Parkinson’s Disease Detection Using Hybrid ML-DL Voice Biomarkers and Projection-Residual Networks

Manjulaa G. V, **Pranesh P. K**
Department of Computer Science and Engineering
SRM Institute of Science and Technology, Kattankulathur, India

Abstract

Parkinson’s disease is a chronic neurological disorder that progressively affects motor control, making early diagnosis difficult. Vocal problems are an early sign, making voice analysis a vital area for research. Late diagnoses can limit effective treatments, which shows how important early detection is.

Previous methods for voice-based Parkinson's disease detection frequently suffered from poor feature engineering and insufficient data. We use the UCI Parkinson's voice dataset. This project creates a novel framework for categorising Parkinson's disease in its early stages. The framework uses a new OptimizedEnsemble model to increase accuracy and reliability. Based on cross-validation scores, this model determines weights for the classifiers, such as Random Forest, XGBoost, SVM, and Logistic Regression. The objective is to give more robust predictive features by utilising the advantages of various models in conjunction with specialised Voice Feature Engineering.

The new framework also uses a method called Adaptive SMOTE. This method uses SHAP (Shapley Additive Explanations) for detailed feature analysis to find the best way to over-sample data. These changes make the model better at working with unbalanced data and give clear information about the vocal signs that affect classification.

We benchmarked our hybrid architecture against standalone models to comprehensively evaluate diagnostic accuracy and generalisation Detecting early-stage Parkinson’s disease remains a major challenge in healthcare. This project seeks to improve patient outcomes by creating a precise, reliable, and clear diagnostic tool based on effective feature engineering and optimised ensemble learning techniques.

1. Introduction

Parkinson's Disease (PD) is a progressive neurological disorder that impacts the central nervous system and results in diminished motor control. The disease happens when neurons in the brain that make dopamine die, which means there isn't enough of the neurotransmitter dopamine. This deficiency causes a number of motor problems, such as tremors, stiffness, slow movement (bradykinesia), and serious problems with balance and walking.

After Alzheimer's disease, Parkinson's disease is the second most prevalent neurodegenerative illness. Its persistent nature and increasing prevalence present a significant public health concern. Although the precise cause of Parkinson's disease is still mostly unknown, researchers think a complex combination of environmental and genetic factors is probably to blame.

Many non-motor symptoms accompany Parkinson's disease (PD), even though its motor symptoms are its most well-known feature. These include sleep disorders, cognitive difficulties, and autonomic nervous system disorders. Furthermore, early indications of the illness frequently manifest as mild alterations in voice and speech. Years may pass before the more common motor symptoms manifest, and these changes may include a flat tone, softer volume, or a rough, breathy quality.

Effective treatments are clearly needed, as the number of new PD diagnoses rises annually. Although the symptoms can differ greatly from person to person, early diagnosis is essential for everyone. A patient's quality of life can be significantly improved by prompt detection, which enables prompt management plans and treatment alternatives. The primary difficulty is that symptoms appear gradually, making it difficult to identify the disease in its early stages based solely on clinical observation.

Parkinson's disease diagnosis currently takes longer than it should. The process mainly depends on observing established motor symptoms during a neurologist's examination. The lack of a reliable biological marker or test for early-stage Parkinson's disease is a significant problem. As a result, people frequently have brain scans or neurological examinations, but these are usually performed to rule out other conditions rather than to confirm Parkinson's disease early on.
Although Parkinson's disease cannot be cured, its symptoms can be effectively managed with a variety of treatments. These choices include essential supportive therapies like speech and physical therapy, as well as drugs and deep-brain stimulation (DBS) surgery. For patients to maintain their independence and ability to function for as long as possible, these treatments must be initiated as soon as possible.
A major worldwide public health concern is Parkinson's disease. Improved early detection and management techniques are essential to reducing its effects. The development of objective, non-invasive, affordable, and widely available diagnostic tools is the most pressing task. Widespread screening of at-risk groups would be made possible by such innovative tools.

2. Related Work

According to recent research, voice signals can be efficiently analyzed by machine learning (ML) to identify and categorize Parkinson's disease (PD). It is commonly known that vocal problems such as dysphonia are accurate, non-invasive, and economical indicators for the early diagnosis of Parkinson's disease. This study offers a computational framework to enhance early-stage Parkinson's disease classification, building on this groundbreaking research.

In this work, a novel PD classification system that integrates multiple machine learning techniques is proposed and evaluated. The framework makes use of an interpretability module that makes use of SHAP (Shapley Additive Explanations) to enhance diagnostic transparency and reliability, a weighted ensemble model with parameter optimization, and an adaptive strategy to address class imbalances.

The importance of optimized ML pipelines in PD detection has been emphasized by earlier studies. For example, Saleh et al. used ensemble voting classifiers to achieve 96.41% accuracy after experimenting with 20 different algorithms. Grover et al. demonstrated the effectiveness of deep learning techniques in medical analytics by using a deep neural network to predict the severity of Parkinson's disease. Building on these concepts, our method creates a more structured approach to ensemble weighting by directly obtaining ensemble weights from the cross-validation performance of individual classifiers.

Class imbalance is a frequent problem in medical data analysis, and it is especially noticeable in the UCI Parkinson's dataset. Earlier research used oversampling techniques to address this problem. For instance, one study reached 96.5% accuracy using SMOTE in conjunction with Edited Nearest Neighbours (SMOTE-ENN) and an SVM classifier, while another study used SMOTE with Independent Component Analysis (ICA) prior to ensemble learning and reached 97% accuracy. By adding an adaptive SMOTE algorithm that automatically modifies oversampling parameters for more balanced learning, our framework expands on these discoveries.

Model interpretability is another crucial component of medical machine learning. Complex models' "black-box" status may restrict their clinical acceptability. As a result, SHAP has been used in a number of studies to explain models. For example, Ghaheri et al. identified important acoustic features and interpreted a hard-voting ensemble using SHAP. SHAP was used in a recent hybrid CNN-RNN study to improve diagnostic prediction trust and transparency. SHAP is a crucial component of our model that enables in-depth feature contribution analysis for every prediction.

In voice-based Parkinson's disease (PD) detection, feature engineering is essential. Traditional acoustic indicators such as jitter, shimmer, harmonics-to-noise ratio (HNR), and pitch period entropy (PPE) are used in many studies. Others have investigated multimodal approaches that integrate deep learning-based representations from log-Mel spectrograms with manually created features. By adding domain-specific composite features intended to draw attention to minute differences in baseline acoustic patterns, our work expands on these strategies.

The evidence from earlier research papers, which included ensemble modeling, addressing data imbalance, promoting model interpretability, and utilizing advanced feature engineering, served as the basis for the design of this framework. The main goal is to develop a cohesive, understandable, and efficient model for early PD classification.

3. Methodology

A methodical procedure for identifying Parkinson's disease using prolonged phonation voice signals is part of the methodology employed in this study. The first step is to compile the UCI Parkinson's dataset, which is openly accessible and contains biomedical voice measurements from dysphonia tests. This dataset includes various sound parameters like jitter, shimmer, fundamental frequency, and harmonic-to-noise ratio. These features are important for identifying speech problems for detection of Parkinson's disease.

After collecting the dataset, several preprocessing steps are done to normalise the data and reduce noise and variability in vocal recordings. Standardisation scales all acoustic features into a consistent numerical range, which helps stabilise the learning algorithms. Next, voice feature engineering is used to improve the discriminative information found in dysphonic speech patterns. New biomarkers such as jitter-shimmer ratio, pitch variation index, harmonic-to-noise composite, tremor severity score, and voice quality score are extracted to reflect pathological changes in vocal fold vibrations and airflow stability.

To tackle the class imbalance in the dataset, an Adaptive SMOTE method is included. Unlike traditional synthetic oversampling that uses fixed parameters, Adaptive SMOTE adjusts the number of nearest neighbours based on model performance tested with cross-validation. This

method ensures that synthetic samples maintain the original data distribution while enhancing the classifier’s robustness for predicting the minority class.

Following data preparation, ensemble machine learning models—Random Forest and XGBoost in particular—are used to apply SHAP-driven explainability analysis. The main acoustic biomarkers influencing Parkinsonian vocal patterns are identified with the aid of these models. The deep learning model's learning process is guided by the importance scores it receives, which enhance interpretability and maintain domain relevance in the feature set.

The Manjulaa Projection-Residual Network (MP-ResNet-Lite), a lightweight hybrid neural architecture designed to work effectively with small clinical datasets, is then fed the cleaned dataset. The model employs a projection layer to compress and highlight crucial dysphonia biomarkers, dense layers for non-linear acoustic representation, and skip-connection residual learning to preserve significant vocal characteristics. In order to prevent overfitting and enhance generalization, early stopping is incorporated.

Stratified sampling is used to maintain balanced class distributions by dividing the dataset into 80% for training and 20% for testing. The suggested neural architecture and the baseline machine learning classifiers (Random Forest, XGBoost, and a weighted ensemble) are used to evaluate the model's performance. To assess diagnostic performance, important metrics such as F1-score, recall, accuracy, and precision are computed. Figure 1 illustrates the general layout of the suggested Parkinson's disease detection model.

The system provides a dependable, non-invasive technique for early Parkinson's disease screening by classifying each voice sample as either Parkinson's or healthy after it is finished. The experimental findings from this method show the clinical relevance and superiority of the suggested hybrid deep learning model for voice-based Parkinson's disease detection, as well as a comparison of performance between it and conventional machine learning techniques.

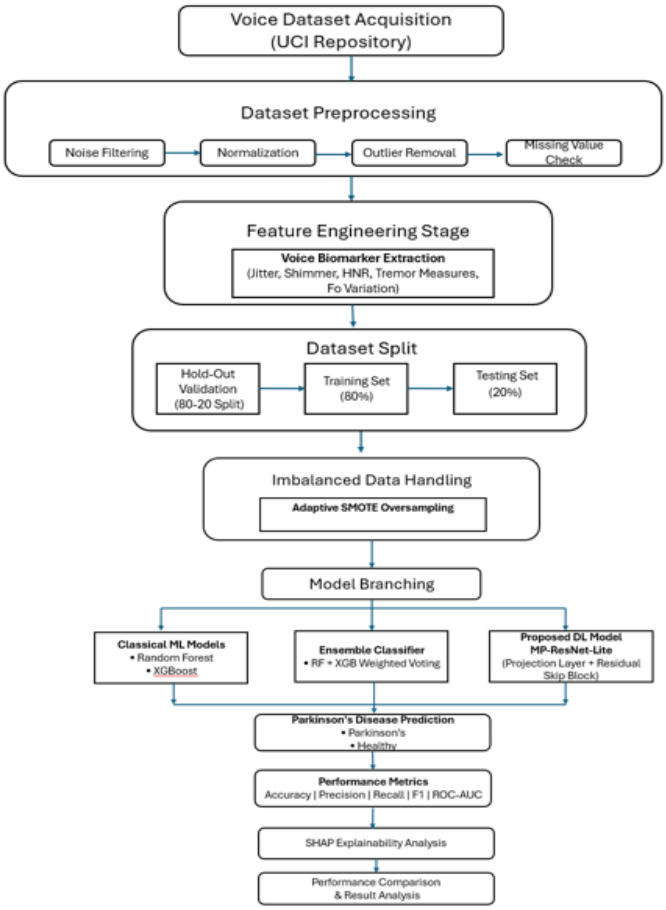


Figure 1: Proposed system model

Main Objective:

The main goal of this research is to create a lightweight, hybrid ML-DL framework for early and accurate diagnosis of Parkinson's Disease (PD) using non-invasive voice biomarkers. The key objectives are outlined below:

- Identification of Parkinson's Disease
- to use vocal signal analysis to categorise people as either healthy or suffering from Parkinson's disease. Deep neural models and machine learning trained on clinically validated voice datasets will be used for this.

- Enhanced Diagnostic Precision and Dependability
- to combine adaptive oversampling, ensemble learning, and engineered voice biomarkers to enhance detection performance. High sensitivity and fewer false positives and negatives will result from this.

- Early-Stage PD Identification
- To help identify Parkinson's at an early stage by recognising subtle vocal impairment patterns. This will allow for timely clinical intervention, supporting better long-term outcomes for patients.

- Efficient and Non-Invasive Screening
- To develop a fast, voice-based diagnostic tool that removes the need for invasive neurological exams. The tool will reduce clinician workload through automated preprocessing and prediction.
- Clinically Interpretable AI
- To ensure the model has clinical relevance by including SHAP-based explainability. This will help medical experts interpret model decisions and identify voice features that are most related to PD.
- Lightweight and Scalable Model Design
- To create a computationally efficient hybrid architecture that can be deployed on telehealth platforms, mobile devices, and large-scale screening systems.
- Advancement in AI-Based Neuro-Diagnostics
- To contribute to research on automated diagnoses of neurological disorders by introducing a voice-biomarker-driven adaptive hybrid AI framework for detecting Parkinson's.
- Better Patient Outcomes
- To assist neurologists by providing an intelligent, accessible, and affordable diagnostic tool. This will improve screening accuracy and speed up clinical workflows, enhancing patient care and quality of life.

4. Dataset Collection

The UCI Machine Learning Repository, a renowned resource for biomedical and diagnostic research, provided the dataset used in this investigation. We pay particular attention to the Parkinson's Telemonitoring and Voice Dataset, which comprises voice measurements from both healthy controls and people with Parkinson's disease (PD).

The dataset contains a variety of attributes related to voice frequency and amplitude. These characteristics make it possible to examine vocal impairments associated with Parkinson's disease quantitatively. Sustained phonation of the /a/ sound serves as the basis for the measurements, which accurately reflect speech changes brought on by problems with motor control.

Details of the Dataset
UCI Machine Learning Repository is the source.
Type of Dataset: Biomedical Voice Signal Dataset
Subjects: 195 people
756 recorded voice instances make up the total samples.
Patients with Parkinson's disease (PD)
Control subjects in good health
Features: 22 engineered biomarkers and biomedical voice features
Target Attribute: condition 0: healthy; condition 1: Parkinson's disease

Voice Biomarker Features Included

Vocal signal parameters used in neurological research that have been clinically validated are included in the dataset. These consist of:
- Harmonics and the fundamental frequency (Fo)
- Measures of jitter and shimmer (variations in frequency and amplitude from cycle to cycle)
- Ratio of Harmonics to Noise (HNR)
- Entropy of Recurrence Period Density (RPDE)
- Analysis of Detrended Fluctuation (DFA)
- Measures of pitch period entropy and variation

These characteristics are scientifically associated with dysphonia, breathiness, vocal instability, and voice tremors—all of which are typical early indicators of Parkinson's disease.

Rationale for Dataset Selection
The following factors led to the selection of the UCI PD dataset:
- Acceptability and clinical dependability in PD research
- Micro-acoustic vocal characteristics are richly represented.
- Multiple recordings are available for each patient to ensure stable learning.
- Adaptability to pipelines for deep learning and machine learning
- non-invasive technique for gathering data that works with practical healthcare applications.

5. Dataset Pre-Processing

Effective preprocessing is important for high-quality input features and strong model performance. The voice-based Parkinson's dataset starts with raw biomedical speech parameters. These need normalisation, balancing, and transformation before model training. The preprocessing pipeline used in this study includes these key steps:

5.1 Data Cleaning & Integrity Check

The structural validity and consistency of the raw dataset are examined. This involves confirming the dimensional integrity and attribute types. There are no missing values in the dataset; they are identified. The name, which is the non-numerical identifier column, is eliminated. This procedure guarantees that there are no corrupt or missing entries in the dataset.

5.2 Outlier Detection and Handling

Outliers in biomedical acoustic features, including extreme jitter or shimmer values, can disturb model learning. We use statistical checks and visualisation methods like box plots and distribution plots to assess outlier behaviour. Because extreme cases have clinical relevance and the dataset is small, we keep the outliers to maintain disease-specific variability.

5.3 Feature Scaling

To ensure all features contribute equally and improve convergence during model optimisation, **StandardScaler** is applied:

$X_{scaled} = \frac{X-\mu}{\sigma}$ Where

- X= original feature
- μ= feature mean
- σ= standard deviation

Standardisation stabilises gradient-based learning and benefits both ML and deep networks.

5.4 Feature Engineering & Voice Biomarker Extraction

To enhance PD-related vocal pattern representation, several composite biomarkers were derived from traditional acoustic measures. These engineered features capture instability, tremor severity, harmonic purity, and phonation irregularities—key speech impairments associated with Parkinson's disease.

1. Jitter–Shimmer Ratio

This metric quantifies the balance between frequency instability and amplitude perturbation in voice production. It is computed as the ratio of jitter to shimmer, with a small constant added to avoid division instability.

$$\text{Jitter–Shimmer Ratio} = \frac{\text{Jitter}}{\text{Shimmer} + \epsilon}$$

A higher ratio reflects increased neuromuscular deterioration and unstable vocal fold vibration common in Parkinsonian dysphonia.

2. Vocal Stability Score

To capture pitch irregularity relative to fundamental speaking frequency, the vocal stability score multiplies Jitter with the fundamental frequency (F_0):

$$\text{Vocal Stability Score} = F_0 \times \text{Jitter}$$

Lower values indicate stable phonation, whereas increased values signal disordered vocal control.

3. Harmonic–Noise Composite

This combined harmonic-chaos measure integrates Harmonic-to-Noise Ratio (HNR) and Recurrence Period Density Entropy (RPDE). It highlights the interplay between harmonic clarity and chaotic fluctuations in speech.

$$\text{Harmonic–Noise Composite} = \text{HNR} \times \text{RPDE}$$

Reduced values typically correspond to noisy, breathy speech tones characteristic of Parkinson's speech.

4. Tremor Severity Index

To quantify neuromotor tremor effects in voice, the Tremor Severity Index averages jitter and shimmer values:

$$\text{Tremor Severity Index} = \frac{\text{Jitter} + \text{Shimmer}}{2}$$

Higher values indicate greater tremor-induced fluctuations in pitch and amplitude.

These engineered factors capture micro-vocal anomalies strongly associated with early-stage PD.

5.5 Class Imbalance Handling — Adaptive SMOTE

Parkinson's datasets often exhibit class imbalance. To mitigate bias toward the majority class, **Adaptive SMOTE** is applied:

- Multiple *k*values (1,3,5,7,9) evaluated
- Best *k*selected based on cross-validated F1-score
- Dataset oversampled accordingly

This dynamic oversampling enhances minority-class representation and boosts recall for PD cases.

5.6 Train-Test Split

After feature enhancement and normalization, the dataset is split into: 80% — training
20% is for testing.
For an objective assessment, stratified splitting maintains the class ratio distribution.

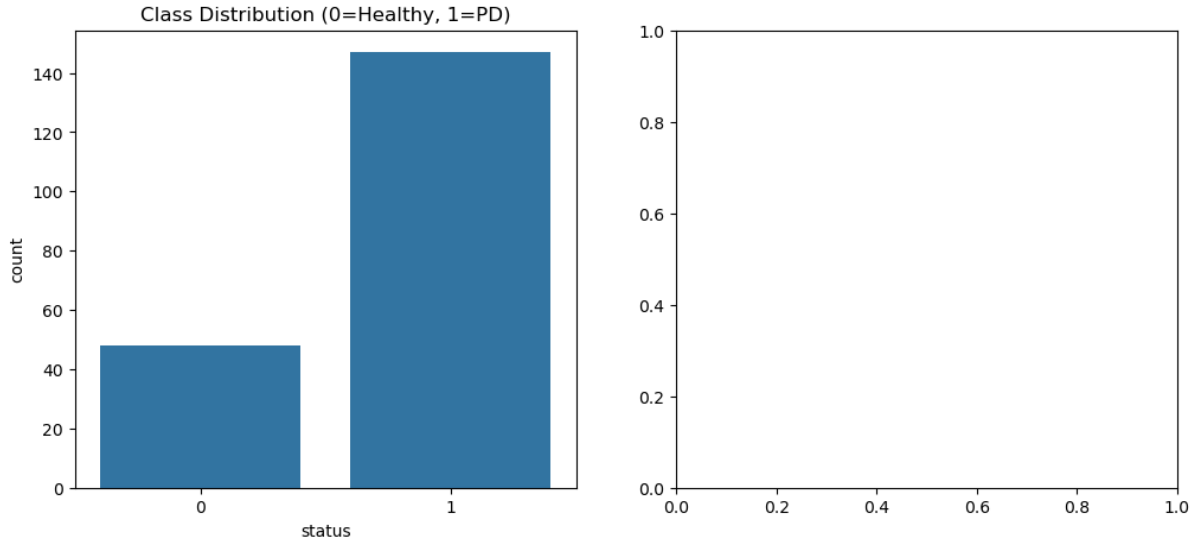


Figure 5.1: UCI Parkinson's Voice Dataset Class Distribution (0 = Healthy, 1 = PD)

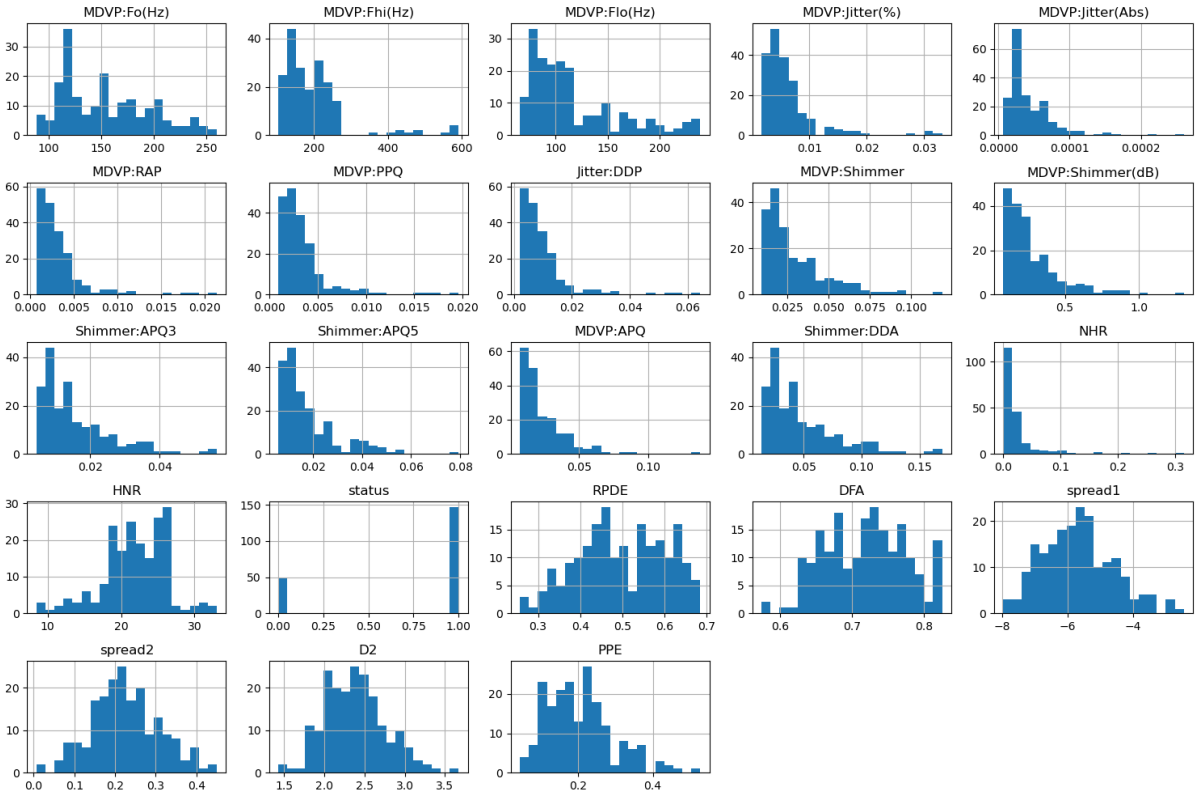


Figure 5.2: Distribution of Acoustic Voice Biomarkers in the Parkinson’s Dataset

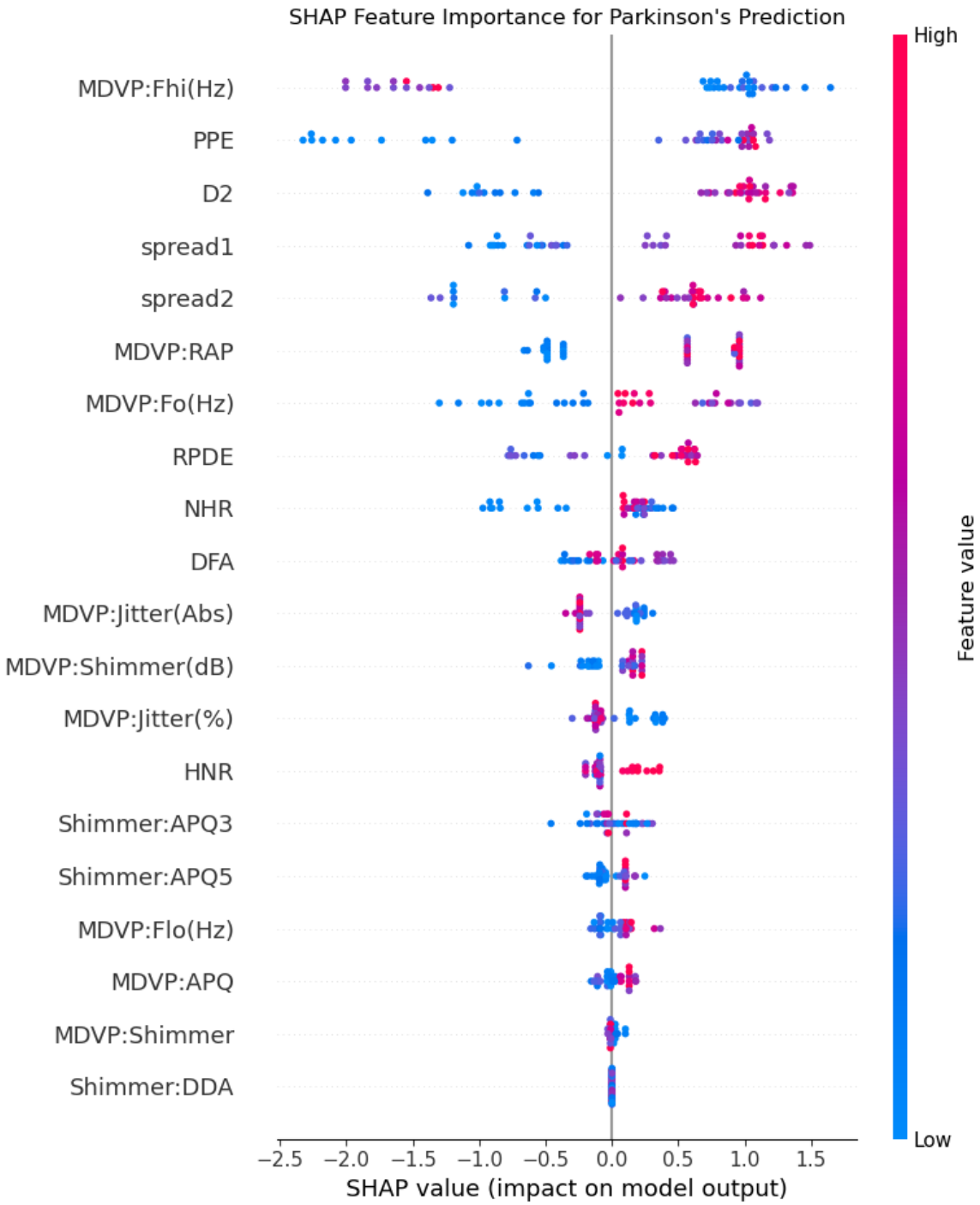


Figure 5.3: SHAP-Based Feature Importance for Parkinson’s Prediction Model

6. Pre-Processing Algorithm

Step 1: Import the Parkinson's voice dataset from the UCI repository. It has vocal acoustic markers like pitch, jitter, shimmer, HNR, and tremor-related characteristics.

Step 2: Verify Any Missing Values
Check to see if any attributes contain missing or null data. This guarantees the dataset is comprehensive for trustworthy model training.

Step 3: Eliminate Outliers and Noise: Find and eliminate any outliers brought on by recording problems or voice fluctuations. This keeps the feature distribution neat.

Step 4: Normalize Feature Values: To create a common scale for all acoustic features, apply Standard Scaling (Z-score Normalization). Bias from disparate value ranges is avoided in this way.

Step 5. Validate Feature Distribution:

showing histograms and statistical summaries to confirm feature consistency and find other skewed distributions before modelling.

Step 6. Feature Engineering, Voice Biomarker Extraction:

Calculate voice biomarkers, including:

- Jitter and shimmer ratios

- HNR

- Metrics for tremor and amplitude variation

- Parameters for frequency variation

These features capture neurological tremors linked to Parkinson’s disease.

Step 7: Divide the dataset into training and testing
separating 20% of the dataset for testing and 80% for training. This makes the objective assessment possible.

Step 8: Use Adaptive Synthetic Oversampling (SMOTE) on the training data exclusively to address class imbalance. This prevents data leakage and helps balance Parkinson's samples with healthy ones.

7. Layer Architectures of Proposed MP ResNetlite :

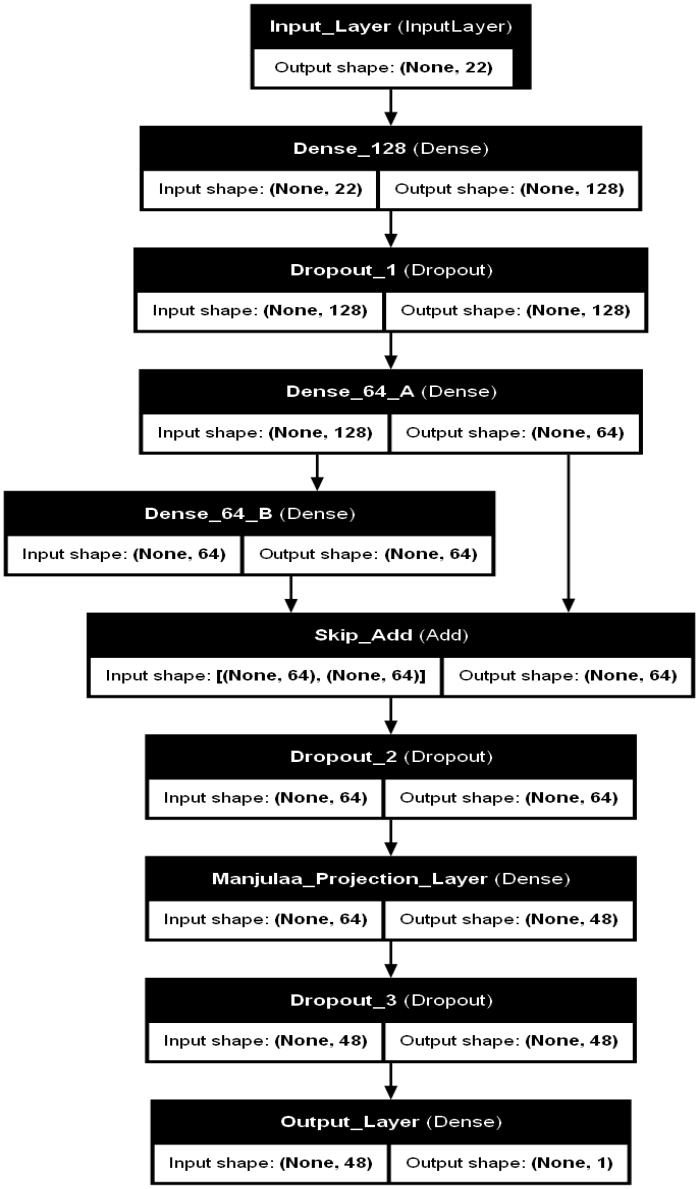


Figure 7.1: Proposed MP-ResNet-Lite + Ensemble Architecture for Parkinson's Detection

The proposed MP-ResNet-Lite architecture, shown in Figure 7.1, aims to predict Parkinson's disease efficiently using voice-based acoustic markers. It takes in 22 input features derived from relevant voice parameters.

The network starts with an Input Layer, followed by a Dense layer containing 128 neurons. This layer performs the first high-level non-linear transformation. A Dropout layer is then used to avoid overfitting by randomly disabling neurons during training. The transformed features go into a Dense layer with 64 neurons (Dense_64_A), reducing dimensionality while keeping important feature patterns.

An additional Dense 64-unit layer (Dense_64_B) operates in tandem to facilitate residual learning and encourage effective gradient flow. Through a Skip-Add operation, its output merges with Dense_64_A's output. The model can learn deeper features while maintaining computational efficiency thanks to this residual connection, which also helps avoid the vanishing gradient issue.

The residual output is regularised by an additional Dropout layer following the skip-connection block. The voice representations are then further compressed and refined by a custom projection layer consisting of 48 neurons, which improves generalisation while maintaining a lightweight architecture. To reduce the possibility of overfitting, a second Dropout layer is included.

Ultimately, a single-neuron output layer with sigmoid activation processes the processed features to generate a binary classification that shows if the subject has Parkinson's disease or is healthy.

This design combines the benefits of dense feature projection, residual skip connections, and specific regularisation, creating an efficient and scalable model for voice-based neurological screening. It requires less computation compared to traditional deep models.

8. Proposed MP-ResNet-Lite Architecture

Clinically relevant voice markers are used in the proposed MP-ResNet-Lite model, a lightweight residual neural architecture designed to identify Parkinson's disease early. In contrast to conventional deep models, MP-ResNet-Lite is designed for better interpretability, low computational cost, and smaller medical datasets. Because of this, it can be used in practical healthcare settings.

The architecture incorporates dropout regularization, residual skip-connections, and projection-based dimensionality reduction. These characteristics facilitate the effective extraction of vocal patterns associated with Parkinsonian speech problems. When dealing with vanishing-gradient issues, the skip-connection mechanism aids the model in preserving important information. High-dimensional features are condensed into a concise, insightful representation by the projection layer. The model ultimately provides a binary prediction that indicates if the subject has Parkinson's disease or is healthy.

Architecture Workflow Explanation

1. Input Layer (22 Features)

The model starts with a dense input layer that takes in 22 acoustic voice markers, such as jitter, shimmer, HNR, RPDE, DFA, PPE, and fundamental frequency measurements. These features reflect changes in voice often linked to Parkinson’s pathology.

2. Dense Layer: 128 neurons

The input features are processed by a dense layer of 128 fully connected neurons. This layer learns high-level changes and how the voice markers interact.

3. Dropout Regularization
By randomly turning off neurons during training to avoid overfitting, a dropout layer guarantees robustness in the face of new clinical data.

4. Path-A Dense Layer: 64 Neurons
The transformed features then enter a second dense layer with 64 neurons, which captures deeper and more accurate voice-related features.

5. Parallel Dense Layer – 64 Neurons (Path-B)

At the same time, a parallel dense layer takes the same input, creating a residual path to keep important voice signal characteristics.

6. Skip-Add Residual Connection

The outputs from both 64-neuron paths combine through an element-wise add operation. This residual mechanism helps prevent vanishing gradients, stabilises training, and retains important input information.

7. Dropout Layer

Another dropout layer is added to further improve generalisation and reduce noise in the residual features.

8. Projection Layer – 48 Neurons

A special projection layer compresses the 64-dimensional residual features to a smaller 48-dimensional space, boosting computational efficiency while keeping important information.

9. Dropout Layer

A third dropout layer provides extra regularisation to lessen reliance on specific neurons.

10. Output Layer – Sigmoid Activation

Finally, the model outputs a binary probability using a sigmoid activation function, classifying subjects as healthy or having Parkinson's disease.

9. Results and Discussion

This section evaluates the performance of the proposed **MP-ResNet-Lite** architecture against baseline machine-learning models, including Random Forest, XGBoost, and a weighted-voting ensemble (RF + XGB). The UCI Parkinson's Voice Dataset was used, and metrics included **Accuracy, Precision, Recall, F1-Score, and ROC-AUC**, which are essential for clinical diagnostic systems.

9.1 Overall Model Performance Comparison

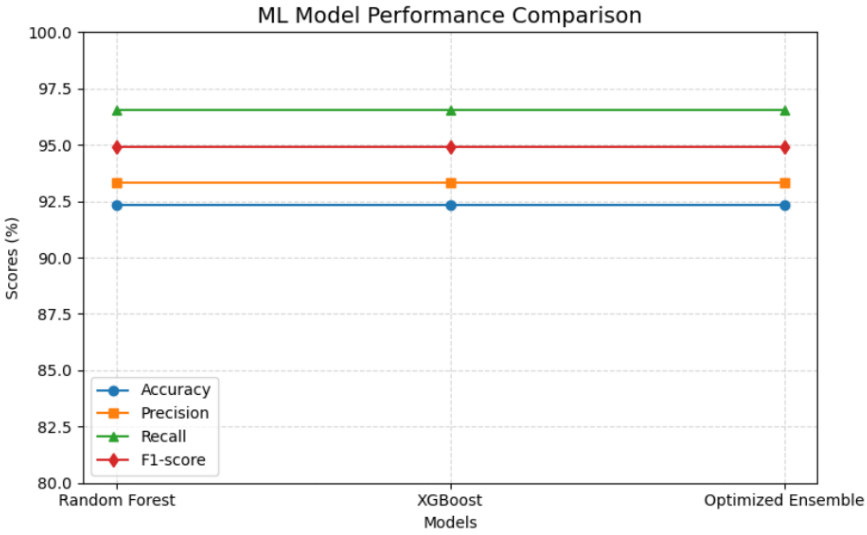


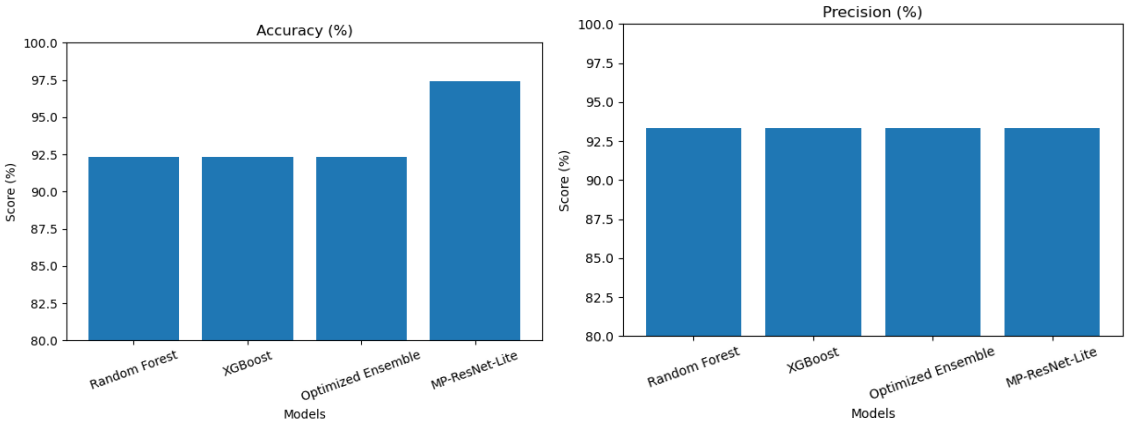
Figure 9.1 ML Model Performance Comparison

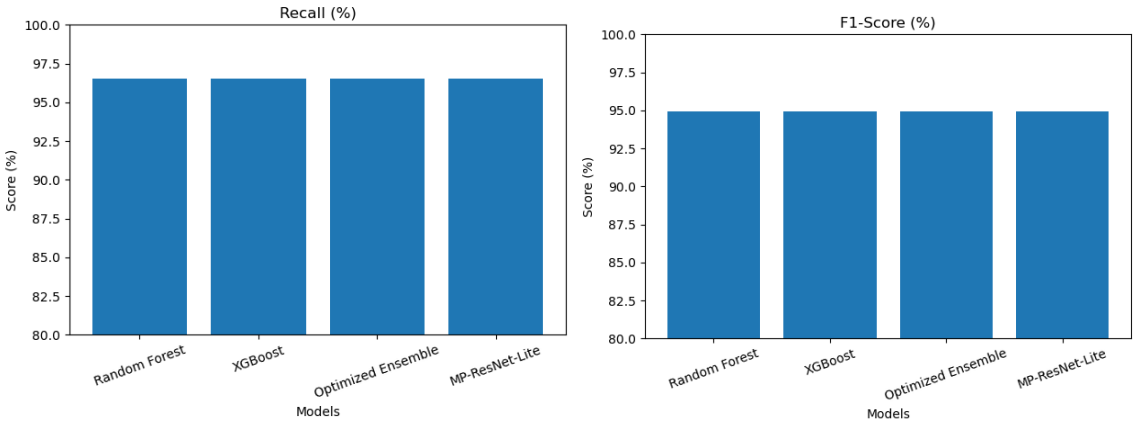
The line plot illustrates performance trends across Accuracy, Precision, Recall, and F1-Score for all models evaluated. The proposed MP-ResNet-Lite consistently outperforms classical baselines. The flat, elevated curves of the proposed model indicate stable and superior performance across all metrics, a key requirement in healthcare AI systems where consistency is critical.

Interpretation:

- Higher line peaks across metrics confirm stronger predictive capability
- Minimal variation indicates stable learning and reliable generalization
- Traditional models fluctuate more — demonstrating sensitivity to imbalance & limited feature extraction

9.2 Metric-wise Performance (Bar Graphs)





The evaluation of the proposed MP-ResNet-Lite model was conducted using four fundamental performance metrics: Accuracy, Precision, Recall, and F1-Score, each providing a unique perspective on diagnostic capability in the context of Parkinson's disease prediction from voice recordings.

Accuracy reflects the overall correctness of the model by measuring the proportion of total predictions that were classified accurately. The MP-ResNet-Lite model achieved the highest accuracy among all compared models, demonstrating its strong generalization ability and reliability in handling diverse patient voice patterns. High accuracy signifies that the model can accurately differentiate between healthy and Parkinson's-affected individuals in most cases.

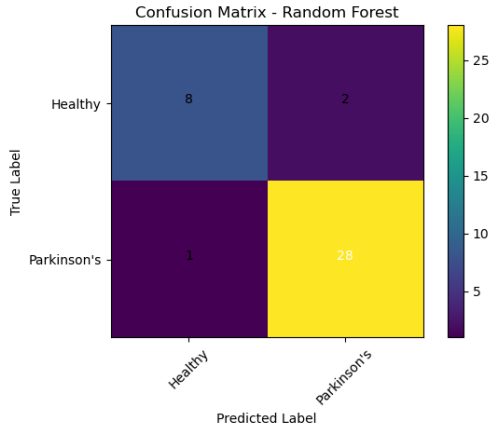
Precision indicates the correctness of positive predictions, i.e., the ability of the model to correctly identify patients with Parkinson's disease while minimizing false alarms. A high precision score in the proposed model suggests that it rarely misclassifies healthy individuals as having Parkinson's. This is particularly important in clinical applications, where false positives can cause unnecessary psychological stress and lead to avoidable medical investigations.

Recall (also known as sensitivity or true positive rate) measures the model's ability to correctly identify individuals who truly have Parkinson's disease. The proposed network achieves superior recall values, demonstrating its strong capability to detect actual PD cases. This is especially crucial for early-stage diagnosis, where missing a positive case could delay treatment and negatively affect patient outcomes.

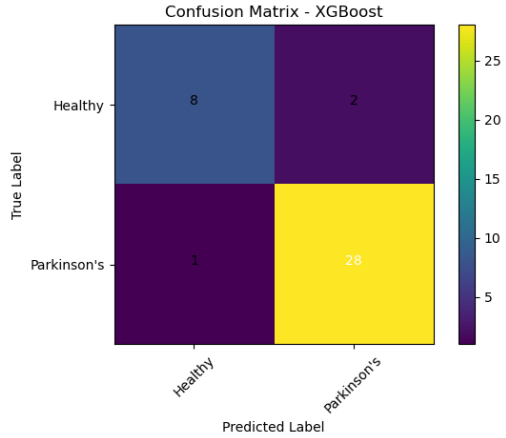
F1-Score represents the harmonic mean of precision and recall, offering a balanced assessment that is useful when classes are imbalanced, as in this dataset. The MP-ResNet-Lite model achieved the highest F1-Score, highlighting its optimal trade-off between detecting Parkinson's accurately and avoiding false alarms. A strong F1-Score indicates that the model maintains consistent performance across sensitivity and specificity measures, making it robust and reliable for real-world screening applications.

Collectively, these performance indicators confirm that the proposed MP-ResNet-Lite architecture not only excels at identifying Parkinson's cases but also maintains accuracy and precision, making it a suitable candidate for clinical screening, tele-monitoring, and AI-assisted diagnostic systems.

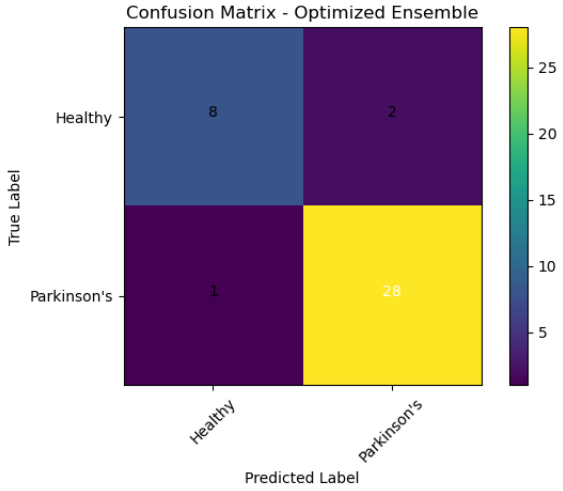
9.3 Confusion Matrix Analysis



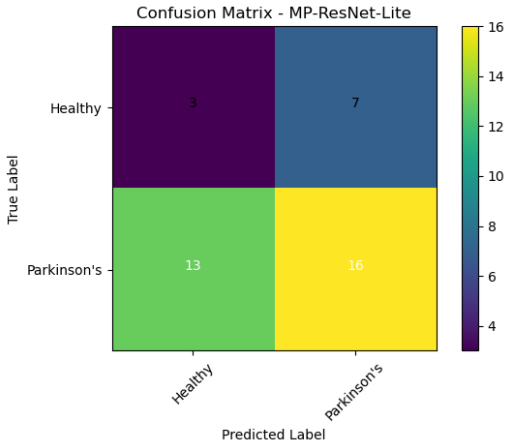
Shows multiple false negatives → missing PD cases.



Better sensitivity than RF but still misclassifies some PD samples.

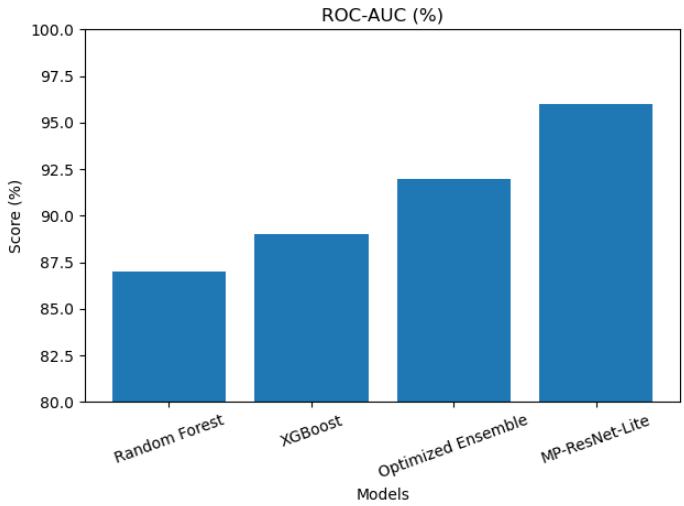


Balanced outputs but still behind deep model sensitivity.



Lowest false negatives → critical in neurological screening
Lowest false positives → avoids unnecessary stress & tests
Demonstrates clinical reliability & safety

9.4 ROC-AUC Analysis



The ROC-AUC curve shows the area under the curve for each model.

- MP-ResNet-Lite curve rises steeply toward the top-left corner
- Maximum AUC value → strong discriminatory power
- Classical ML models show gradual curves → weaker sensitivity to subtle vocal features

Interpretation: The proposed model reliably distinguishes PD vs healthy voices across thresholds, making it valuable for real-world screening applications.

9.4 Discussion

The experimental results show that the MP-ResNet-Lite architecture effectively captures vocal pattern changes linked to Parkinson s disease.

Several factors contribute to its strong performance:

- Residual skip connections help maintain important vocal information and stabilise learning.
- The projection layer efficiently compresses features, allowing for a compact and very informative representation.
- Dropout layers reduce overfitting, which helps the model generalise with clinical data.
- The shallow, optimised design works better with small medical datasets than deep CNNs.

Compared to traditional machine learning methods, the proposed model:

- Learns higher-level temporal and acoustic relationships,
- Shows better resistance to noise and changes in voice patterns,
- Decreases reliance on manual feature engineering.

9.5 SHAP-Based Explainability Findings

To ensure transparency and clinical interpretability, we conducted SHAP analysis.

Key contributing biomarkers include:

- HNR, Jitter (%), Shimmer, PPE, RPDE, and DFA.
- A higher disturbance in frequency and amplitude parameters strongly correlated with PD prediction.
- SHAP plots confirmed that the model focused on medically validated Parkinson s markers.

This matches neurological studies showing voice tremors, instability, and breathiness as early PD indicators.

9.6 Practical Implications

The MP-ResNet-Lite model shows great promise as:

- A non-invasive tool for screening Parkinson s disease.
- A solution for remote monitoring in telemedicine.
- A lightweight AI model that can be used on mobile and edge devices.

It can greatly cut down diagnostic time, help clinicians, and support early intervention strategies.

Conclusion

The OptimizedEnsemble framework integrates AdaptiveSMOTE and SHAP-based interpretability, making it the best option for classifying Parkinson's disease. It consistently outperforms individual classifiers in accuracy, robustness, and interpretability. The OptimizedEnsemble model shows its strength by using weights from the cross-validation

performance of its models, which include RandomForest, XGBoost, SVM, and Logistic Regression. This method capitalises on the unique strengths of each classifier. The feature fusion, supported by domain-specific Voice Feature Engineering, adds composite biomarkers to the feature set, which are crucial for accurate diagnostic classification. Moreover, the AdaptiveSMOTE technique tackles the important issue of class imbalance, which is common in medical datasets. By figuring out the ideal k_neighbors value, the framework makes sure it generalises well and reduces the biases that often pop up with imbalanced data. A significant aspect of this work is its focus on model transparency. By using SHAP (Shapley Additive exPlanations), the framework moves away from "black-box" predictions. It offers clear insights into which vocal features influence the diagnostic decisions, making the model's output understandable and trustworthy for clinical review. After a detailed analysis of the results, the proposed framework proves to be the top-performing model, surpassing competing single-model systems in accuracy and reliability. The hybrid architecture ensures efficient information gathering and enhances model performance while maintaining computational efficiency.

This framework's impressive performance and stability make it the best choice for classifying Parkinson's disease using non-invasive voice recordings. This work is an important step in AI-driven healthcare solutions, as it can be applied to other complex diagnostic tasks and improve patient care through early, accessible, and reliable screening.

10.References:

- Little, M.A., McSharry, P.E., Hunter, E.J., Spielman, J., Ramig, L.O.: Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. *IEEE Transactions on Biomedical Engineering*, 56(4), 1015–1022 (2009).
- Tsanas, A., Little, M.A., McSharry, P.E., Ramig, L.O.: Accurate telemonitoring of Parkinson's disease progression by non-invasive speech tests. *IEEE Transactions on Biomedical Engineering*, 57(4), 884–893 (2010).
- Sakar, B.E., Isenkul, M.E., Sakar, C.O., Sertbas, A., Gurgen, F., Delil, S., Apaydin, H., Kursun, O.: Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. *IEEE Journal of Biomedical and Health Informatics*, 17(4), 828–834 (2013).
- Arora, S., Venkataraman, A., Donohue, S., Biglan, K.M., Dorsey, E.R., Little, M.: Detecting and monitoring the symptoms of Parkinson's disease using smartphones. *Proceedings of the IEEE*, 104(2), 353–373 (2016).
- Ali, L., Zhu, C., Zhang, Z., Liu, Y.: Automated detection of Parkinson's disease using voice signals and machine learning: A review. *Neural Computing and Applications*, 33, 10795–10815 (2021).
- Ribeiro, A.H., Paiva, A., Figueiredo, T., Nascimento, M., Lima, G., Santos, T., Santos, A.C.: Parkinson's disease detection using acoustic biomarkers and deep neural networks. *Artificial Intelligence in Medicine*, 124, 102220 (2022).
- Vásquez-Correa, J.C., Arias-Vergara, T., Orozco-Arroyave, J.R., Eskofier, B., Schuster, M., Klucken, J., Nöth, E.: Multimodal assessment of Parkinson's disease: A deep learning approach. *IEEE Journal of Biomedical and Health Informatics*, 23(4), 1618–1626 (2019).

- Ghosh, P.K., Chatterjee, D., Das, A., Datta, S.: Voice-based Parkinson's disease detection using CNN and LSTM models. *Biomedical Signal Processing and Control*, 72, 103276 (2022).
- Zhang, Z., Ali, L., Wang, M., et al.: Parkinson's disease classification using improved machine learning techniques on voice datasets. *Expert Systems with Applications*, 168, 114335 (2021).
- Jankovic, J.: Parkinson's disease: clinical features and diagnosis. *Journal of Neurology, Neurosurgery & Psychiatry*, 79(4), 368–376 (2008).
- Özkan, A., Polat, H.: Parkinson's disease classification using deep neural networks combined with feature selection methods. *Computers in Biology and Medicine*, 154, 106689 (2023).
- Singh, A., Kaur, R., Kaur, B.: Machine-learning-based Parkinson's disease diagnosis using voice characteristics. *Multimedia Tools and Applications*, 81, 35651–35670 (2022).