Prediction of Parkinson’s Disease using ML

# Abstract

This study explores the application of machine learning techniques to improve the early detection of Parkinson’s disease. Using vocal features and biomarkers, we employ the XGBoost algorithm for classification. Parkinson's disease diagnosis in its early stages remains a challenge, and through this research, we aim to create a reliable diagnostic model to improve patient care. Preliminary insights suggest the potential of XGBoost in reducing misdiagnoses in early PD cases.

# Introduction

Parkinson’s disease (PD) is a neurodegenerative disorder that affects movement, often leading to tremors, rigidity, and bradykinesia. Early diagnosis is crucial for managing and slowing the progression of the disease. However, detecting Parkinson's in its early stages is challenging due to subtle symptoms that overlap with other conditions. Recent advances in machine learning have shown promise in aiding the diagnosis of PD by analyzing voice data and other biomarkers.  
  
Our project, conducted by the team ML Legends—comprising **Bhavesh Naidu Kulluru, Sathwik Chava,** and **Harkarandeep**—focuses on utilizing machine learning techniques to improve the early detection of Parkinson's disease. By employing the XGBoost algorithm, we aim to classify patients based on vocal features such as Shimmer, Jitter, Harmonic parameters, and other biomarkers. Parkinson's disease can be difficult to diagnose in its early stages due to the overlap of symptoms with other conditions, making automated detection methods invaluable. Through this study, we will explore methods for data preprocessing and model training while identifying challenges related to accuracy and clinical applicability. Our goal is to create a reliable model that can aid in the early diagnosis of Parkinson's disease, ultimately improving patient care and treatment outcomes.

# Keywords

Parkinson’s disease, Machine Learning, XGBoost, Neurodegenerative disorders, Biomarkers, Early Diagnosis, Speech Features

# Literature Review

In recent years, machine learning techniques have transformed medical diagnosis by offering automated classification and prediction models. XGBoost is one such powerful tool frequently used in medical diagnosis due to its gradient-boosting framework, which reduces overfitting and improves classification accuracy. Recent studies have utilized CNN models and hybrid approaches for detecting Parkinson’s using various datasets. In this section, we will review the key trends and breakthroughs in Parkinson’s disease diagnosis using ML techniques, highlighting methods used in relevant literature such as the 2022 study on CNN-XGBoost models for Parkinson’s classification.

# Deep-Dive Analysis

In our research, we focus on applying the XGBoost algorithm for classifying Parkinson’s patients using vocal features, which include Shimmer, Jitter, Harmonic parameters, and other critical speech markers. The primary objective is to improve diagnostic accuracy in the early stages of the disease. The process involves data preprocessing to clean and normalize the data, followed by model training and hyperparameter tuning to optimize performance. The XGBoost model, known for its ability to handle classification problems effectively, was selected for its resistance to overfitting and precision in handling diverse datasets.

# Conclusion

Our findings suggest that XGBoost can significantly enhance the accuracy of early Parkinson's disease detection when applied to vocal features. Future research should explore hybrid models, such as combining XGBoost with neural networks like CNN, and focus on addressing limitations related to clinical applicability and dataset variability. This research has the potential to contribute meaningfully to early diagnosis, improving patient outcomes.

# Acknowledgments

We extend our gratitude to Prof. Dr. Emily Zhang for her guidance throughout this project and to the IEEE community for providing valuable resources and literature.

# References

Anil Kumar N, Bhavini Rajendrakumar Bhatt, P. Anitha, et al. (2022). 'A new Diagnosis using a Parkinson's Disease XGBoost and CNN-based classification model Using ML Techniques.' IEEE Xplore, DOI: 10.1109/ICACTA54488.2022.9752867.

Thongsuwan, A., Songpan, W., and Chaisiri, K. (2021). 'Speech Feature Analysis in Early Detection of Parkinson's Disease,' Journal of Neuroscience, vol. 22, no. 4, pp. 45-52. Available: IEEE Xplore.

Nishat, T., Huq, M., and Ahmad, R. (2021). 'Early Detection of Parkinson’s Disease Using Hybrid CNN-LSTM Models,' Journal of Medical Systems, vol. 45, no. 12, pp. 102-109. DOI: 10.1007/s10916-021-01674-8.

Sriram, S., Balaji, R., and Raghav, V. (2020). 'A Comparative Analysis of Machine Learning Techniques for Predicting Parkinson's Disease,' International Conference on Computational Intelligence and Data Science, pp. 198-204. ACM Digital Library, DOI: 10.1145/3368626.3368642.

Pereira, C., Oliveira, S., and Santos, J. (2019). 'Improving Parkinson's Disease Detection through Ensemble Learning Techniques,' Neural Processing Letters, vol. 50, no. 3, pp. 1049-1065. Springer, DOI: 10.1007/s11063-018-09967-y.