IRJCS: International Research Journal of Computer Science
Volume 13, Issue 01, January 2026
ISSN: 2393-9842
https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1301.09

Innovative Approaches to Enhance Accuracy of Parkinson's Disease

Prof.Gopala krishna P
Assistant Professor, Department of Computer Science and Engineering,
Vemana Institute of Technology, Bengaluru, India
gopala.krishna@vemanait.edu.in
https://orcid.org/0009-0000-8645-1109
Nishithav reddy , Rakshitha B, Ranjitha P, Srividya BV
UG Students, Department of Computer Science and Engineering,
Vemana Institute of Technology, Bengaluru, India
nishithavreddy.cs2022@vemanait.edu.in, rakshithab.cs2022@vemanait.edu.in,
ranjithap.cs2022@vemanait.edu.in, srividyabv.cs2022@vemanait.edu.in

Publication History
Manuscript Reference: IRJCS/RS/Vol.13/Issue01/CSJA26.JACS10088
Research Article | Open Access | Double-Blind Peer Reviewed Article ID: IRJCS/RS/Vol.13/Issue01/CSJA26.JACS10088
Received:12,December 2025,Revised:24,December 2025,Accepted:02 January 2026 Published Online:20 January 2026
https://www.irjcs.com/volumes/Vol13/iss-01/09.CSJA26.JACS10088.pdf
Article Citation:Prof.Gopala,Nishithav,Rakshitha,Ranjitha,Srividya(2026),Innovative Approaches to Enhance Accuracy of Parkinson's Disease, IRJCS: International Research Journal of Computer Science, Volume 13, Issue 01 of 2026 pages 43-47
Doi:>https://doi.org/10.26562/irjcs.2026.v1301.09
BibTeX Key Prof.Gopala@2026Innovative
IRJCS papers should be cited as IRJCS (International Research Journal of Computer Science, AM Publications, India 2026, ISSN 2393-9842, https://doi.org/10.26562/ijrcs.2025.v1301.09 The journal's official abbreviation is IRJCS.
Orcid: https://orcid.org/0009-0004-9398-7488

Abstract: Parkinson's disease (PD) is a slowly worsening neurological condition that affects a person's ability to control movement. Detecting PD at an early stage can make treatment more effective and help patients manage their symptoms better. In this work, we develop a real-time system that analyzes a person's voice to predict the likelihood of Parkinson's disease using machine learning. The system is trained on the UCI Parkinson's dataset, which contains several important vocal features such as jitter, shimmer, and harmonic-to-noise ratio. Multiple machine learning models including Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), and Logistic Regression (LR) are trained and evaluated using accuracy, F1-score, ROC, and AUC metrics. The model with the best performance is deployed in a user-friendly Stream lit web application that allows real-time predictions from simple voice recordings. The proposed system is non-invasive, affordable, and easy to use, making it a promising tool for preliminary screening and early identification of Parkinson's disease.

1. INTRODUCTION

Parkinson's disease (PD) is a long-term neurological condition that slowly reduces a person's ability to control their movements. This occurs as particular neurons that manage movement gradually degenerate. Leading to indications such as shaking, stiffness, slowed motion, and imbalance. Apart from these movement-related symptoms, many individuals may also experience non-motor challenges including sleep disturbances, mood changes, and cognitive difficulties, which affect everyday life. Because PD develops gradually, detecting it early can make a major difference in how well patients respond to treatment. Traditional diagnosis relies on clinical examinations and rating scales like the Unified Parkinson's Disease Rating Scale (UPDRS). Although effective, these methods usually identify PD only after noticeable neural damage has already occurred. With the rise of machine learning (ML), researchers now explore new ways to identify early signs of PD using measurable data such as speech patterns, hand writing, and motion signals. Voice-based diagnosis is especially promising because PD often causes subtle changes in speech long before major motor symptoms appear. Machine learning and deep learning models such as SVM, Random Forest, CNNs, and RNNs are widely used to automatically learn these patterns and improve diagnostic accuracy.

2. LITERATURE REVIEW

Parkinson's disease (PD) ranks as one of the most wide spread disorders involving the degeneration of nerve cells. Affecting millions of individuals worldwide, particularly with in the aging population [1]. The condition leads to progressive motor and non-motor impairments, including tremors, rigidity, bradykinesia, and speech difficulties [2]. Because PD progresses gradually and its symptoms often resemble those of other disorders, early and reliable diagnosis remains a significant clinical challenge [3]. Recent research trends emphasize non-invasive, voice-based diagnostic approaches supported by machine learning (ML) techniques to identify subtle vocal changes associated with early PD stages [4].Many studies have given innovative models, feature sets, and datasets to enhance diagnostic accuracy and accessibility. Speech signals have gained substantial attention as promising biomarkers for early PD detection.

RumanaIslam et al. examined PD classification using 752 audio features taken from steady vowel phonations and compared the performance of K-Nearest Neighbors (kNN) and Support Vector Machine (SVM) classifiers [1]. Their results showed that an optimized kNN model achieved superior accuracy and faster computation time, demonstrating its potential for practical screening. In another study, Linlin Yuan et al. focused on telemedicine applications by analyzing dysphonia patterns in ongoing speech data collected from 252 individuals [6].

3. SYSTEM DESIGN AND METHODOLOGY

The proposed system is designed to operate as a complete pipeline that processes voice recordings and predicts whether the speaker may show signs of Parkinson's disease. The architecture is organized into three layers: the presentation layer, the application processing layer, and the data layer. The process begins with collecting the dataset. The UCI Parkinson's dataset is selected because it provides reliable biomedical voice measurements related to PD. Key parameters such as jitter, shimmer, fundamental frequency, and harmonic-to-noise ratio act as strong indicators of vocal impairments linked to the disease. Before model training, the dataset undergoes preprocessing to remove noise, normalize feature scales, and handle missing values. This ensures consistent input quality and improves overall model performance. The next stage involves extracting meaningful features from audio signals. Using the Librosa library, several acoustic features including MFCCs, zero-crossing rate, pitch frequency, and various spectral and harmonic characteristics are computed. These features are then formed into structured feature sets suitable for ML models. The final stage integrates the best performing model into a Stream lit web interface, allowing users to upload voice samples and receive predictions in real time.

4. IMPLEMENTATION

The system is implemented using Python due to its rich support for machine learning, audio processing, and web deployment. Libraries such as Librosa, Scikit-learn, NumPy, Pandas, and Stream lit form the core development tools.

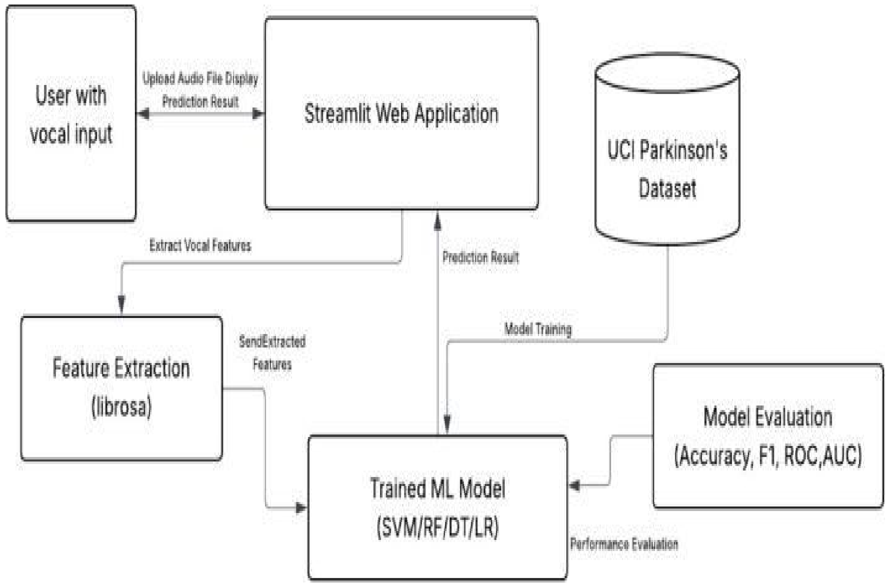

Fig.1. System architecture of the put-forward innovative approaches for enhancing the accuracy of PD

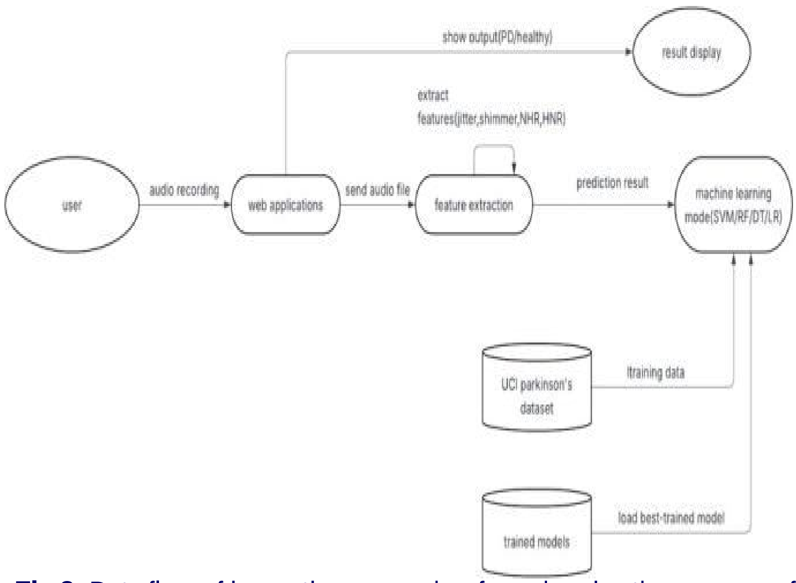

Fig.2. Data flow of innovative approaches for enhancing the accuracy of PD.

IRJCS: International Research Journal of Computer Science ISSN: 2393-9842
Volume 13, Issue 01, January 2026 https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1301.09

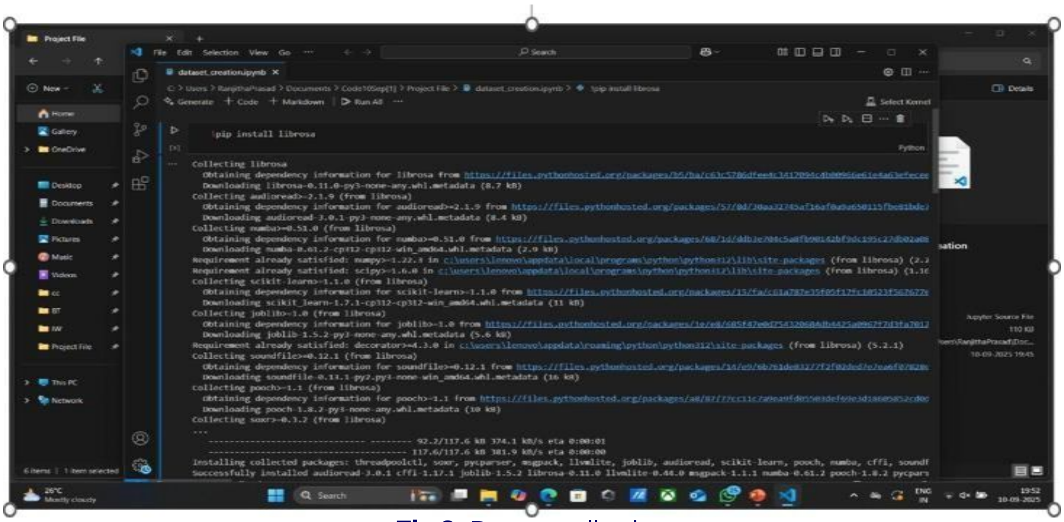

Fig.3. Dataset collection

This step ensures that the extracted audio features are organized in a clear and consistent format, allowing the machine-learning models to process them effectively. For real-time use, the system applies the same feature extraction procedure to the voice recordings uploaded by users, ensuring that the live inputs match the structure of the training data. The machine-learning models are built using the Scikit-learn library. Several commonly used classifiers such as Support Vector Machines (SVM), Random Forest (RF), Logistic Regression (LR), and Decision Trees (DT) are trained and fine tuned. Their key hyperparameters, including the SVM kernel choice, the number of trees in Random Forest, and the depth of Decision Trees, are adjusted using grid search and cross-validation to improve performance and reduce sensitivity to noise. After training, each model is tested based on criteria such as accuracy, precision, recall, F1-score, and ROC–AUC. The system interface also shows confidence levels and feature insights, making the results clear and easy to interpret for both clinicians and users who want an initial assessment.

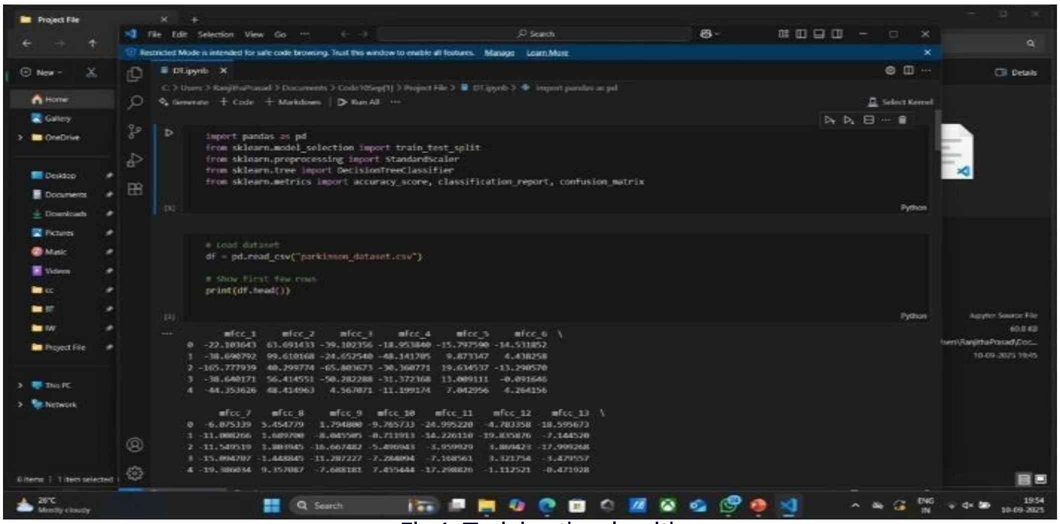

Fig.4. Training the algorithms

5. RESULTS AND DISCUSSION

The system's performance is evaluated using the UCI dataset, which is widely used in PD research. After separating the dataset for training and evaluation portions, multiple machine learning models are compared based on accuracy, precision, recall, F1- score, and ROC–AUC metrics. Among all the models tested, the Support Vector Machine (SVM) achieved the highest accuracy and delivered the most consistent results. Its outstanding performance highlights that SVM can effectively capture the non-linear relationships present in vocal biomarkers of Parkinson's disease. Random Forest also showed competitive accuracy due to its ensemble structure. However, models of lower complexity such as Logistic Regression and Decision Trees delivered lower performance, indicating their limitations in handling complex speech variations. The results demonstrate that integrating vocal features with machine learning can reliably differentiate between healthy and Parkinson's-affected speech. The strong ROC–AUC values further highlight the system's discriminative power.

6. CONCLUSION

Among all the models evaluated, the Support Vector Machine (SVM) delivered the most accurate and reliable results, showing its strength in capturing the complex patterns present in biomedical voice data. This performance is consistent with observations reported in earlier studies. The Random Forest classifier also produced strong outcomes, while simpler models such as Logistic Regression and Decision Trees were less effective in detecting subtle vocal variations linked to Parkinson's disease.

ISSN: 2393-9842
https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1301.09

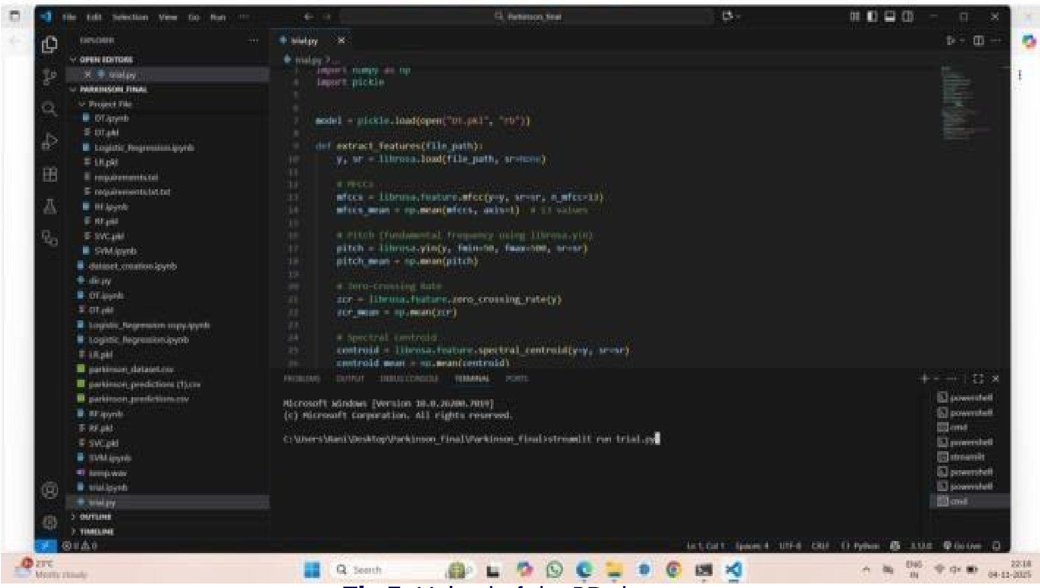

Fig.5. Upload of the PD dataset.

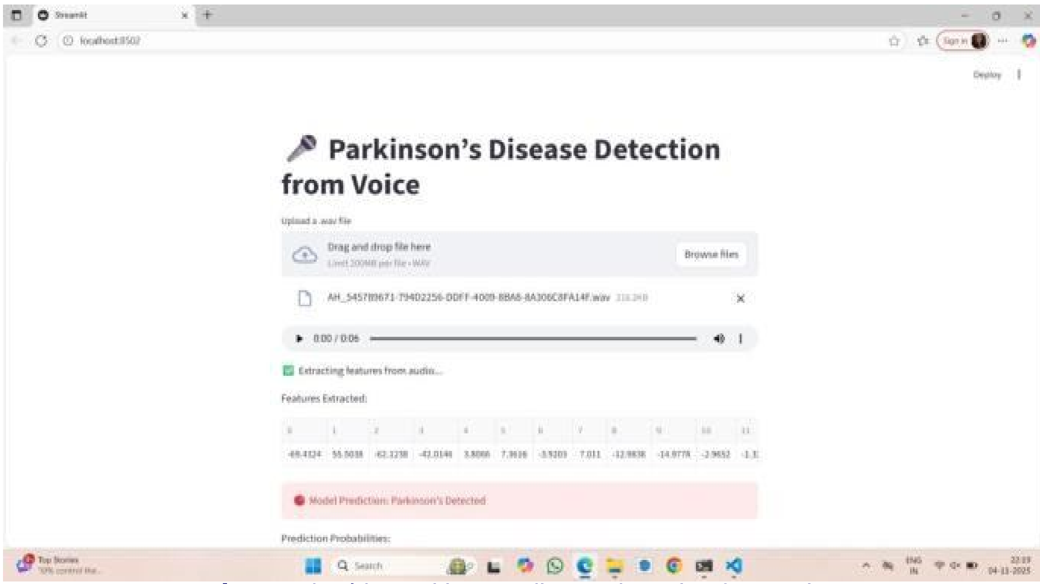

Fig.6. Unhealthy Parkinson's disease detection from voice

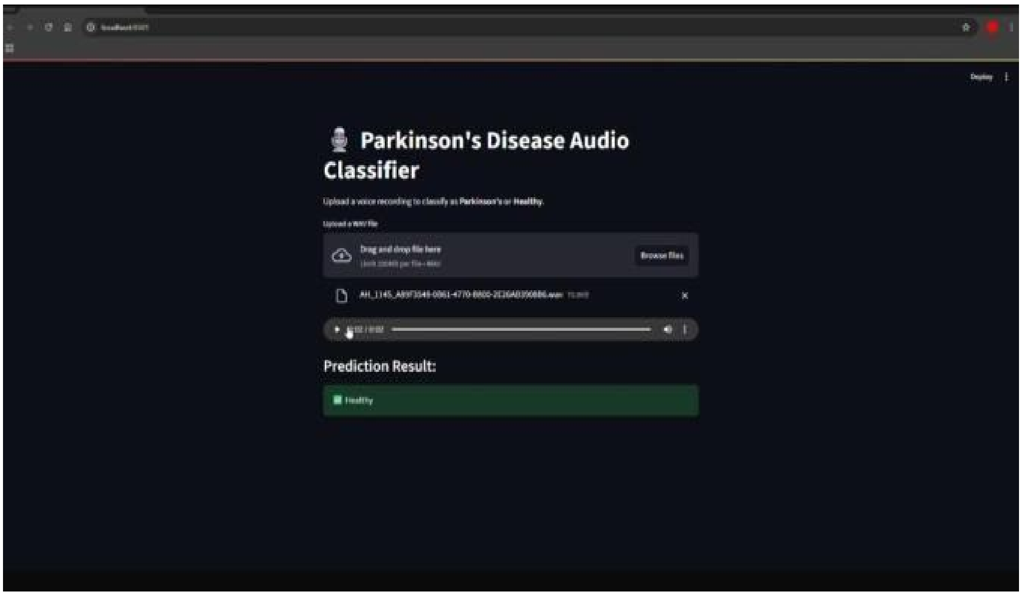

Fig.7. Healthy Parkinson's disease detection from voice.

The best-performing the model was later embedded into a Stream lit web application, allowing the system to generate predictions instantly. This makes the tool suitable for telemedicine and easy preliminary screening. Testing with real-world audio samples confirmed that the application can efficiently analyze user-uploaded recordings and provide quick, meaningful results, demonstrating its potential for use in remote monitoring and clinical decision support. Although the overall performance is promising, there is still room for improvement. Future enhancements may include using larger and more diverse datasets, in corporating various types of speech tasks, and exploring advanced deep-learning models to improve robustness and generalization. Additional developments could involve support for multiple languages, continuous tracking of patients over time, and integration with wearable devices to build a more comprehensive neurological assessment system.

7. FUTURE SCOPES

Although the current system performs effectively, several enhancements can make it even more reliable and clinically useful. Advanced deep learning models, such as CNNs and RNNs, can be explored to automatically learn complex speech patterns and reduce the need for manual feature extraction. Deploying the system on mobile and IoT platforms such as smart phones, smart speakers, or wearable devices could enable continuous voice monitoring and early detection. Collaboration with medical professionals will be essential for clinical validation, ensuring that the system meets medical standards and can be used confidently in real-world healthcare settings. The system can be further improved by making it compatible with mobile devices and Internet of Things (IoT) platforms. Integrating the model into smart phones, wearable sensors, or smart home assistants would allow continuous and real-time monitoring of a user's voice, helping with early detection and regular follow-up assessments. In addition, working closely with medical professionals to clinically verify the system and align it with established diagnostic standards is essential. Such collaboration would support the transition of this technology from a research prototype to a practical tool used in everyday healthcare settings.

ACKNOWLEDGMENT

The authors wish to whole heartedly thank their project supervisor and faculty members for their constant support, guidance, and constructive feedback throughout the course of this work. Their suggestions played a key role in shaping the problem statement, methodology, and experimental design of the system. The authors also extend their gratitude to the team behind the University of California, Irvine Machine Learning Repository for providing the open-source dataset that made the implementation and evaluation possible. Finally, the authors appreciate the encouragement and help received from their institution, classmates, and families, whose support greatly contributed to the successful completion of this project. The authors are also grateful to the laboratory staff for providing the necessary resources and a supportive environment to carry out the experiments smoothly.

REFERENCES

1. R.Islam and E.Abdel-Raheem," Parkinson's disease detection using voice Features and machine learning algorithms,"2023.https://consensus.app/papers/parkinson-%E2%80%99-s-disease-detection-using-voice-features-and-islam-abdel-raheem/79b31c34a8f15e3 1a86153a2764 23d48/
2. S.S.Nayak and A.D.Darji," Identification of Parkinson's disease from speech signal using machine learning approach," International Journal of Speech Technology, vol.26,no.4,pp.981–990,2023 https://consensus.app/papers/identification-of-parkinsons-disease-from-speech-signal-nayak-darji/e123ad011d635223bf906ec1bf3505b1/
3. M.Mahmud et al.," Vocal feature-guided detection of Parkinson's disease,"2023. https://consensus.app/papers/vocal-feature-guided-detection-of-parkinson-%E2%80%99-s-disease-mamun-mahmud/ff868c51d8005ff799c246c25c5774ed/
4. H.Alshammri and G.Alharbi,"Machine learning approaches to identify Parkinson's disease using voice signal features,"2023.https://consensus.app/papers/machine-learning-approaches-to-identify-parkinsons-alshammri-alharbi/ffb1a8135f26502fbdd16 bebec0af2cf/
5. Y.Liu,"Parkinson's disease prediction using machine learning,"2023. https://consensus.app/papers/parkinson-disease-predictionusingmachineyuanliu/8dccb8d1bf61534daa2e93d818f2 f1f8/
6. O.Almeida-Filho, "Detecting Parkinson's disease with sustained phonation,"2023. https://consensus.app/papers/detecting-parkinsons-disease-with-sustained-phonation-almeida-filho/bb9f25063c565785b8432f92b71ed2a2/
7. S.Bukhari and R.Ogudo," Ensemble machine learning approach for Parkinson's disease detection,"2023.[Online]. Availablehttps://consensus.app/papers/ensemble-machine-learning-approach-for-parkinson-%E2%80%99-s-bukhari-ogudo/e4224cfd 3b6e5bac9076be9dbf8de84a/
8. P.Govindu and S.Palwe," Early detection of Parkinson's disease using machine learning,"2023.[Online].Available: https://consensus.app/papers/early-detection-of-parkinsons-disease-using-machine-govindupalwe/bf9b8030e1775e469ea3394b5d3532d6/