the size of the dataset is small, it is necessary to evaluate models' performance in different conditions and test models' performance against every sample. The K-Fold Cross Validation (KFCV) is a common approach when the data size is small. In this study, the stratified K-Fold CV (SKFCV) was used for model performance evaluation. The main distinction between KFCV and SKFCV is that the Stratified ensures that each fold of test data has the same proportion of observations with a given label. In addition to S10FCV (K=10), the whole process was run 10 times using different random states, 10 prime numbers less than 200, to have different training and testing sets in each evaluation of S10FCV. When the dataset is imbalanced, using S10FCV makes sure every chunk of data reflects that imbalance properly—so the model gets a fairer test—while regular 10-fold splitting can shuffle things unevenly and throw off the results. By this was the stability of the model will also be checked. Finally, the average of 10 runs of S10FCV (A10S10FCV) was considered as a model performance measure. The Accuracy, Precision, Recall, F1-Score, and ROC-AUC were used as models' evaluation and comparison metrics.

3. Results

Initially, to solve the imbalance problem of the dataset, SMOTE generated new synthetic data and added it to the dataset. As a result, the positive and negative class distribution of the samples in the dataset (147,48) changed to (147,146).

The performance of the basic XGB and XGB models optimized using TPE (XGB-TPE) was investigated, and then the results of their combination with SMOTE methods for data augmentation and the results of their combination with optimal feature selection by MI are presented in Table 2. The stability and robustness of the proposed models and their performance were evaluated using A10S10FCV, and the average value of 10 different runs was considered as the final performance criterion of the model.

Table 2

The Prediction Performance of Different Models using A10S10FCV

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
XGB	92.15	93.21	97	94.92	96.81
SOMTE-XGB	95.47	96.31	94.77	95.41	99.23
SMOTE-XGB-TPE (SXT)	96.76	97.70	95.91	96.70	98.72

The results presented in Table 2 showed that SMOTE and TPE have improved the performance of the XGB model in early PD prediction. Based on various evaluation criteria, the proposed SXT model has higher precision, accuracy, and F1 score than other models, which are 97.70%, 96.76%, and 96.70%, respectively, and the XGB model has higher recall, 97%, and the SMOTE-XGB model has higher ROC-AUC criterion, 99.23%.

The results of 10 different runs of the proposed SXT model are presented in Table 3. As shown, the accuracy of S10FCV with random mode = 179 is 97.62%, which is higher than the average value of 96.77%. However, as can be seen in Table 2, the average value based on A10S10FCV is considered the evaluation criterion of the SXT model.

Table 3

The Proposed SXT Model Stability Checking Using S10FCV

Random State	Accuracy	Precision	Recall	F1-Score	AUC
17	97.28	98.62	95.90	97.14	98.47
29	96.93	96.79	97.29	96.96	98.89
37	96.60	98.20	95.19	96.51	98.56
42	96.93	97.37	96.62	96.94	99.25
53	96.59	98.08	95.19	96.47	98.74
89	95.57	95.88	95.33	95.56	98.27
101	97.61	98.71	96.57	97.53	98.76
139	96.26	97.33	95.24	96.20	98.49
179	97.62	98.08	97.24	97.58	99.24
199	96.28	97.95	94.57	96.20	98.57
Average	96.77	97.70	95.91	96.71	98.72

To compare the prediction accuracy of the proposed model in the present study with some published models (e.g., models evaluated using different training and test data splits, such as 80:20), the results of implementing the proposed SXT model with the best performance of S10FCV, with random mode = 179, are shown in Table 4. The presented results showed that the proposed model achieved 100% accuracy in different test folds.

Table 4

The Proposed SXT Model Best Run Performance based on S10FCV

Test Fold	Accuracy	Precision	Recall	F1-Score	AUC
1	100.00	100.00	100.00	100.00	100.00
2	96.67	100.00	93.33	96.55	100.00
3	96.67	93.75	100.00	96.77	99.11
4	96.67	93.75	100.00	96.77	100.00
5	96.55	100.00	93.33	96.55	96.67
6	100.00	100.00	100.00	100.00	100.00
7	100.00	100.00	100.00	100.00	100.00
8	96.55	100.00	92.86	96.30	99.05
9	96.55	100.00	92.86	96.30	99.05
10	96.55	93.33	100.00	96.55	98.57
Average	97.62	98.08	97.24	97.58	99.24

4. Discussions

To compare the proposed SXT model performance with previous published works, different factors were considered, including data augmentation method, type of classifier, and way in which the models were evaluated. Table 5 provides complete information about these methods and their comparison with the proposed SXT model.

Table 5

Comparison of the Proposed SXT PD Early Prediction Model with Other Works.

Work	DA	Classifier	Tuning Alg.	Testing Method	Best Acc (%)	Recall (%)
Lamba et al. (Lamba et al.)	SMOTE	Naïve Bayes, KNN, RF	-	10FCV	95.58	93.19
Jain et al. (Jain et al., 2021)	SMOTE	DNN	-	80:20	91.47	97.12
Alsham… et al. (Alshammri et al., 2023)	SMOTE	MLP	GSCV	70-30	98.31	96
Reddy et al. (H. Reddy et al., 2024)	SMOTE	RF, XGBoost	-	80-20	95	100
Saha et al (Saha & Nath, 2025)	SMOTE	Ensemble of PD-CNN		70:20	99.47 96.18*	98.19 95.63*
Baruah et al. (Baruah et al., 2025)	SMOTE	RF, LR, SVM	GSCV	5-FCV	97.44	100
Proposed SXT	SMOTE	XGB	TPE	A10S10FCV	96.76	95.91
				Best S10FCVRun	97.62	97.24
				Best Fold Testing	100	100

*- Our implementation performance of their work using A10S10FCV., # - Number of selected features

In most published works, including the present study, the performance of models has usually been compared only in terms of accuracy. While in the context of medical problems, accurate detection of positive cases is more important than the overall accuracy of the model. Therefore, it is better to compare the overall performance of models based on their sensitivity (recall). In analyzing and comparing the results of this criterion, it was found that the sensitivity of some models was not taken into account; however, comparing the sensitivity of the proposed SXT model with other models showed that in this criterion, the proposed model performs much better than other models.

The first main distinction of our work with others is using the A10S10FCV strategy for model evaluation, as described in section 2.6. By this way, we tested stability and reliability of proposed model, which are very important issue in designing ML models. As shown in Table 5, the method of evaluation was not specified in some studies. Other studies like (Alshammri et al., 2023; Saha & Nath, 2025) used the 70:30, and (Jain et al., 2021; H. Reddy et al., 2024) used the 80:20 strategy to train and evaluate model performance, but it is not clear whether they reported best results or the average of different executions. The presented result in Table 4 showed that the proposed SXT model, with a 90:10 data splitting strategy, had 100% accuracy in 3 different fold testing. Therefore, it could be said that it outperforms other published works tested using an 80(70):20(30) data splitting strategy. By considering A10S10FCV or similar strategies for model evaluation, only (Kadam & Jadhav, 2019) and (Balaha et al., 2025) have used a similar strategy for evaluation and reported accuracies of 93.84% and 95.67%, respectively, which is weaker than the performance of the proposed SXT model with an accuracy of 96.77%. In comparison with other works evaluated using 10-FCV or 5-FCV, it must be noted that only (Lamba et al.) used 10-FCV, and (Baruah et al., 2025) used 5-FCV. To have a fair comparison, the proposed SXT best run with random-state =179 was considered. As shown in Table 2, the SXT accuracy was 97.62% showing superiority over (Baruah et al., 2025; Lamba et al.) with 95.89%, 95.58% and 97.44% accuracy, respectively. The (Thirapanish et al., 2024) didn't report accuracy, but the reported recall, 88.75% is also weaker than the SXT recall, 98.79%.

In comparison with works that used XGB as a classifier, the (H. Reddy et al., 2024) used SMOTE for data augmentation and XGB for PD early prediction and evaluated their model once time based on 80:20 data splitting. Their work accuracy is reported as 95%, which is weaker than the proposed model result. Also, the (Balaha et al., 2025) used PSO for feature selection and TPE for hyperparameter tuning and used XGB and majority voting classifier for prediction, their XGB model accuracy was 90.20% and their best ensemble model accuracy was 95.67% which are weaker than the presented results in this study.

The performance of the proposed SXT prediction model shows that the combination of SMOTE as a data augmentation tool to solve the data imbalance problem, and XGB hyperparameter tuning using TPE has helped to create a more stable and predictable tool that can detect PD early and efficiently. The SMOTE, XGB, and TPE combination were used for the first time for PD early prediction. The proposed combination, as the SXT model, can be used as a reliable and practical tool by healthcare providers in daily operations.

The main limitation of this study is that the model was trained using the available dataset, which is small in size. Also, external validation is not possible because there is no similar dataset at hand. To have a general-purpose tool that can be used in daily care, it is necessary to train the model using a large dataset of individuals of different genders and ages from different countries and cultures to gain the trust of physicians for operational use as an intelligent assistant.

5. Conclusions

In order to have an efficient, reliable, stable, and usable medical decision support tool to assist care providers in PD early prediction, this study has proposed an advanced model based on XGB to provide a reliable and efficient method that is able to predict PD using speech signal processing. This method is low-cost and highly efficient compared to other methods, and does not require invasive medical intervention.

In this study, for the first time, a combination of SMOTE, XGB, and TPE (SXT) methods has been used to solve the problem of early prediction of PD. After preprocessing the speech signal data, the problem of low data size and an imbalanced dataset was investigated and balanced using the SMOTE technique. In order to increase the accuracy of the model, the parameters of the XGB model were optimized using the TPE method. The proposed combination has created a reliable model.

To test the stability of the model and examine its performance in different situations, the model was evaluated using the A10S10FCV method. In this strategy, the S10FCV process was run 10 times with different random states to accurately assess the stability and validity of the model. The presented results and the discussions conducted show that the proposed model, while having appropriate stability, also has higher efficiency than all the proposed models. It is recommended that the A10S10FCV strategy be used by all researchers/practitioners to evaluate the performance of models, especially when the data volume is small.

The proposed results showed that selecting a proper algorithm at each step of designing ML tools could have a great impact on the final model performance.

In future work, we will evaluate the SXT model's performance on other speech datasets. We will consider feature selection and other data augmentation techniques for PD early prediction. Also, we try to develop a web-based application that can receive the patient's voice by a Microphone, extract features, and use the proposed SXT algorithm to detect whether he/she healthy or sick.

Authors' Contributions

Authors contributed equally to this article.

Declaration

In order to correct and improve the academic writing of our paper, we have used the language model ChatGPT.

Transparency Statement

The collected voice speech signal data by the University of Oxford is publicly available in the UCI repository (Little et al., 2008).

Acknowledgments

We would like to express our gratitude to all individuals helped us to do the project.

Declaration of Interest

The authors report no conflict of interest.

Funding

According to the authors, this article has no financial support.

Ethics Considerations

In this research, ethical standards including obtaining informed consent, ensuring privacy and confidentiality were considered.

References

Aarsland, D., Batzu, L., Halliday, G. M., Geurtsen, G. J., Ballard, C., Ray Chaudhuri, K., & Weintraub, D. (2021). Parkinson disease-associated cognitive impairment. *Nature Reviews Disease Primers*, *7*(1), 47. https://pubmed.ncbi.nlm.nih.gov/34210995/

Akila, B., & Nayahi, J. J. V. (2024). Parkinson classification neural network with mass algorithm for processing speech signals. *Neural Computing and Applications*, *36*(17), 10165-10181. https://link.springer.com/article/10.1007/s10462-025-11347-y

Almeida, J. S., Rebouças Filho, P. P., Carneiro, T., Wei, W., Damaševičius, R., Maskeliūnas, R., & de Albuquerque, V. H. C. (2019). Detecting Parkinson's disease with sustained phonation and speech signals using machine learning techniques. *Pattern Recognition Letters*, *125*, 55-62. https://www.sciencedirect.com/science/article/pii/S0167865519301163

Alshammri, R., Alharbi, G., Alharbi, E., & Almubark, I. (2023). Machine learning approaches to identify Parkinson's disease using voice signal features. *Frontiers in Artificial Intelligence*, *6*, 1084001. https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2023.1084001/full

Armstrong, M. J., & Okun, M. S. (2020). Diagnosis and treatment of Parkinson disease: a review. *JAMA*, *323*(6), 548-560. https://pubmed.ncbi.nlm.nih.gov/32044947/

Balaha, H. M., Hassan, A. E.-S., Ahmed, R. A., & Balaha, M. H. (2025). Comprehensive multimodal approach for Parkinson's disease classification using artificial intelligence: insights and model explainability. *Soft Computing*, 1-33. https://dl.acm.org/doi/10.1007/s00500-025-10463-9

Baqer, N. R., & Rashidi-Khazaee, P. (2025). Residential Building Energy Usage Prediction Using Bayesian-Based Optimized XGBoost Algorithm. *IEEE Access*. https://ieeexplore.ieee.org/iel8/6287639/10820123/10900361.pdf

Baruah, D., Rehman, R., Bora, P. K., Mahanta, P., Dutta, K., & Konwar, P. (2025). Performance Evaluation of Classification Algorithms for Parkinson's Disease Diagnosis: A Comparative Study. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, *7*(3), 692-712. https://jeeemi.org/index.php/jeeemi/article/view/713

Ben-Shlomo, Y., Darweesh, S., Llibre-Guerra, J., Marras, C., San Luciano, M., & Tanner, C. (2024). The epidemiology of Parkinson's disease. *The lancet*, *403*(10423), 283-292. https://pubmed.ncbi.nlm.nih.gov/38245248/

Bergstra, J., Yamins, D., & Cox, D. (2013). Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. International conference on machine learning, https://proceedings.mlr.press/v28/bergstra13.html

Beriich, M., Ouhmida, A., Alouani, Z., Saleh, S., Cherradi, B., & Raihani, A. (2025). Advancing Parkinson's Disease Detection: A Review of AI and Deep Learning Innovations. 2025 5th International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), https://jglobal.jst.go.jp/en/detail?JGLOBAL_ID=202502252256845613

Bloem, B. R., Okun, M. S., & Klein, C. (2021). Parkinson's disease. *The lancet*, *397*(10291), 2284-2303. https://pubmed.ncbi.nlm.nih.gov/33848468/

Braga, D., Madureira, A. M., Coelho, L., & Ajith, R. (2019). Automatic detection of Parkinson's disease based on acoustic analysis of speech. *Engineering Applications of Artificial Intelligence*, *77*, 148-158. https://www.sciencedirect.com/science/article/abs/pii/S0952197618302045

Cantürk, İ., & Günay, O. (2024). Investigation of scalograms with a deep feature fusion approach for detection of Parkinson's disease. *Cognitive Computation*, *16*(3), 1198-1209.

https://link.springer.com/article/10.1007/s12559-024-10254-8

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. https://doi.org/10.1145/2939672.2939785

Das, P., Nanda, S., & Panda, G. (2020). Automated improved detection of Parkinson's disease using ensemble modeling. 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), https://www.proceedings.com/content/057/057976webtoc.pdf

Dorsey, E. R., Sherer, T., Okun, M. S., & Bloem, B. R. (2018). The emerging evidence of the Parkinson pandemic. *Journal of Parkinson's disease*, *8*(s1), S3-S8. https://pubmed.ncbi.nlm.nih.gov/30584159/

Gupta, D., Julka, A., Jain, S., Aggarwal, T., Khanna, A., Arunkumar, N., & de Albuquerque, V. H. C. (2018). Optimized cuttlefish algorithm for diagnosis of Parkinson's disease. *Cognitive Systems Research*, *52*, 36-48. https://www.sciencedirect.com/science/article/pii/S1389041718301876

Islam, M. A., Majumder, M. Z. H., Hussein, M. A., Hossain, K. M., & Miah, M. S. (2024). A review of machine learning and deep learning algorithms for Parkinson's disease detection using handwriting and voice datasets. *Heliyon*, *10*(3). https://www.sciencedirect.com/science/article/pii/S2405844024015007

Jain, D., Mishra, A. K., & Das, S. K. (2020). Machine learning based automatic prediction of Parkinson's disease using speech features. Proceedings of International Conference on Artificial Intelligence and Applications: ICAIA 2020, https://www.researchgate.net/publication/342640627_Machine_Learning_Based_Automatic_Prediction_of_Parkinson's_Disease_Using_Speech_Features

Jain, D., Mishra, A. K., & Das, S. K. (2021). Machine learning based automatic prediction of Parkinson's disease using speech features. Proceedings of International Conference on Artificial Intelligence and Applications: ICAIA 2020,

Kadam, V. J., & Jadhav, S. M. (2018). Feature ensemble learning based on sparse autoencoders for diagnosis of Parkinson's disease. In *Computing, Communication and Signal Processing: Proceedings of ICCASP 2018* (pp. 567-581). Springer. https://dl.acm.org/doi/abs/10.1007/s00521-021-05741-0

Kadam, V. J., & Jadhav, S. M. (2019). Feature ensemble learning based on sparse autoencoders for diagnosis of Parkinson's disease. Computing, Communication and Signal Processing: Proceedings of ICCASP 2018, https://dl.acm.org/doi/abs/10.1007/s00521-021-05741-0

Kadhim, M. N., Al-Shammary, D., & Sufi, F. (2024). A novel voice classification based on Gower distance for Parkinson disease detection. *International Journal of Medical Informatics*, *191*, 105583. https://www.sciencedirect.com/science/article/pii/S1386505624002466

Kardan, R., Nazari, M., Hemmati, J., Ahmadi, A., & Ashab, M. (2024). A Novel Therapeutic Strategy for Parkinson's Disease based on the Gut Microbiota: A Rreview Article [Review]. *Scientific Journal of Kurdistan University of Medical Sciences*, *29*(3), 127-138. https://doi.org/10.61186/sjku.29.3.11

Lamba, R., Gulati, T., Alharbi, H. F., & Jain, A. A hybrid system for Parkinson's disease diagnosis using machine learning techniques. *International Journal of Speech Technology*, 1-11. https://dl.acm.org/doi/10.4018/IJSI.292027

Lamba, R., Gulati, T., & Jain, A. (2022). A hybrid feature selection approach for parkinson's detection based on mutual information gain and recursive feature elimination. *Arabian Journal for Science and Engineering*, *47*(8), 10263-10276. https://www.springerprofessional.de/en/a-hybrid-feature-selection-approach-for-parkinson-s-detection-ba/20046808

Little, M., McSharry, P., Hunter, E., Spielman, J., & Ramig, L. (2008). Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. *Nature Precedings*, 1-1. https://pmc.ncbi.nlm.nih.gov/articles/PMC3051371/

Mei, J., Desrosiers, C., & Frasnelli, J. (2021). Machine learning for the diagnosis of Parkinson's disease: a review of literature. *Frontiers in Aging Neuroscience*, *13*, 633752. https://www.frontiersin.org/journals/aging-neuroscience/articles/10.3389/fnagi.2021.633752/full

Pahuja, G., & Nagabhushan, T. (2021). A comparative study of existing machine learning approaches for Parkinson's disease detection. *IETE Journal of Research*, *67*(1), 4-14. https://www.shs-conferences.org/articles/shsconf/ref/2022/09/shsconf_etltc2022_03027/shsconf_etltc2022_03027.html

Patel¹, N., Srividhya, R., Linda, P. E., & Rajesh¹, S. (2025). Parkinson's Insight: Leveraging CNN and LSTM Networks for Enhanced Diagnostic Accuracy. Proceedings of the International Conference on Advancements in Computing Technologies and Artificial Intelligence (COMPUTATIA 2025), https://www.atlantis-press.com/proceedings/computatia-25/126010054

Poewe, W., Seppi, K., Tanner, C. M., Halliday, G. M., Brundin, P., Volkmann, J., Schrag, A.-E., & Lang, A. E. (2017). Parkinson disease. *Nature Reviews Disease Primers*, *3*(1), 17013. https://doi.org/10.1038/nrdp.2017.13

Rahman, S., Hasan, M., Sarkar, A. K., & Khan, F. (2023). Classification of Parkinson's disease using speech signal with machine learning and deep learning approaches. *European Journal of Electrical Engineering and Computer Science*, *7*(2), 20-27. https://ejece.org/index.php/ejece/article/view/488

Reddy, A., Reddy, R. P., Roghani, A. K., Garcia, R. I., Khemka, S., Pattoor, V., Jacob, M., Reddy, P. H., & Sehar, U. (2024). Artificial intelligence in Parkinson's disease: Early detection and diagnostic advancements. *Ageing Research Reviews*, *99*, 102410. https://ejece.org/index.php/ejece/article/view/488

Reddy, H., Jagadeesh, D. V. S., Pati, P. B., & Kn, B. P. (2024). Parkinson's Disease Diagnosis from Patients Speech Analysis. 2024 IEEE 9th International Conference for Convergence in Technology (I2CT), https://www.semanticscholar.org/paper/Parkinson's-Disease-Diagnosis-from-Patients-Speech-HarshithaReddy-Aryagopal/e24528e3b84b9b2f6d65f7c8821d1bc9a9f16639

Saha, D. K., & Nath, T. D. (2025). A lightweight CNN-based ensemble approach for early detecting Parkinson's disease with enhanced features. *International Journal of Speech Technology*, 1-15. https://pubmed.ncbi.nlm.nih.gov/28592904/

Schapira, A. H., Chaudhuri, K. R., & Jenner, P. (2017). Non-motor features of Parkinson disease. *Nature Reviews Neuroscience*, *18*(7), 435-450. https://pubmed.ncbi.nlm.nih.gov/28592904/

Senturk, Z. K. (2020). Early diagnosis of Parkinson's disease using machine learning algorithms. *Medical Hypotheses*, *138*, 109603. https://pubmed.ncbi.nlm.nih.gov/32028195/

Sharma, P., Sundaram, S., Sharma, M., Sharma, A., & Gupta, D. (2019). Diagnosis of Parkinson's disease using modified grey wolf optimization. *Cognitive Systems Research*, *54*, 100-115. https://www.sciencedirect.com/science/article/abs/pii/S1389041718308726

Srinivasan, S., Ramadass, P., Mathivanan, S. K., Panneer Selvam, K., Shivahare, B. D., & Shah, M. A. (2024). Detection of Parkinson disease using multiclass machine learning approach. *Scientific reports*, *14*(1), 13813. https://www.nature.com/articles/s41598-024-64004-9

Thirapanish, W., Kantavat, P., Wanvarie, D., Chuangsuwanich, E., & Punyabukkana, P. (2024). Evaluating Machine Learning-Based Feature Selection Methods for Diagnosing Parkinson's Disease Under the SVM Framework. 2024 7th International Conference on Artificial Intelligence and Big Data (ICAIBD), https://www.researchgate.net/publication/382718440_Evaluating_Machine_Learning-Based_Feature_Selection_Methods_for_Diagnosing_Parkinson's_Disease_Under_the_SVM_Framework

Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, *415*, 295-316. https://doi.org/10.1016/j.neucom.2020.07.061

Yang, Z., Zhou, H., Srivastav, S., Shaffer, J. G., Abraham, K. E., Naandam, S. M., & Kakraba, S. (2025). Optimizing parkinson's disease prediction: A comparative analysis of data aggregation methods using multiple voice recordings via an automated artificial intelligence pipeline. *Data*, *10*(1), 4. https://www.mdpi.com/2306-5729/10/1/4

Zolin, A., Ooi, H., Zhou, M., Su, C., Wang, F., & Sarva, H. (2025). Liver fibrosis associated with more severe motor deficits in early Parkinson's disease. *Clinical Neurology and Neurosurgery*, *252*, 108861. https://scholar.google.com/citations?user=P4PgpD4AAAAJ&hl=en