Soft voting ensemble model to improve Parkinson's disease prediction with SMOTE

Jumanto Unjung [a,1,*], Rofik [a,2], Endang Sugiharti [a,3], Alamsyah [a,4], Riza Arifudin [a,5], Budi Prasetiyo [a,6], Much Aziz Muslim [a,b,7]

[a] Department of Computer Science, Universitas Negeri Semarang, Indonesia
[b] Faculty of Technology Management, Universiti Tun Hussein Onn Malaysia, Johor, Malaysia
[1] jumanto@mail.unnes.ac.id; [2] rofikn4291@students.unnes.ac.id; [3] endangsugiharti@mail.unnes.ac.id; [4] alamsyah@mail.unnes.ac.id;
[5] rizaarifudin@mail.unnes.ac.id; [6] bprasetiyo@mail.unnes.ac.id; [7] a212muslim@mail.unnes.ac.id
* corresponding author

ARTICLE INFO

Article history
Received June 16, 2024
Revised February 12, 2025
Accepted February 12, 2025
Available online February 28, 2025

Keywords
Soft-voting ensemble
SMOTE
Parkinson's disease prediction
Cross-validation

ABSTRACT

Parkinson's disease is one of the major neurodegenerative diseases that affect the central nervous system, often leading to motor and cognitive impairments in affected individuals. A precise diagnosis is currently unreliable, plus there are no specific tests such as electroencephalography or blood tests to diagnose the disease. Several studies have focused on the voice-based classification of Parkinson's disease. These studies attempt to enhance the accuracy of classification models. However, a major issue in predictive analysis is the imbalance in data distribution and the low performance of classification algorithms. This research aims to improve the accuracy of speech-based Parkinson's disease prediction by addressing class imbalance in the data and building an appropriate model. The proposed new model is to perform class balancing using SMOTE and build an ensemble voting model. The research process is systematically structured into multiple phases: data preprocessing, sampling, model development utilizing a voting ensemble approach, and performance evaluation. The model was tested using voice recording data from 31 people, where the data was taken from OpenML. The evaluation results were carried out using stratified cross-validation and showed good model performance. From the measurements taken, this study obtained an accuracy of 97.44%, with a precision of 97.95%, recall of 97.44%, and F1-Score of 97.56%. This study demonstrates that implementing the soft-voting ensemble-SMOTE method can enhance the model's predictive accuracy.

1. Introduction

Parkinson's disease (PD), a neurodegenerative disease that affects the central nervous system, is an increasingly pressing global health problem [1]. The disease often triggers motor and cognitive dysfunction in affected individuals. PD is a chronic and progressive condition where brain cells that produce dopamine are damaged [2]. The destruction of these cells causes irregular body movement due to a lack of dopamine, and about 70% to 80% of dopamine cells are damaged when the main symptoms of Parkinson's appear [3]. One of the symptoms is frequent shaking of the hands and feet while at rest [4]. In addition, Muscle stiffness, voice changes, slowed movements, and depression are also symptoms that can occur in people affected by this PD disease.

The disease often appears in the elderly [5] around 60 years old, although one in 20 samples showed major symptoms under 50 years. This causes many people to attribute these symptoms to age-related

changes. Parkinson's is a disease that is still untreatable [6]. However, existing medications can significantly reduce symptoms, especially in the early stages of the disease [3], [7].

Although the characteristics of Parkinson's disease have been discovered, early diagnosis in these cases is a complicated and time-consuming challenge [8]. Early and proper diagnosis is essential for more effective treatment and better care. However, the uncertainty associated with the early-stage diagnosis of Parkinson's disease has emerged as a critical concern in this study. The challenge arises from the difficulty in accurately diagnosing Parkinson's disease during its early stages, and accurate diagnoses tend to be late. There are many cases where patients only get diagnosed after the symptoms of the disease have progressed to a more serious level.

Numerous studies and medical reports have also expressed the urgency of detection in these cases. In recent decades, the disease outbreak has become the world's most common neurological health issue [9]. In the UK, for example, the prevalence of individuals diagnosed with Parkinson's disease continues to rise, and recent data suggests that more than 145,000 Britons are currently living with the disease. In Indonesia, the news of the increase in Parkinson's disease is also increasingly alarming. In some recent cases, it has been reported that in one family, almost all of them suffer from this disease [10]. It has also been projected that the prevalence of individuals affected by PD will continue to rise at an accelerated rate until the year 2050. This issue is expected to become a significant concern in many developed countries due to the substantial healthcare costs of managing this disease [11]. More accurate and efficient predictions can significantly benefit understanding, addressing, and processing the widespread Parkinson's disease outbreak.

Several prior studies have explored the application of voice analysis for the detection of Parkinson's disease, this is particularly relevant given that alterations in voice are among the most prominent clinical indicators of Parkinson's Disease [9], [12], [13]. This voice-based approach has also been considered quite effective in recent studies [14]–[17]. The results of these studies show that voice characteristics and human speech patterns can be used as potential indicators of Parkinson's disease. However, developing voice detection models that only utilize traditional tools is problematic.

Machine learning, including data mining methods, has emerged as a highly effective instrument in disease prediction [18]. These methods use data as an input tool for computers to process and make predictions based on mathematical and statistical calculations performed by machines. Data mining works by selecting, extracting, and modeling unknown hidden patterns from large data sets [19]. Thus, Machine Learning also allows computers to learn from the voice data converted to numeric from this Parkinson's disease prediction case. It can identify patterns associated with PD symptoms and make more accurate predictions. With computational capabilities and powerful algorithms, machine learning is a promising alternative to creating a good Parkinson's disease prediction model [20].

Several studies have focused on detecting Parkinson's disease through voice analysis. As done by Yaman *et al.* [21], it utilizes acoustic features by applying the weight feature technique for classification using the Support Vector Machine (SVM) method. This research shows quite high performance, where the accuracy reaches 91.25%. Furthermore, the research conducted by Ghaheri *et al.* [22] used SHAP and Hard Voting Ensemble methods for voice analysis to detect Parkinson's disease. They utilized Pearson's correlation coefficient to analyze the relationships among features, and their study achieved an accuracy rate of 85.42%.

Research conducted by Solana-Lavalle and Rosas-Romero [23] took an approach that differentiated Parkinson's Disease detection results according to the patient's gender. This study found different factors in Parkinson's disease detection depending on the patient's gender. The results of this study show that high-frequency sound content is very significant in helping Parkinson's disease detection in women. At the same time, low frequency is more effective in disease detection in men, and the accuracy rate obtained reaches 95.9% in women and 100% in men.

Meanwhile, Ahmed *et al.* [24] conducted a study that tested the intensity and spectrum of sound in Parkinson's patients by involving 6 machine learning algorithms. The results of this study show that the Random Forest algorithm achieves high accuracy, which is 97%. Research conducted by Sheibani *et al.*

[3] also predicted Parkinson's disease based on sound recordings. The research used an ensemble learning model, which combines the KNN, SVM, Decision Tree, and Naive Bayes algorithms. With the model built and the application of cross-validation, the research achieved the greatest accuracy of 90.65%. Unfortunately, few previous studies still apply the voting ensemble learning model and handle imbalanced classes appropriately. It shows the potential to maximize the performance obtained in predicting this disease. Therefore, this study aims to enhance the accuracy of Parkinson's disease prediction by implementing the soft voting ensemble learning model and SMOTE.

2. Method

The development of a Parkinson's disease detection model involved multiple stages, including data collection, data preprocessing, data oversampling, model construction, and performance evaluation. These stages are carried out sequentially. A detailed description of each stage of this research is presented in Fig. 1.

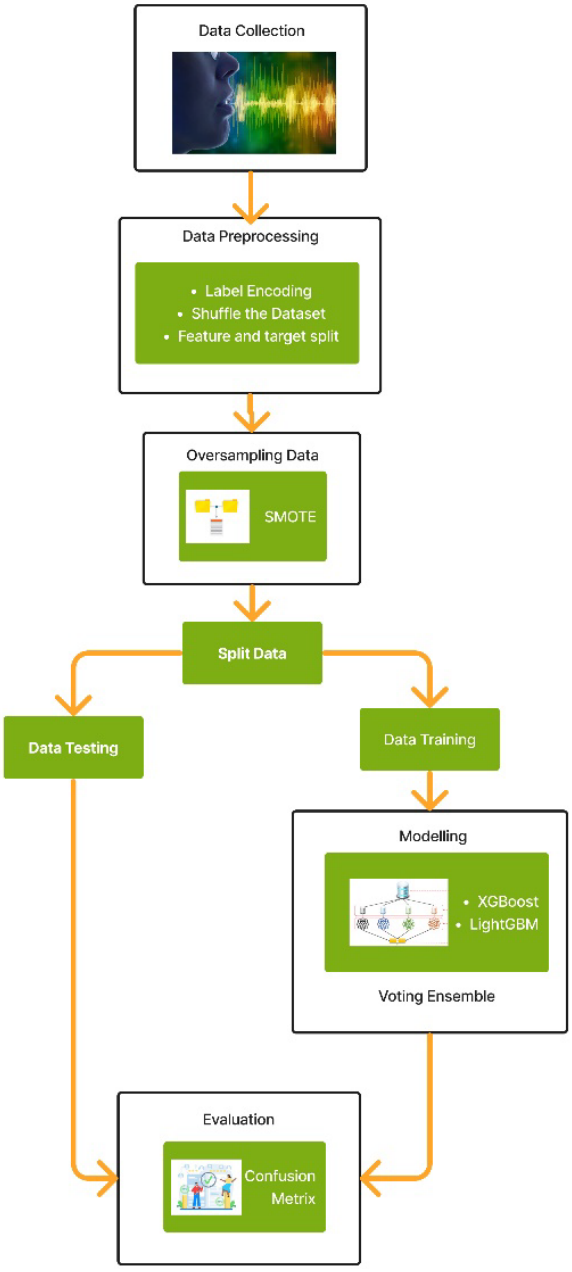

Fig. 1. Research flow

2.1. Data Collection

At this stage, a dataset is obtained, comprising a collection of voice recordings from individuals diagnosed with Parkinson's disease and healthy individuals. The dataset is obtained from a publicly available platform on OpenML. The dataset is accessible through the provided URL link: https://www.openml.org/search?type=data&status=active&id=1488&sort=runs.This dataset results from data collection involving voice recordings from 31 individuals. In the dataset, there are 23 individuals suffering from Parkinson's disease, with 16 of them being male and 7 females. Meanwhile, the other 8 people were healthy, with 3 males and 5 females. Each voice recording has been converted into numerical data presenting relevant voice attributes. The dataset consists of 48 records derived from voice recordings of healthy individuals and 147 records obtained from voice recordings of individuals diagnosed with Parkinson's disease. The dataset consists of 22 features and 1 label. Table 1 shows the details of the features in the dataset.

Table 1. Description of each feature in the dataset

Label	Feature Name	Description
V1	MDVP:Fo(Hz)	The mean fundamental frequency of the vocal voice
V2	MDVP:Fhi(Hz)	The maximum fundamental frequency of the vocal voice
V3	MDVP:Flo(Hz)	The minimum fundamental frequency of the vocal voice.
V4	MDVP:Jitter (%)	Various indicators of fluctuation in the fundamental frequency.
V5	MDVP:Jitter (Abs)	The fundamental frequency, particularly elevated in pathological sounds.
V6	MDVP:RAP	Disturbance in the relative amplitude within the Kay Pentax MDVP.
V7	MDVP:PPQ	The perturbation quotient measured over a five-point period in the Kay Pentax MDVP system.
V8	Jitter:DPP	The mean absolute difference of variances among cycles, normalized by the average period.
V9	MDVP:Shimmer	Several metrics for assessing amplitude variation.
V10	MDVP:Shimmer(dB)	The local shimmer measurement in decibels as analyzed using the Kay Pentax MDVP system.
V11	Shimmer:APQ3	Perturbation quotient derived from three-point amplitude.
V12	Shimmer:APQ5	Perturbation quotient calculated from five-point amplitude.
V13	MDVP:APQ	Perturbation quotient derived from eleven-point amplitude in the Kay Pentax MDVP.
V14	Shimmer:DDA	The mean absolute difference between the variances of amplitudes across successive periods.
V15	NHR	Two metrics indicating the proportion of noise to tonal components in the voval signals.
V16	HNR	HNR values typically exhibit smaller measurements in individuals with PD compared to those who are healthy.
V17	RPDE	Two measures of nonlinear dynamical complexity.
V18	D2	The relevance dimension (D2) quantifies the degree of irregularity present in the the system's reconstructed state space.
V19	DFA	Fractal scaling exponent of the signal.
V20	spread1	Three non-linear indicators assessing the variation in fundamental frequency.
V21	spread2	Variation
V22	PPE	Pitch period entropy.
Class	Status	Subject's health condition: 1 for Parkinson's, 0 for healthy.

2.2. Data Pre-Processing

At this stage, the dataset, which is already in numeric form and stored as a CSV file, is imported into Google Colab. After that, the separation between features and target is done. Label encoding is done to ensure that the data in the label is already in numeric form. The target in this dataset is identified by an attribute called 'Class', where the value '1' indicates an individual who is healthy, and '2' indicates an individual who is affected by PD. Shuffling is performed to randomize the data. This process is conducted

to ensure that the data utilized for training the model does not contain specific patterns that could influence its performance.

2.3. Data balancing using SMOTE

Oversampling is one of the techniques used to overcome the issue of disproportionate class distribution in datasets [25]. In this case, there is a class imbalance between samples of individuals with PD and individuals who are healthy. The sample of individuals with PD is larger than that of healthy individuals. Therefore, oversampling using the Synthetic Minority Oversampling Technique (SMOTE) is used to overcome this issue by creating synthetic samples from the minority class [26], namely healthy individuals. Oversampling the sample of healthy individuals prevents the model from focusing too much on the sample of PD sufferers and underlearning the sample of healthy individuals. This can lead to mispredictions where individuals predicted to be in the healthy class are instead classified as PD sufferers.

2.4. Modelling

In the modeling stage, a classification model is developed to identify individuals with PD based on the extracted voice features from the dataset. There are 2 classification models used, namely XGBoost, and LightGBM. XGBoost is utilized to manage high-dimensional classification tasks and enhance performance effectively [27]. XGBoost is essentially a combined machine-learning system that relies on decision tree-based models and uses a gradient-boosting framework [28]. The algorithm functions by constructing decision trees during the training phase, simultaneously minimizing the error function to achieve optimal efficiency and scalability [29]. This research ensures the model can understand the data well by oversampling first. LightGBM is also a popular [30] and efficient performing classification model [31]–[35], which is also utilized in this research. This model exhibits advantages in terms of computational efficiency and its capability to process large-scale datasets with high-dimensional features [36]. LightGBM utilizes parallel multi-threaded histogram-based techniques to accelerate the training phase and preprocess data through gradient-based one-sided sampling and exclusive feature bundling, thereby improving computational efficiency [37].

The Voting Ensemble method [38]–[40] is a powerful machine-learning technique that integrates several models' predictive capabilities to enhance the model's performance [41]. This method predicts by taking the most votes from the base mode used [42], [43]. The method used in ensemble voting is soft voting. Soft Voting considers the probability score or confidence level given by each model. The confidence scores of the predicted results of the base models, namely XGBoost and LightGBM, are taken, and the final decision is determined by aggregating the computed scores. So, if the XGBoost and LightGBM algorithms predict the opposite class, the algorithm's decision with the highest confidence probability is taken. The proposed model is illustrated in Fig. 2. The application of this method was carried out because, in some previous studies, it was able to improve the predictive results of [44] and overcome the weaknesses of individual models.

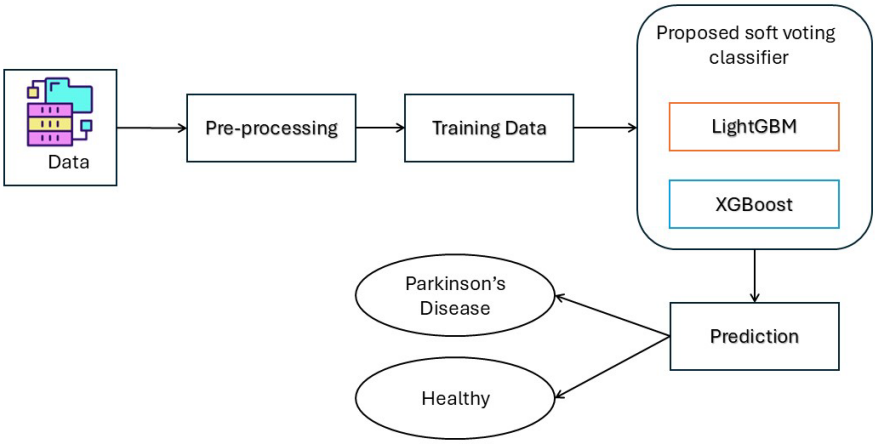

Fig. 2. Proposed soft voting classification method

The dataset was previously partitioned into training and testing sets using an 80:20 ratio. Training is carried out using the Stratified K-fold Cross Validation technique to be able to train. This method divides a dataset into subsets and then iteratively trains and tests the model. Stratified is applied to ensure the proportion of classes in each fold [45]. Then after going through the training stage, the data is tested using testing data that the model has never seen before.

2.5. Model Evaluation

The model evaluation is conducted utilizing a confusion matrix. A confusion matrix is a technique that can describe how well the model predicts positive and negative cases and how the model can make correct predictions. The confusion matrix consists of 4 main metrics. The confusion matrix comprises four primary metrics. The first metric, True Positive (TP), represents the number of actual positive instances correctly classified by the model. True Negative (TN) refers to the number of actual negative instances accurately identified by the model. False Positive (FP) denotes the number of negative instances incorrectly classified as positive. Conversely, False Negative (FN) indicates the number of positive instances that were misclassified as negative by the model. This confusion matrix calculates model performance through accuracy, precision, recall, and F1-Score. The calculations for model performance are outlined in Formulas 1 to 4.

- Accuracy

  Accuracy is a metric that quantifies the overall capability of a model to classify instances across different classes correctly. Accuracy is measured by the equation (1).

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (1)$$

- Precision

  Precision is a value that indicates the extent to which the model can predict positive cases correctly. Which is useful for understanding the extent to which the model's positive predictions are reliable. Precision is measured by equation (2).

$$Precision = \frac{TP}{(TP+FP)} \quad (2)$$

- Recall

  Recall is a value that indicates the extent to which the model is successful in identifying all positive cases. Recall is measured by equation (3).

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

- F1-score

  F1-score is a combined metric of precision and recall used to measure overall model performance. F1-score is measured by equation (4).

$$F1 - Score = \frac{2*(Precision*Recall)}{(Precision+Recall)} \quad (4)$$

3. Results and Discussion

Parkinson's disease detection results from a process involving several important steps in data analysis. The first is data collection. The dataset used comes from the publicly accessible OpenML platform, which consists of voice recordings of 31 people, with 23 of them suffering from PD and the other 8 in good health. The dataset consists of 195 voice data recordings, comprising 23 existing features. The second stage is Data Preprocessing. Label encoding was performed, where previously, the label features contained data that was not yet numeric. Data shuffling is performed to ensure that the data does not

have a certain pattern that can interfere with the prediction process by the model. The third stage is class balancing using SMOTE. To prevent the model from being too skewed to one of the classes. The fourth stage is modeling, which involves creating an ensemble voting learning model from the basic XGBoost and LightGBM models. The data is trained using Stratified K-fold cross-validation. Ensuring that the model performs training and testing on each piece of data. The training data is divided into 5 folds, each fold equally contains each class. Then the fifth stage is model evaluation using a confusion matrix, which can produce performance in terms of accuracy, precision, recall, and f1-score of the model.

To see how the features in the dataset correlate (see Fig. 3), a heatmap was created to illustrate the degree of correlation between features. Some features show a significant positive correlation, which indicates that a change in one feature will result in a change in another feature, and these changes are directly proportional. Conversely, other features show a strong negative correlation, indicating that an increase in one feature corresponds to a decrease in another.

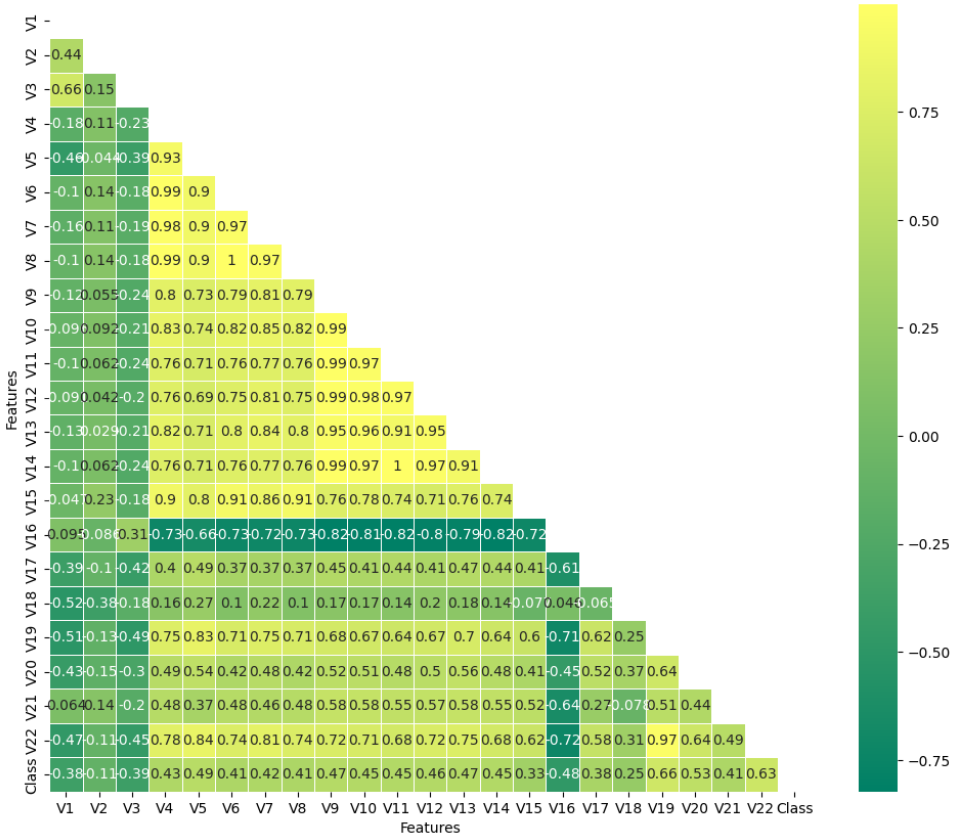

Fig. 3. Relationship between features in the dataset

The heatmap identifies features with a strong positive correlation and a strong negative correlation with the target. Based on the heatmap, the features with a strong positive correlation with the target are feature v19, with a correlation level of 0.57; feature v22, which has a correlation of 0.54; and feature v20, which correlates with 0.46. This means that when these features increase, the probability of the target class will also increase, and vice versa. Features that have a fairly strong negative correlation with the target are v1, with a negative correlation level of -0.4, v16 correlation of -0.39, and v3 correlation of -0.38. This negative correlation indicates that when the feature value of this feature increases, the probability of the target class will decrease and vice versa. In addition, the heatmap also shows the correlation between features, such as in v8 and v6. It features v14 and v11, which have a very strong positive correlation of 1, indicating that they move hand in hand.

Data balancing is done with SMOTE, which is believed to be a good method, as it synthesizes data from minority classes. Instead of duplicating the data. The result of SMOTE can balance the data well. The change in data before and after SMOTE can be seen in Fig. 4 and Fig. 5, respectively.

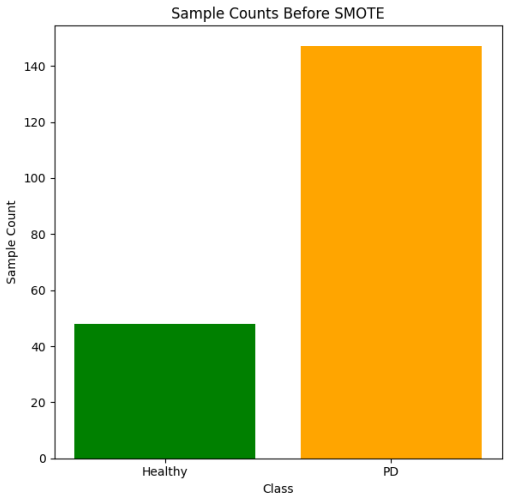

Fig. 4. Comparison between classes on the dataset

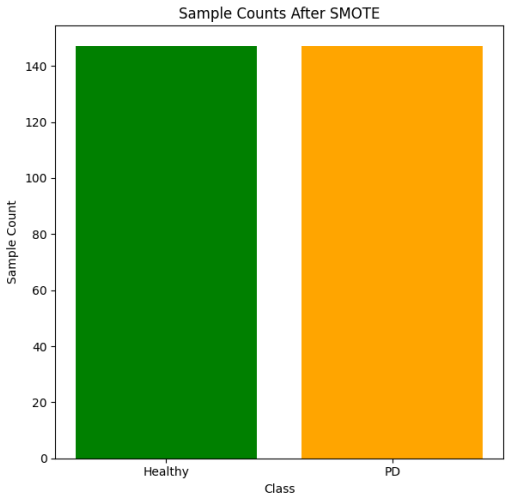

Fig. 5. Comparison between classes on the dataset after SMOTE

Before SMOTE, there is a significant difference in the amount of data between healthy people and people with PD. The data was divided into training data and testing data. It turns out that with a division of 80% and 20%, the training data for the class of healthy individuals is 44 and people with PD is 112. If class balancing is not done, of course the model that will be created will lean towards the class of individuals with PD, which will allow misclassification. This happens because the model learns less from healthy individual data or negative classes. So SMOTE is carried out on the training data. so that the training data has several healthy individual data and individuals with PD each 112 data records. Based on how SMOTE works, the newly generated data added to the training dataset is the result of data synthesis from the negative class or healthy individual data. Not by duplicating it. SMOTE results show that the data becomes balanced between the 2 classes so that the model can learn in a balanced and fair way from the 2 classes.

After the data is balanced, the ensemble soft voting model is built using the XGBoost and LightGBM base models. The model is trained by applying Stratified K-Fold Cross Validation, to ensure that the model works well not because it happens to get easy testing data or vice versa. The data division is set to 5 folds. After the model has been trained, the model is tested on the testing data. Fig. 6 illustrates the model performance during training and testing. When training using CV, the model shows a significant change in performance. This can be seen from the accuracy value obtained per fold which experiences ups and downs.

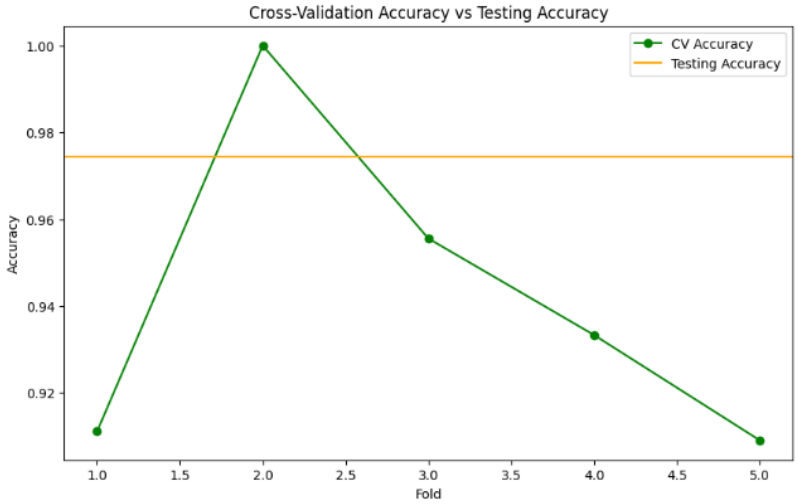

Fig. 6. The Comparison of accuracy during model training and testing

When testing accuracy on training data that has not applied CV, the resulting accuracy is 100%. This indicates that the model is overfitting. However, when stratified K-fold cross-validation was applied, the model gave an average accuracy of 94.18%. When compared to the testing data of 97.44%, it can be said that the model has good performance. This shows that the model is good enough to generalize between positive and negative classes and individuals with PD and healthy conditions. The precision value shows the model's performance, which is 97.95%. The recall is 97.44%, and F1-Score is 97.56%. From these results, it can be said that the model that has been built can distinguish the target class very well and can be relied upon to detect positive cases (Parkinson's disease classification) with a low error rate. And this also shows that the model created can potentially have real-world applications in detecting Parkinson's disease. Table 2 compares the model results from this study with previous research on the same topic, specifically Parkinson's disease prediction.

Table 2. Comparative performance of the proposed method and existing related studies

Similar Study	Method	Performance
Sheibani *et al.* [3]	Using the ensemble approach, with internal classifications KNN, SVM, and NB. The Ultimate Classification utilizes MLP, AB, Voting, and RF.	Accuracy Ensemble model: 90.6%
Avuçlu and Elen [5]	Random Forest with Statistical measurement results for 75 × 25.	Accuracy Random Forest: 85.81%
Solana-Lavalle and Rosas-Romero [23]	A comparative analysis of the Extreme Gradient Boosting (XGB) classifier, Logistic Regression classifier, Stochastic Gradient Descent (SGD) classifier, Random Forest classifier, Decision Tree (DT) classifier, and K-Nearest Neighbor (KNN) classifier.	Accuracy Random Forest: 97%
Peker and Kubat [46]	The ensemble method, utilizing stratified 10-fold cross-validation, integrates four discretization algorithms: Chi2, ChiMerge (ChiM), Modified Chi2 (ModChi2), and Extended Chi2 (ExtChi2).	Accuracy Ensemble model: 88.03%
Despotovic *et al.* [47]	Gaussian Processes, enhanced with Automatic Relevance Determination (ARD), are applied using 10-fold cross-validation	Accuracy Gaussian model: 96.92%
Proposed Method	**Utilizing ensemble voting method with XGBoost and LightGBM algorithms, SMOTE data balancing, and Cross-validation.**	**Accuracy Ensemble Voting: 97.44%**

Based on the comparative analysis of model performance between the proposed method and prior related studies, the soft voting ensemble approach demonstrated superior performance, achieving an accuracy of 97.44%. These findings indicate that the soft voting ensemble combining XGBoost and LightGBM with SMOTE is the most promising approach to evaluating performance metrics. Furthermore, this study contributes to the body of knowledge by providing a computational framework for evaluating experimental datasets. Additionally, it offers a comparative study of different machine learning models that can be utilized for predicting Parkinson's disease.

Implementing the SMOTE technique enhances data synthesis while preserving class classification accuracy within the generated dataset. Moreover, using stratified k-fold cross-validation in the training phase does not contribute to significant variations in performance outcomes. This study develops a soft voting ensemble model and conducts extensive evaluations to assess its effectiveness. However, further comprehensive investigations are necessary to verify the model's ability to generalize new data. It is important to recognize that the optimal results obtained may be influenced by the dataset's limited size, particularly due to the scarcity of instances in the healthy class. Future research is recommended to validate the model's performance using larger and more diverse datasets.

4. Conclusion

This study aims to predict Parkinson's disease using the Soft Voting Ensemble model. We measure the effectiveness of the ensemble model formed by two base models, namely XGBoost and LightGBM. We also address the issue of class imbalance between healthy individuals and those with Parkinson's

disease by applying the SMOTE technique. The training was performed with Startifid K-Fold Cross Validation for 5 folds. Model evaluation was performed using the confusion matrix. This study has achieved promising results, with 97.44% accuracy, 97.95% precision, 97.44% recall, and 97.56% F1 score. These performance results show that our model can effectively distinguish between healthy individuals and those with Parkinson's disease. Thus, this model can potentially be an effective tool for disease detection in clinical practice.

Acknowledgment

We express our gratitude for the support provided by the Budgetary Execution Checklist (DIPA) of the Faculty of Mathematics and Natural Sciences (FMIPA), Universitas Negeri Semarang, Indonesia, in funding this research through the grant (Grant No. DPA 40.14.4/UN37/PPK.4.4/2022).

Declarations

Author contribution. Jumanto was responsible for the conceptualization, methodology, and preparation of the original draft. Rofik contributed to the original draft, developing the program, and testing. Sugiharti was involved in writing, editing, and supervision. Alamsyah contributed to writing and supervision. Arifudin provided supervision and editing. Prasetiyo participated in writing and editing. Muslim was engaged in writing, editing, and supervision.

Funding statement. This research was funded by FMIPA Universitas Negeri Semarang for Grant Contract Number DPA 40.14.4/UN37/PPK.4.4/2022 Year 2022.

Conflict of interest. The authors affirm that there are no conflicts of interest associated with this research.

Additional information. No further information is available regarding this paper.

References

[1] E. J. Alqahtani, F. H. Alshamrani, H. F. Syed, and S. O. Olatunji, "Classification of Parkinson's Disease Using NNge Classification Algorithm," in *2018 21st Saudi Computer Society National Computer Conference (NCC)*, Apr. 2018, pp. 1–7, doi: 10.1109/NCG.2018.8592989.

[2] Z. Karapinar Senturk, "Early diagnosis of Parkinson's disease using machine learning algorithms," *Med. Hypotheses*, vol. 138, p. 109603, May 2020, doi: 10.1016/j.mehy.2020.109603.

[3] R. Sheibani, E. Nikookar, and S. Alavi, "An ensemble method for diagnosis of Parkinson's disease based on voice measurements," *J. Med. Signals Sensors*, vol. 9, no. 4, p. 221, Oct. 2019, doi: 10.4103/jmss.JMSS_57_18.

[4] Y. Liu, Y. Li, X. Tan, P. Wang, and Y. Zhang, "Local discriminant preservation projection embedded ensemble learning based dimensionality reduction of speech data of Parkinson's disease," *Biomed. Signal Process. Control*, vol. 63, p. 102165, Jan. 2021, doi: 10.1016/j.bspc.2020.102165.

[5] E. Avuçlu and A. Elen, "Evaluation of train and test performance of machine learning algorithms and Parkinson diagnosis with statistical measurements," *Med. Biol. Eng. Comput.*, vol. 58, no. 11, pp. 2775–2788, Nov. 2020, doi: 10.1007/s11517-020-02260-3.

[6] R. Mathur, V. Pathak, and D. Bandil, "Emerging Trends in Expert Applications and Security," in *ICE-TEAS 2024*, 2024, vol. 841, pp. 357–363, doi: 10.1007/978-981-13-2285-3.

[7] K. M. Alalayah, E. M. Senan, H. F. Atlam, I. A. Ahmed, and H. S. A. Shatnawi, "Automatic and Early Detection of Parkinson's Disease by Analyzing Acoustic Signals Using Classification Algorithms Based on Recursive Feature Elimination Method," *Diagnostics*, vol. 13, no. 11, p. 1924, May 2023, doi: 10.3390/diagnostics13111924.

[8] A. U. Haq *et al.*, "Feature Selection Based on L1-Norm Support Vector Machine and Effective Recognition System for Parkinson's Disease Using Voice Recordings," *IEEE Access*, vol. 7, pp. 37718–37734, 2019, doi: 10.1109/ACCESS.2019.2906350.

[9] R. Khaskhoussy and Y. Ben Ayed, "Improving Parkinson's disease recognition through voice analysis using deep learning," *Pattern Recognit. Lett.*, vol. 168, pp. 64–70, Apr. 2023, doi: 10.1016/j.patrec.2023.03.011.

[10] S. K. Bhasin and I. U. Bharadwaj, "Perceptions and meanings of living with Parkinson's disease: an account of caregivers lived experiences," *Int. J. Qual. Stud. Health Well-being*, vol. 16, no. 1, Jan. 2021, doi: 10.1080/17482631.2021.1967263.

[11] S. Aich, H.-C. Kim, K. Younga, K. L. Hui, A. A. Al-Absi, and M. Sain, "A Supervised Machine Learning Approach using Different Feature Selection Techniques on Voice Datasets for Prediction of Parkinson's Disease," in *2019 21st International Conference on Advanced Communication Technology (ICACT)*, Feb. 2019, vol. 2019-Febru, pp. 1116–1121, doi: 10.23919/ICACT.2019.8701961.

[12] J. Jumanto *et al.*, "Optimizing Support Vector Machine Performance for Parkinson's Disease Diagnosis Using GridSearchCV and PCA-Based Feature Extraction," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 10, no. 1, pp. 38–50, Feb. 2024, doi: 10.20473/jisebi.10.1.38-50.

[13] C. O. Sakar *et al.*, "A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform," *Appl. Soft Comput.*, vol. 74, pp. 255–263, Jan. 2019, doi: 10.1016/j.asoc.2018.10.022.

[14] P. Klumpp *et al.*, "The phonetic footprint of Parkinson's disease," *Comput. Speech Lang.*, vol. 72, p. 101321, Mar. 2022, doi: 10.1016/j.csl.2021.101321.

[15] T. Khan *et al.*, "Assessing Parkinson's disease severity using speech analysis in non-native speakers," *Comput. Speech Lang.*, vol. 61, p. 101047, May 2020, doi: 10.1016/j.csl.2019.101047.

[16] J. M. Tracy, Y. Özkanca, D. C. Atkins, and R. Hosseini Ghomi, "Investigating voice as a biomarker: Deep phenotyping methods for early detection of Parkinson's disease," *J. Biomed. Inform.*, vol. 104, p. 103362, Apr. 2020, doi: 10.1016/j.jbi.2019.103362.

[17] J. Unjung and M. R. Ningsih, "Optimized Handwriting-based Parkinson's Disease Classification Using Ensemble Modeling and VGG19 Feature Extraction," *Sci. J. Informatics*, vol. 10, no. 4, pp. 489–498, Nov. 2023. [Online]. Available at: https://journal.unnes.ac.id/nju/sji/article/view/47108.

[18] L. Sigcha *et al.*, "Deep learning and wearable sensors for the diagnosis and monitoring of Parkinson's disease: A systematic review," *Expert Syst. Appl.*, vol. 229, p. 120541, Nov. 2023, doi: 10.1016/j.eswa.2023.120541.

[19] H. Syamsudin, S. Khalidah, and J. Unjung, "Lepidoptera Classification Using Convolutional Neural Network EfficientNet-B0," *Indones. J. Artif. Intell. Data Min.*, vol. 7, no. 1, p. 47, Nov. 2023, doi: 10.24014/ijaidm.v7i1.24586.

[20] S. Saravanan *et al.*, "A Systematic Review of Artificial Intelligence (AI) Based Approaches for the Diagnosis of Parkinson's Disease," *Arch. Comput. Methods Eng.*, vol. 29, no. 6, pp. 3639–3653, Oct. 2022, doi: 10.1007/s11831-022-09710-1.

[21] O. Yaman, F. Ertam, and T. Tuncer, "Automated Parkinson's disease recognition based on statistical pooling method using acoustic features," *Med. Hypotheses*, vol. 135, p. 109483, Feb. 2020, doi: 10.1016/j.mehy.2019.109483.

[22] P. Ghaheri, H. Nasiri, A. Shateri, and A. Homafar, "Diagnosis of Parkinson's disease based on voice signals using SHAP and hard voting ensemble method," *Comput. Methods Biomech. Biomed. Engin.*, vol. 27, no. 13, pp. 1858–1874, Oct. 2024, doi: 10.1080/10255842.2023.2263125.

[23] G. Solana-Lavalle and R. Rosas-Romero, "Analysis of voice as an assisting tool for detection of Parkinson's disease and its subsequent clinical interpretation," *Biomed. Signal Process. Control*, vol. 66, p. 102415, Apr. 2021, doi: 10.1016/j.bspc.2021.102415.

[24] I. Ahmed, S. Aljahdali, M. Shakeel Khan, and S. Kaddoura, "Classification of Parkinson Disease Based on Patient's Voice Signal Using Machine Learning," *Intell. Autom. Soft Comput.*, vol. 32, no. 2, pp. 705–722, Nov. 2022, doi: 10.32604/iasc.2022.022037.

[25] T. D.K., P. B.G, and F. Xiong, "Auto-detection of epileptic seizure events using deep neural network with different feature scaling techniques," *Pattern Recognit. Lett.*, vol. 128, pp. 544–550, Dec. 2019, doi: 10.1016/j.patrec.2019.10.029.

[26] S. Davoudi, A. Ahmadi, and M. R. Daliri, "Frequency–amplitude coupling: a new approach for decoding of attended features in covert visual attention task," *Neural Comput. Appl.*, vol. 33, no. 8, pp. 3487–3502, Apr. 2021, doi: 10.1007/s00521-020-05222-w.

[27] M. Huu Nguyen, T.-A. Nguyen, and H.-B. Ly, "Ensemble XGBoost schemes for improved compressive strength prediction of UHPC," *Structures*, vol. 57, p. 105062, Nov. 2023, doi: 10.1016/j.istruc.2023.105062.

[28] C. Jatoth, N. E., M. A.V.R., and S. R. Annaluri, "Effective monitoring and prediction of Parkinson disease in Smart Cities using intelligent health care system," *Microprocess. Microsyst.*, vol. 92, p. 104547, Jul. 2022, doi: 10.1016/j.micpro.2022.104547.

[29] Y. Huang, Y. Wang, P. Wang, and Y. Lai, "An XGBOOST predictive model of void ratio in sandy soils with shear-wave velocity as major input," *Transp. Geotech.*, vol. 42, p. 101100, Sep. 2023, doi: 10.1016/j.trgeo.2023.101100.

[30] J. Ye, B. Zhao, and H. Deng, "Photovoltaic Power Prediction Model Using Pre-train and Fine-tune Paradigm Based on LightGBM and XGBoost," *Procedia Comput. Sci.*, vol. 224, pp. 407–412, Jan. 2023, doi: 10.1016/j.procs.2023.09.056.

[31] Z. Cui, X. Qing, H. Chai, S. Yang, Y. Zhu, and F. Wang, "Real-time rainfall-runoff prediction using light gradient boosting machine coupled with singular spectrum analysis," *J. Hydrol.*, vol. 603, p. 127124, Dec. 2021, doi: 10.1016/j.jhydrol.2021.127124.

[32] L. Tian, L. Feng, L. Yang, and Y. Guo, "Stock price prediction based on LSTM and LightGBM hybrid model," *J. Supercomput.*, vol. 78, no. 9, pp. 11768–11793, Jun. 2022, doi: 10.1007/S11227-022-04326-5.

[33] X. Wang, N. Xu, X. Meng, and H. Chang, "Prediction of Gas Concentration Based on LSTM-LightGBM Variable Weight Combination Model," *Energies*, vol. 15, no. 3, p. 827, Jan. 2022, doi: 10.3390/en15030827.

[34] Y. Dasril, M. A. Muslim, M. F. Al Hakim, J. Jumanto, and B. Prasetiyo, "Credit Risk Assessment in P2P Lending Using LightGBM and Particle Swarm Optimization," *Regist. J. Ilm. Teknol. Sist. Inf.*, vol. 9, no. 1, pp. 18–28, Feb. 2023, doi: 10.26594/register.v9i1.3060.

[35] M. A. Muslim *et al.*, "New model combination meta-learner to improve accuracy prediction P2P lending with stacking ensemble learning," *Intell. Syst. with Appl.*, vol. 18, p. 200204, May 2023, doi: 10.1016/j.iswa.2023.200204.

[36] D. GhoshRoy, P. A. Alvi, and K. Santosh, "Explainable AI to Predict Male Fertility Using Extreme Gradient Boosting Algorithm with SMOTE," *Electronics*, vol. 12, no. 1, p. 15, Dec. 2022, doi: 10.3390/electronics12010015.

[37] L. Li, Z. Liu, J. Shen, F. Wang, W. Qi, and S. Jeon, "A LightGBM-based strategy to predict tunnel rockmass class from TBM construction data for building control," *Adv. Eng. Informatics*, vol. 58, p. 102130, Oct. 2023, doi: 10.1016/j.aei.2023.102130.

[38] M. A. Khan, N. Iqbal, Imran, H. Jamil, and D.-H. Kim, "An optimized ensemble prediction model using AutoML based on soft voting classifier for network intrusion detection," *J. Netw. Comput. Appl.*, vol. 212, p. 103560, Mar. 2023, doi: 10.1016/j.jnca.2022.103560.

[39] A. U. Dullah, F. N. Apsari, and J. Jumanto, "Ensemble learning technique to improve breast cancer classification model," *J. Soft Comput. Explor.*, vol. 4, no. 2, pp. 107–114, Jun. 2023, doi: 10.52465/joscex.v4i2.166.

[40] M. A. Muslim *et al.*, "An Ensemble Stacking Algorithm to Improve Model Accuracy in Bankruptcy Prediction," *J. Data Sci. Intell. Syst.*, vol. 2, no. 2, pp. 79–86, Mar. 2023, doi: 10.47852/bonviewJDSIS3202655.

[41] N. Rai, N. Kaushik, D. Kumar, C. Raj, and A. Ali, "Mortality prediction of COVID-19 patients using soft voting classifier," *Int. J. Cogn. Comput. Eng.*, vol. 3, pp. 172–179, Jun. 2022, doi: 10.1016/j.ijcce.2022.09.001.

[42] Y. Jin, Y. Liu, W. Zhang, S. Zhang, and Y. Lou, "A novel multi-stage ensemble model with multiple K-means-based selective undersampling: An application in credit scoring," *J. Intell. Fuzzy Syst.*, vol. 40, no. 5, pp. 9471–9484, Apr. 2021, doi: 10.3233/JIFS-201954.

[43] R. Verma and S. Chandra, "RepuTE: A soft voting ensemble learning framework for reputation-based attack detection in fog-IoT milieu," *Eng. Appl. Artif. Intell.*, vol. 118, p. 105670, Feb. 2023, doi: 10.1016/j.engappai.2022.105670.

[44] P. Tavana, M. Akraminia, A. Koochari, and A. Bagherifard, "An efficient ensemble method for detecting spinal curvature type using deep transfer learning and soft voting classifier," *Expert Syst. Appl.*, vol. 213, p. 119290, Mar. 2023, doi: 10.1016/j.eswa.2022.119290.

[45] M. T R, V. K. V, D. K. V, O. Geman, M. Margala, and M. Guduri, "The stratified K-folds cross-validation and class-balancing methods with high-performance ensemble classifiers for breast cancer classification," *Healthc. Anal.*, vol. 4, p. 100247, Dec. 2023, doi: 10.1016/j.health.2023.100247.

[46] N. Peker and C. Kubat, "Application of Chi-square discretization algorithms to ensemble classification methods," *Expert Syst. Appl.*, vol. 185, p. 115540, Dec. 2021, doi: 10.1016/j.eswa.2021.115540.

[47] V. Despotovic, T. Skovranek, and C. Schommer, "Speech Based Estimation of Parkinson's Disease Using Gaussian Processes and Automatic Relevance Determination," *Neurocomputing*, vol. 401, pp. 173–181, Aug. 2020, doi: 10.1016/j.neucom.2020.03.058.