

Early Detection of Parkinson's Disease Using Deep Learning-Based Voice Analysis

Nithya.M¹, Sridharshan.J.²

Assistant Professor, Department of Computer Engineering, K.L.N. College of Engineering, Pottapalayam,
Tamil Nadu, India¹

U.G. Student, Department of Computer Engineering, K.L.N. College of Engineering, Pottapalayam,
Tamil Nadu, India²

ABSTRACT: Parkinson's disease (PD) is a progressive neurodegenerative disorder that primarily affects movement and speech, making early detection essential for effective management and improved patient outcomes. This project proposes an AI-based hybrid system for the early detection of Parkinson's disease by integrating both machine learning and deep learning approaches. The system operates in two stages: a mini project that analyzes structured numerical voice features and a final project that processes raw audio signals. In the machine learning stage, the UCI Parkinson's dataset is utilized, where features are preprocessed through normalization and optimized using Recursive Feature Elimination (RFE) to select the most relevant attributes. Multiple classifiers, including Support Vector Machine (SVM), Random Forest, and Gradient Boosting, are trained, and a stacking ensemble model is implemented to improve prediction performance. In the deep learning stage, speech recordings are preprocessed, augmented, and converted into Mel Spectrograms, which are then used to train a Convolutional Neural Network (CNN) for accurate classification. The entire system is integrated into a web-based application using Flask, allowing users to upload CSV files, individual audio files, multiple files, or compressed ZIP folders for real-time prediction. The output is generated in a structured CSV format, indicating whether each input corresponds to a healthy individual or a Parkinson's patient. The proposed system demonstrates the effectiveness of combining machine learning and deep learning techniques in biomedical voice analysis and provides a scalable, user-friendly, and non-invasive solution for early Parkinson's disease detection.

KEYWORDS: Parkinson's Disease, Machine Learning, Deep Learning, Voice Analysis, Mel Spectrogram, CNN, Ensemble Learning, RFE, Flask, Early Detection

I. INTRODUCTION

Parkinson's disease (PD) is a progressive neurodegenerative disorder that primarily affects movement, coordination, and speech due to the gradual loss of dopamine-producing neurons in the brain. Early symptoms such as tremors, stiffness, slowed movement, and subtle changes in speech are often difficult to detect in the initial stages. Traditional diagnostic methods rely on clinical observation and neurological expertise, which can be subjective, time-consuming, and dependent on the experience of medical professionals. Therefore, there is a strong need for automated, objective, and efficient systems that can assist in the early detection of Parkinson's disease.

With the rapid advancement of artificial intelligence (AI), machine learning (ML), and deep learning (DL), it has become possible to analyze complex biomedical data and identify hidden patterns associated with diseases. Among various diagnostic approaches, voice analysis has emerged as a highly effective and non-invasive technique. Individuals affected by Parkinson's disease often exhibit distinct vocal impairments such as reduced loudness, monotonic speech, irregular pitch, and vocal tremors. These changes can be captured either as structured numerical features or directly from raw audio signals, making them suitable for computational analysis.

This project proposes a hybrid AI-based Parkinson's Disease Detection System that integrates both machine learning and deep learning techniques. The system is divided into two major components: a mini project that processes structured numerical voice features using machine learning, and a final project that analyzes speech recordings using deep learning. In the machine learning module, the UCI Parkinson's dataset is utilized, where features are preprocessed

through normalization and optimized using Recursive Feature Elimination (RFE). Multiple classifiers such as Support Vector Machine (SVM), Random Forest, and Gradient Boosting are combined using a stacking ensemble model to improve prediction performance.

In the deep learning module, audio recordings are preprocessed by removing noise, applying augmentation techniques, and converting them into Mel Spectrograms. These spectrograms are then used to train a Convolutional Neural Network (CNN), which is capable of capturing complex patterns in speech signals. The entire system is integrated into a web-based application using Flask, allowing users to upload CSV files, individual audio files, multiple files, or ZIP folders for prediction. The system generates results in a structured CSV format, indicating whether each input corresponds to a healthy individual or a Parkinson's patient.

This integrated approach enhances the flexibility, accuracy, and usability of the system by supporting multiple input formats and leveraging the strengths of both machine learning and deep learning. The proposed system aims to provide a reliable, non-invasive, and scalable solution for early detection of Parkinson's disease, contributing to improved diagnosis and better patient care.

II. LITERATURE REVIEW

[1] W. S. Lim, S. I. Chiu, P. L. Peng, et al., "A Cross-Language Speech Model for Detection of Parkinson's Disease," *Journal of Neural Transmission*, Springer, vol. 132, pp. 579–590, 2025.

This research presents a speech-based model capable of detecting Parkinson's disease across multiple languages. It focuses on extracting speech patterns that remain consistent regardless of language differences. The results show that speech-based detection models can generalize well across diverse datasets, improving their applicability in real-world scenarios.

[2] S. Srinivasan, P. Ramadass, S. K. Mathivanan, K. P. Selvam, and M. A. Shah, "Detection of Parkinson Disease Using Multiclass Machine Learning Approach," *Scientific Reports*, vol. 14, Article 13813, 2024.

This study proposes a multiclass machine learning framework for Parkinson's disease detection using biomedical data. It evaluates multiple algorithms and compares their performance in classification tasks. The findings indicate that machine learning models can effectively differentiate between disease stages, enhancing early detection and diagnosis.

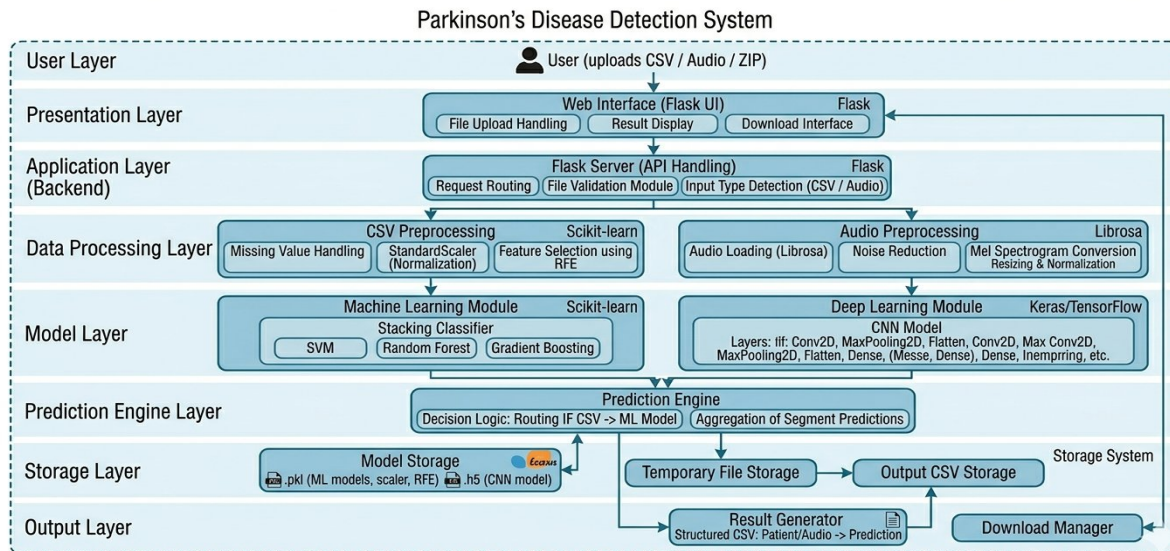
[3] I. Tougui, M. Zakroum, O. Karrakchou, and M. Ghogho, "Transformer-Based Transfer Learning on Self-Reported Voice Recordings for Parkinson's Disease Diagnosis," *Scientific Reports*, vol. 14, 2024.

This paper introduces a transformer-based deep learning approach for analyzing voice recordings in Parkinson's detection. By applying transfer learning, the model improves performance even with limited data. The study demonstrates that advanced deep learning techniques can capture complex speech patterns and significantly enhance diagnostic accuracy.

[4] H. Xu, W. Xie, M. Pang, Y. Li, L. Jin, F. Huang, and X. Shao, "Non-Invasive Detection of Parkinson's Disease Based on Speech Analysis and Interpretable Machine Learning," *Frontiers in Aging Neuroscience*, vol. 17, 2025.

This research focuses on interpretable machine learning models for Parkinson's detection using speech analysis. It emphasizes transparency in predictions, allowing better understanding of how features influence results. The study confirms that non-invasive speech-based systems are effective and can be used as reliable tools in clinical decision support systems.

III. PROPOSED SYSTEM



The proposed system is an AI-based Parkinson’s Disease Detection System designed to provide an accurate, non-invasive, and user-friendly solution for early diagnosis. The system integrates both machine learning and deep learning techniques to analyze two different types of input data—structured numerical features and unstructured audio signals. This hybrid approach enhances the overall capability of the system by leveraging the strengths of both methodologies. The system is divided into two main modules: the mini project module and the final project module. The mini project focuses on analyzing structured numerical data obtained from voice measurements. In this module, the dataset is first preprocessed by removing irrelevant attributes and normalizing the feature values using standard scaling techniques. Feature selection is then performed using Recursive Feature Elimination (RFE), which identifies the most significant features that contribute to accurate classification. By reducing the number of features, the system minimizes noise and improves computational efficiency. The selected features are then fed into multiple machine learning models, including Support Vector Machine (SVM), Random Forest, and Gradient Boosting. These models are combined using a stacking ensemble technique, where the predictions of individual models are used as input to a final meta-classifier. This ensemble approach improves prediction accuracy and reduces the risk of overfitting.

The final project module focuses on analyzing speech recordings using deep learning techniques. Audio data is first preprocessed using a series of steps to ensure quality and consistency. Silence removal is performed to eliminate unnecessary parts of the signal, followed by data augmentation techniques such as noise addition and pitch shifting to increase the diversity of the dataset. The processed audio is then converted into Mel Spectrograms, which represent the frequency distribution of the signal over time. These spectrograms serve as input to a Convolutional Neural Network (CNN), which is designed to automatically extract spatial features and identify patterns associated with Parkinson’s disease. The CNN model consists of multiple convolutional and pooling layers, followed by fully connected layers that perform the final classification. Both modules are integrated into a unified web-based application developed using Flask. The user interface is designed to be simple and intuitive, allowing users to upload CSV files containing numerical data or audio files for analysis. The system supports multiple input formats, including individual files, multiple files, and compressed ZIP folders. Once the input is provided, the backend processes the data using the appropriate model and generates predictions. The results are then compiled into a structured CSV file, where each entry is labeled as either “Healthy Individual” or “Parkinson Individual.”

The proposed system offers several advantages, including flexibility, scalability, and ease of use. It eliminates the need for invasive medical procedures and provides quick results, making it suitable for preliminary screening and remote healthcare applications. Additionally, the combination of machine learning and deep learning ensures high reliability

and robustness. Overall, the system serves as an efficient diagnostic support tool that can assist healthcare professionals in the early detection and monitoring of Parkinson's disease.

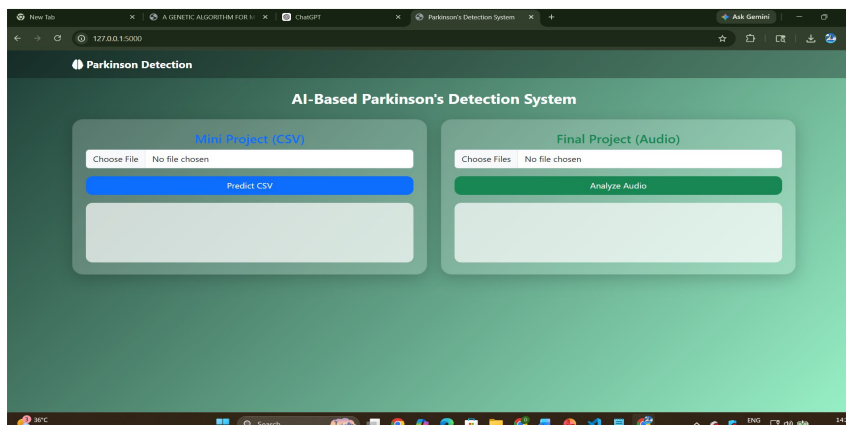
IV. METHODOLOGY AND WORKFLOW

The methodology of the proposed Parkinson's Disease Detection System is designed as a structured pipeline that processes input data, extracts meaningful features, applies intelligent models, and generates accurate predictions. The system follows a hybrid approach by combining machine learning for structured numerical data and deep learning for audio-based analysis. The workflow ensures systematic data processing, model training, and prediction in an efficient and scalable manner. The overall workflow begins with data collection and input handling. The system accepts two types of inputs: CSV files containing numerical voice features and audio files in .wav format. For the mini project, the input consists of structured biomedical voice measurements, while for the final project, raw speech recordings are used. The system is designed to handle multiple files as well as compressed ZIP folders, making it flexible for real-world usage. Once the data is uploaded, it is directed to the respective processing pipeline based on its type.

In the machine learning pipeline, the uploaded CSV data undergoes preprocessing steps. Initially, irrelevant attributes such as identifiers are removed, and the remaining features are aligned with the original training dataset. The data is then normalized using a pre-trained scaler to ensure consistency in feature values. After normalization, Recursive Feature Elimination (RFE) is applied to select the most important features that contribute to classification. These selected features are passed into a trained stacking ensemble model, which combines multiple classifiers to generate the final prediction. The output is then formatted into a structured CSV file indicating whether each sample corresponds to a healthy individual or a Parkinson's patient. In the deep learning pipeline, audio data undergoes a series of preprocessing steps. The audio signal is first cleaned by removing silence and unwanted noise. Data augmentation techniques such as adding background noise and pitch shifting are applied to improve model generalization. The processed audio is then converted into Mel Spectrograms, which represent the signal in both time and frequency domains. These spectrograms are normalized and resized to maintain uniformity. The prepared data is then fed into a Convolutional Neural Network (CNN), which extracts hierarchical features and performs classification. For multiple augmented versions of the same audio, a majority voting mechanism is used to determine the final prediction.

The trained models are integrated into a web-based application developed using Flask. The frontend provides an intuitive interface for users to upload data, while the backend handles preprocessing, prediction, and result generation. After processing, the system generates output in the form of downloadable CSV files, ensuring easy interpretation and record-keeping. Each entry in the output clearly indicates the input file or sample along with its predicted class. The workflow ensures seamless interaction between different components of the system, from data input to final output. It emphasizes automation, accuracy, and usability, making the system suitable for both research and practical applications. By following this structured methodology, the proposed system achieves efficient processing and reliable detection of Parkinson's disease using both numerical and audio data.

V. RESULTS AND DISCUSSION



Volume 15, Issue 4, April 2026**|DOI: 10.15680/IJRSET.2026.1504070|**

The machine learning model developed for the mini project shows strong performance in classifying Parkinson's disease using numerical voice features. By applying preprocessing techniques such as normalization and feature selection using Recursive Feature Elimination (RFE), the system successfully reduces irrelevant information and focuses on the most significant attributes. The ensemble approach, which combines multiple classifiers through stacking, enhances the prediction capability of the model.

When tested with unseen data, the model produces consistent and reliable predictions. The system is also capable of processing uploaded CSV files and generating structured output results. The ability to handle real-time input through the web interface makes the system practical and user-friendly. The results indicate that numerical voice features are effective indicators for detecting Parkinson's disease.

The machine learning model developed for the mini project shows strong performance in classifying Parkinson's disease using numerical voice features. By applying preprocessing techniques such as normalization and feature selection using Recursive Feature Elimination (RFE), the system successfully reduces irrelevant information and focuses on the most significant attributes. The ensemble approach, which combines multiple classifiers through stacking, enhances the prediction capability of the model.

When tested with unseen data, the model produces consistent and reliable predictions. The system is also capable of processing uploaded CSV files and generating structured output results. The ability to handle real-time input through the web interface makes the system practical and user-friendly. The results indicate that numerical voice features are effective indicators for detecting Parkinson's disease.

VI. CONCLUSIONS

This study presents an effective AI-based system for the early detection of Parkinson's disease by integrating both machine learning and deep learning techniques. The system utilizes voice-based data in two forms: structured numerical features and raw audio signals. In the machine learning component, feature selection using Recursive Feature Elimination (RFE) combined with ensemble methods such as stacking improves prediction reliability and efficiency. In the deep learning component, audio signals are processed into Mel Spectrograms and analyzed using a Convolutional Neural Network (CNN), enabling the system to capture complex speech patterns associated with Parkinson's disease. The combined approach enhances overall system performance and demonstrates that voice analysis is a powerful and non-invasive biomarker for early diagnosis.

The proposed system eliminates the need for invasive medical procedures and provides a user-friendly web-based platform capable of handling multiple input formats, including CSV files, individual audio files, and ZIP folders. This makes the system suitable for practical applications such as preliminary screening and remote healthcare support. In the future, the system can be further improved by incorporating advanced deep learning architectures, expanding datasets with diverse and multilingual speech samples, and developing real-time mobile or cloud-based applications. These enhancements will increase accessibility, scalability, and effectiveness, making the system more suitable for widespread clinical use.

VII. ACKNOWLEDGMENT

The author expresses sincere gratitude to Ms. M. Nithya, Assistant Professor, Department of Computer Science and Engineering, K.L.N. College of Engineering, for her valuable guidance, continuous support, and encouragement throughout the course of this project. Her insightful suggestions and constant motivation played a vital role in the successful completion of this work. The author also extends heartfelt thanks to the faculty members and peers of the department for their cooperation and assistance during the research process.

REFERENCES

1. Alshammri R, Alharbi G, Alharbi E, Almubark I (2023), 'Machine Learning Approaches to Identify Parkinson's Disease Using Voice Signal Features' in *Frontiers in Artificial Intelligence*, Vol. 6, pp. 1–12.
2. Lim WS, Chiu SI, Peng PL, et al. (2025), 'A Cross-Language Speech Model for Detection of Parkinson's Disease' in *Journal of Neural Transmission*, Vol. 132, pp. 579–590.

3. Srinivasan S, Ramadass P, Mathivanan SK, Selvam KP, Shah MA (2024), 'Detection of Parkinson Disease Using Multiclass Machine Learning Approach' in Scientific Reports, Vol. 14, Article 13813, pp. 1–12.
4. Tougui I, Zakroum M, Karrakchou O, Ghogho M (2024), 'Transformer-Based Transfer Learning on Self-Reported Voice Recordings for Parkinson's Disease Diagnosis' in Scientific Reports, Vol. 14, pp. 1–12.
5. Xu H, Xie W, Pang M, Li Y, Jin L, Huang F, Shao X (2025), 'Non-Invasive Detection of Parkinson's Disease Based on Speech Analysis and Interpretable Machine Learning' in Frontiers in Aging Neuroscience, Vol. 17, pp. 1–12.
6. Mohammed MH, Kadhim MN, Al-Shammery D, Ibaida A (2025), 'Novel Voice Signal Segmentation Based on Clark Distance to Improve Intelligent Parkinson Disease Detection' in Journal of Voice, pp. 1–10.
7. Rizvi SM, Siyam MAU, Jibon FA, Shorna SA, Akter R (2025), 'Voice-Based Parkinson Disease Detection Using SVM & XGBoost' in Proc. International Conference on Quantum Photonics, Artificial Intelligence and Networking (QPAIN), pp. 1–6.
8. Pah ND, Indrawati V, Kumar DK (2023), 'Voice-Based SVM Model Reliability for Identifying Parkinson's Disease' in IEEE Access, Vol. 11, pp. 144296–144305.
9. Iyer A, et al. (2023), 'A Machine Learning Method to Process Voice Samples for Parkinson's Disease Detection' in Scientific Reports, Vol. 13, pp. 1–10.
10. Swain KP, et al. (2024), 'Detecting Parkinson's Disease through Voice Analysis with Machine Learning Algorithms' in Open Biomedical Engineering Journal, Vol. 18, pp. 1–10.
11. Singh N, et al. (2024), 'An Ensemble Technique to Predict Parkinson's Disease Using Optimized Machine Learning Algorithms' in Neurocomputing, Vol. 467, pp. 1–10.
12. Bukhari SNH, et al. (2024), 'Ensemble Machine Learning Approach for Parkinson's Disease Detection Using Voice Signals' in Mathematics, Vol. 12, Issue 10, Article 1575, pp. 1–14.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details