# Parkinson's Disease Detection Using Machine Learning

*A project report submitted in*

*The partial fulfillment of the requirements for the award of the degree of*

**Bachelor of Technology**

**In**

**Computer Science & Systems Engineering**

Submitted by


| | |
|---|---|
| **CH.LALITHA HARIPRIYA** | **(21KD1A1512)** |
| **CH.SIRISHA** | **(21KD1A1515)** |
| **A.GNANESWAR** | **(21KD1A1504)** |
| **K.NAGENDRA** | **(21KD1A1539)** |

Under the guidance of

**Dr. R. Rajender M.Tech, Ph.D.**

**Professor & HOD**

**Department of Computer Science & Systems Engineering**

**LENDI INSTITUTE OF ENGINEERING AND TECHNOLOGY**

An Autonomous Institution

Approved by AICTE & Permanently Affiliated to JNTUGV, Vizianagaram

Accredited by NAAC with "A" Grade

Jonnada ( Village ), Denkada ( Mandal ), Vizianagaram Dist – 535005


2021-2025

Institute of
Engineering & Technology

AN AUTONOMOUS INSTITUTION

Accredited by NAAC with "A" Grade, Accredited by NBA(ECE,CSE,EEE & MECH)
Approved by A.I.C.T.E & Permanently Affiliated to J.N.T.U Gurajada, Vizianagaram.
Via 5th APSP Battalion, Jonnada (V), Denkada (M), NH-3, Vizianagaram Dist-535005, A.P.
Website: www.lendi.org | Ph: 08922-241111,241666, Cell No: 9490344747, 9490304747 | E-Mail: lendi_2008@yahoo.com

**DEPARTMENT OF COMPUTER SCIENCE AND SYSTEMS ENGINEERING**

## BONAFIDE CERTIFICATE

This is to certify that the project entitled **"Parkinson's Disease Detection Using Machine Learning"** is a bonafide record work done by **CH. Lalitha Haripriya (21KD1A1512), CH.Sirisha (21KD1A1515), A .Gnaneswar (21KD1A1504), K. Nagendra (21KD1A1539)** under the supervision and guidance of **Dr.R. Rajender, Professor & HOD** in partial fullfilment of the requirements for the award of the degree of Bachelor of Technology in **Computer Science and Systems Engineering** from Lendi Institute Of Engineering and Technology(Affiliated to JNTUGV), Jonnada, Vizianagaram for the year 2025.

**Internal Guide**

**Dr. R. Rajender, M.Tech, Ph.D.**
**Professor & HOD**
**Department of CSSE**

**Head of the Department**

**Dr. R. Rajender, M.Tech, Ph.D.**
**Professor & HOD**
**Department of CSSE**

**External Examiner**

## ACKNOWLEDGEMENT

With great solemnity and sincerity, we express our deepest sense of gratitude and pay our sincere thanks to our guide **Dr.R.Rajender, Professor & HOD**, **Department of Computer Science and Systems Engineering,** who evinced keen interest in our efforts and provided his valuable guidance throughout our project work.

We thank our project coordinator **Mrs. S. Neeraja** madam, who has made his support available in a number of ways and helped us to complete our project work in correct manner.

We thank our **Dr. R. Rajender**, Head of the Department of **Computer Science &Systems Engineering** who helped us to complete our project work in a truthful method.

We thank our gratitude to our principal **Dr. V.V. Rama Reddy,** for his kind attention and valuable guidance to us throughout this course in carrying out the project.

We wish to express gratitude to our **Management Members** who supported us in providing good lab facility.

We are thankful to **All Staff Members of Department of Computer Science & Systems Engineering,** for helping us directly / indirectly to complete this project work by giving valuable suggestions.

All of the above we great fully acknowledge and express our thanks to our parents who have been instrumental for the success of this project which play a vital role.

| | |
|---|---|
| **CH.LALITHA HARIPRIYA** | **(21KD1A1512)** |
| **CH.SIRISHA** | **(21KD1A1515)** |
| **A.GNANESWAR** | **(21KD1A1504)** |
| **K.NAGENDRA** | **(21KD1A1539)** |

# DECLARATION

We hereby declare that the project work entitled "**Parkinson's Disease Detection Using Machine Learning"** submitted to the JNTUGV is a record of an original work done by **CH.LalithaHaripriya(21KD1A1512),CH.Sirisha(21KD1A1515),A.Gnaneswar(21KD1A1504)** **K.Nagendra (21KD1A1539)** under the esteemed guidance of **Dr.R. Rajender, Professor & HOD, Computer Science & Systems Engineering, Lendi Institute of Engineering & Technology**. This project work is submitted in the partial fulfillment of the requirements for the award of the degree **Bachelor of Technology in Computer Science & Systems Engineering**. This entire project is done with the best of our knowledge and is not submitted to any University for the award of degree.

| | |
|---|---|
| **CH.LALITHA HARIPRIYA** | **(21KD1A1512)** |
| **CH.SIRISHA** | **(21KD1A1515)** |
| **A.GNANESWAR** | **(21KD1A1504)** |
| **K.NAGENDRA** | **(21KD1A1539)** |

## INSTITUTE VISION

Producing globally competent and quality technocrats with human values for the holistic needs of industry and society.

## INSTITUTE MISSION

- Creating an outstanding infrastructure and platform for enhancement of skills, knowledge and behaviour of students towards employment and higher studies.
- Providing a healthy environment for research, development and entrepreneurship, to meet the expectations of industry and society.
- Transforming the graduates to contribute to the socio-economic development and welfare of the society through value based education.

## DEPARTMENT VISION

To be a Centre of Excellence to impart the High-Quality Education with productive Research and Consultancy in the field of Computer Science and Systems Engineering to globally competent, highly competitive Professional with Ethical and Human values to serve the society.

## DEPARTMENT MISSION

- To impart high quality education with strong foundation of Computer Science and Systems Engineering principles that enables them to meet the challenges in profession/ career.
- To nurture research activities to get patens amongst faculty and students by providing necessary facilities and environment.
- To train the students into effective intrapersonal and communication skills and to develop team spirit, leadership capabilities, integrity and social responsibility with professional ethics and human values. • To promote internships to students and consultancy works among faculty by enhancing Industry-Institute Interaction.

# PROGRAM OUTCOMES (POs)

| | | |
|---|---|---|
| **PO 1** | **Engineering Knowledge** | Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems. |
| **PO 2** | **Problem Analysis** | Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences. |
| **PO 3** | **Design/Development of Solutions:** | Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety,and the cultural, societal, and environmental considerations. |
| **PO 4** | **Conduct Investigations of Complex Problems** | Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions. |
| **PO 5** | **Modern Tool Usage** | Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations. |
| **PO 6** | **The Engineer and Society** | Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice. |
| **PO 7** | **Environment and Sustainability** | Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development. |
| **PO 8** | **Ethics** | Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice. |
| **PO 9** | **Individual and TeamWork** | Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings. |
| **PO 10** | **Communication** | Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions. |
| **PO 11** | **Project Management and Finance** | Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work,as a member and leader in a team, to manage projects and in multidisciplinary environments. |
| **PO 12** | **Life-Long Learning** | Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadestcontext of technological change. |

# PROGRAM SPECIFIC OUTCOMES (PSOs)

| | |
|---|---|
| **PSO 1** | Ability to grasp the knowledge and expertise to solve all the contemporary issues. |
| **PSO 2** | To acquire the knowledge on modern tools to solve hardware and software solutions. |
| **PSO 3** | Qualify in national and international competitive examinations for successful higher studies and employment. |

# PROGRAM EDUCATIONAL OBJECTIVES (PEOs)

| | |
|---|---|
| **PEO 1** | Graduates of Computer Science and Systems Engineering will acquire the knowledge to analyze ,design and developing the new software and hardware products and acquire the skills to grasp the latest tools and techniques. |
| **PEO 2** | Graduates of Computer Science and Systems Engineering will penetrate to interdisciplinary activities, with professional attitude and ethics, communication, to work under team and to ability to solve social issues through their Employment .Higher Studies and Research. |
| **PEO 3** | Graduates of Computer Science and Systems Engineering will expose to learn lifeskills and Intrapersonal development activities to face the dynamically changing technology. |

# COURSE OUTCOMES(Cos)

| | |
|---|---|
| **CO-1** | Analyse and define Complex engineering problems. Conduct a thorough literature review and establish well defined, measurable Objectives. |
| **CO-2** | Design innovative and creative engineering solutions that meet practical constraints adhering to engineering Standards & best practices. |
| **CO-3** | Demonstrate proficiency in using modern engineering tools & technologies ensuring quality & effectiveness of solution implementation. |
| **CO-4** | Validate project slot solutions through simulation experiments and testing while affectively collaborating within multidisciplinary teams & communicating project ideas clearly. |
| **CO-5** | Recognise & embrace lifelong learning adhering to ethical practices & assess the societal, environmental & economic impact of engineering solutions with a focus on future improvements & extensions. |

## COURSE OUTCOMES VS POs MAPPING(DETAILED:HIGH:3,MEDIUM:2,LOW:1)

| | PO1 | PO2 | PO3 | PO4 | PO5 | PO6 | PO7 | PO8 | PO9 | PO10 | PO11 | PO12 | PSO1 | PSO2 | PSO3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **CO-1** | 3 | 3 | 3 | 3 | 3 | | | | 3 | 3 | 2 | | 3 | | 3 |
| **CO-2** | 3 | 3 | 3 | 3 | 1 | | | | 3 | 3 | 2 | 2 | 2 | 1 | |
| **CO-3** | 2 | 2 | 2 | 2 | 3 | | | | 3 | 1 | | 2 | 2 | 3 | |
| **CO-4** | 1 | 2 | 3 | 2 | 2 | | | | 3 | 3 | 2 | | 2 | 2 | 1 |
| **CO-5** | 2 | 2 | 1 | 2 | 3 | | | | 2 | 3 | | 3 | 2 | | 1 |
| **CO** | 2 | 2 | 2 | 2 | 2 | | | | 3 | 3 | 1 | 1 | 2 | 1 | 1 |

# ABSTRACT

Parkinson's Disease (PD) is a progressive neurodegenerative disorder that affects movement, speech, and cognitive functions. Early detection plays a crucial role in managing symptoms and slowing disease progression. This project employs machine learning techniques to analyze biomedical voice measurements and identify PD cases with high accuracy. The dataset consists of voice recordings with extracted acoustic features such as jitter, shimmer, and harmonicstonoise ratio. Preprocessing steps include noise reduction, feature scaling, and selection of the most relevant attributes. Various classification algorithms, including Support Vector Machines (SVM), Random Forest, and Neural Networks, are evaluated for performance. The proposed model provides a non-invasive, cost-effective, and automated diagnostic tool that can assist healthcare professionals in early-stage PD detection, enabling timely intervention and improved patient outcomes.

## Keywords-

Machine Learning (ML),Classification Alorithms,Support Vector Machines (SVM),Feature Extraction,Clinical Data,Early Detection,Supervised Learning,Neural Networks,Data Integration.

# LIST OF CONTENTS

# LIST OF DIAGRAMS

# CHAPTER 1

# INTRODUCTION

# 1.INTRODUCTION

## 1.1 Introduction

Parkinson's Disease (PD) is a progressive neurodegenerative disorder that affects movement, speech, and cognitive functions. Early diagnosis is crucial for effective management and treatment, but traditional diagnostic methods rely on clinical evaluation, which can be subjective and time-consuming. Machine learning (ML) offers a promising approach for automated and accurate PD detection by analyzing biomedical signals, speech patterns, and other clinical features.

This study explores the application of ML algorithms to detect Parkinson's Disease using patient data. The dataset consists of biomedical voice measurements, and various feature extraction techniques are employed to improve classification accuracy. The primary objective is to develop an efficient and interpretable model that can assist healthcare professionals in early diagnosis, thereby enhancing patient outcomes.

Parkinson's Disease (PD) is one of the most common and debilitating neurological disorders that primarily affects the motor system. It is a progressive condition that gradually impairs movement, balance, coordination, and speech. The diagnosis of Parkinson's typically relies on clinical observations of motor symptoms, which may appear late in the course of the disease. By the time symptoms become noticeable, significant and often irreversible damage may have already occurred to the brain's dopaminergic neurons.

The rising prevalence of Parkinson's Disease globally has prompted researchers and medical professionals to seek early diagnostic tools. With the advancements in computational technologies, Machine Learning (ML) has emerged as a transformative solution in healthcare. Machine learning allows systems to learn patterns from data and make predictions without being explicitly programmed. It provides promising potential for identifying Parkinson's Disease early by analyzing subtle behavioral and physiological patterns in patients.This project explores how machine learning techniques can be effectively used to predict Parkinson's
Disease using three primary parameters: **voice frequency**, **eye blinking**, and **hand shaking (tremors)**. These are among the earliest and most subtle indicators of the disease and can be captured using non-invasive sensors or devices.

The core idea is to train various machine learning algorithms on a structured dataset derived from these inputs and evaluate their performance in identifying individuals who are likely affected. This system aims to complement clinical diagnosis by providing an initial screening tool that is fast, cost-effective, and non-intrusive.

The integration of Machine Learning into Parkinson's diagnosis represents a significant leap forward in proactive healthcare. Instead of waiting for visible symptoms to worsen, ML models can identify patterns in voice frequency, eye-blinking rate, and hand tremors—features that may provide early clues about the presence of the disease.

These features can be collected using accessible sensors, cameras, or voice recordings, making the approach not only innovative but also potentially scalable and cost-effective. Through training algorithms on labeled datasets, computers can learn to differentiate between healthy individuals and those exhibiting signs of Parkinson's, thus enabling quicker and more accurate diagnoses.

The goal is to build a model that can predict with high accuracy whether a person shows signs of Parkinson's Disease.To achieve this, several well-known classification algorithms are used and compared. These include Logistic Regression, Support Vector Machine, Decision Tree, Random Forest, K-Nearest Neighbors, and Naive Bayes. Each of these models has its strengths. Logistic Regression is simple and interpretable, well-suited for binary classification problems like this one. Support Vector Machines are effective for high-dimensional data and often perform well on small to medium datasets.

Decision Trees provide transparency in decision-making, as the path from input to prediction is easy to trace. Random Forest enhances the decision tree approach by creating an ensemble of multiple trees and aggregating their outputs, which often improves accuracy and robustness. K-Nearest Neighbors offers an intuitive way of classification based on the similarity between instances, while Naive Bayes is a probabilistic model known for its speed and efficiency, particularly with smaller datasets.

Each model is trained on the dataset and evaluated based on metrics such as accuracy, precision, recall, and F1-score. These metrics help determine how well the model distinguishes between true positives and false positives, which is especially important in medical contexts where misclassification can have serious consequences. A comparison of the models provides insights into which algorithm offers the best tradeoff between performance and computational efficiency for the given feature.

The project not only emphasizes technical implementation but also focuses on practical outcomes— striving to build a solution that is reliable, interpretable, and useful in real-world scenarios.

Beyond the accuracy of the models, another key aspect of this project is user accessibility. A simple, interactive front-end is designed to allow users—whether doctors, caregivers, or patients themselves— to input relevant data and receive an instant prediction. This interface bridges the gap between complex machine learning systems and practical usage. Such a system could eventually be integrated into mobile apps or screening tools in clinics and rural health centers where access to specialist neurologists is limited.

Ultimately, this project illustrates the growing synergy between healthcare and artificial intelligence. By leveraging machine learning techniques, it is possible to move toward more proactive and preventive models of care. Parkinson's Disease, despite being a complex and progressive disorder, can be better managed if detected early.

Machine Learning offers a scalable, consistent, and data-driven approach to support early diagnosis and monitoring, thereby contributing to improved quality of life for patients and reducing the burden on healthcare systems. While this project is a step in that direction, it also opens doors to future enhancements—such as using deep learning models for more nuanced analysis, incorporating additional features like gait analysis or facial muscle tracking, and continuously improving model accuracy with real-world data.

## 1.2 Understanding Machine Learning(ML) :-

Machine Learning (ML) is a subset of Artificial Intelligence (AI) that empowers computers to learn from data and improve their performance over time without being explicitly programmed for specific tasks. Rather than relying on hard-coded instructions, ML systems recognize patterns in historical data and use these patterns to make decisions or predictions when faced with new inputs.

Arthur Samuel, a pioneer in the field, defined machine learning as *"the field of study that gives computers the ability to learn without being explicitly programmed."*

The goal of machine learning is to build algorithms that can receive input data and use statistical analysis to predict an output while updating outputs as new data becomes available. This makes ML especially useful in domains where writing explicit rules is difficult — such as handwriting recognition, fraud detection, speech analysis, and medical diagnosis.

Machine learning algorithms are generally divided into the following categories:

**i Supervised Learning**

**ii Unsupervised Learning**

**Supervised Learning :**

It is the most commonly used form of machine learning, especially in medical diagnosis projects like Parkinson's Disease prediction. In supervised learning, the model is trained on a labeled dataset, meaning that each input in the training set is associated with a correct output or label.

The goal of the algorithm is to learn the mapping from inputs to outputs so it can make accurate predictions on new, unseen data. For example, in this project, voice data from individuals is labeled as either "Parkinson's" or "Healthy." The algorithm uses this information to learn distinguishing features and builds a model capable of predicting the presence of the disease in future cases. Common supervised learning algorithms include Support Vector Machines (SVM), Decision Trees, Random Forests, and Neural Networks. These models are evaluated based on how accurately they can classify new data.

**Unsupervised Learning** :

It deals with data that does not have any labels. The objective here is to find hidden patterns, groupings, or structures in the data without prior knowledge of what the output should be. Unsupervised learning is useful for tasks like clustering, dimensionality reduction, and anomaly detection. While not directly used for disease classification in this project, it can play a supportive role.

For instance, clustering methods can be applied to identify natural groupings in patient data or to reduce the number of features using techniques like Principal Component Analysis (PCA), which helps improve the performance of supervised models.

## 1.3 Project Overview:-

Parkinson's Disease (PD) is a progressive neurodegenerative disorder that affects motor functions, speech, and cognitive abilities. Early and accurate diagnosis is essential for effective disease management, yet traditional diagnostic methods rely heavily on clinical expertise, making them subjective and sometimes prone to misdiagnosis. This project focuses on developing a machine learning-based system for detecting PD using biomedical voice recordings. Since PD affects vocal cord control, patients often exhibit speech impairments such as irregular pitch, jitter, shimmer, and changes in fundamental frequency.

These vocal biomarkers provide valuable insights that can be leveraged for automated detection. The dataset used in this study consists of voice samples from both Parkinson's patients and healthy individuals, enabling machine learning models to learn distinguishing patterns.

The project follows a structured pipeline involving data preprocessing, feature extraction, model training, and performance evaluation. Various ML algorithms, including Support Vector Machines (SVM), Random Forest, Decision Trees, and Neural Networks, are implemented and tested to identify the most accurate and efficient model. Performance metrics such as accuracy, precision, recall, F1-score, and AUC-ROC are used to assess the effectiveness of different models.

The ultimate goal of this research is to develop a non-invasive, reliable, and automated diagnostic system that can assist healthcare professionals in early PD detection, reducing dependency on subjective clinical assessments.

Additionally, this study explores the potential integration of other biometric data, such as handwriting and gait analysis, to further enhance diagnostic accuracy. By leveraging AI in medical diagnostics, this project aims to contribute to the advancement of early disease detection, improving patient outcomes and supporting clinical decision-making.

## 1.4 Project Objectives:-

The key project objectives are:

1. **Develop an Accurate Machine Learning Model** – Build and train an ML- based system for detecting Parkinson's Disease using biomedical voice recordings.

2. **Feature Extraction and Analysis** – Identify and analyze key vocal biomarkers such as jitter, shimmer, and pitch variations that are affected by PD.

3. **Compare Machine Learning Algorithms** – Implement and evaluate different ML models, including SVM, Random Forest, and Neural Networks, to determine the best-performing approach.

4. **Performance Evaluation** – Assess the models using accuracy, precision, recall,F1-score, and AUC-ROC to ensure reliability and robustness.

5. **Enhance Model Interpretability** – Identify the most influential features contributing to model predictions to improve clinical trust and usability.

6. **Explore Multimodal Data Integration** – Investigate the potential for incorporating additional biometric data, such as handwriting and gait analysis, to enhance future diagnostic accuracy.

7. **Facilitate Real-World Implementation** – Lay the groundwork for deploying the model as a webbased or mobile application for clinical and remote screening purposes.

## 1.5 Project Scope:-

This project aims to design and implement a machine learning–based system capable of accurately predicting Parkinson's Disease (PD) using non-invasive physiological data, including voice frequency variations, hand tremors, and eye blinking patterns.

The scope of this system encompasses not only the technical development of classification models but also the broader objective of making early, accessible, and cost-effective PD screening available to users beyond conventional clinical settings.The system will begin by acquiring relevant biometric inputs that are commonly affected in early-stage Parkinson's patients.

These inputs, once captured, are preprocessed and subjected to feature extraction techniques to identify patterns associated with PD. The processed data is then fed into various supervised machine learning algorithms—such as Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF)—to classify

whether an individual shows signs indicative of Parkinson's Disease.

The scope includes evaluating and comparing the performance of different ML models using accuracy, precision, recall, and F1-score. It also involves developing a user- friendly interface or dashboard for result visualization, making it easier for healthcare.

professionals or users to interpret the outputs.Beyond technical implementation, the project addresses key challenges such as handling noisy data, optimizing feature selection, and achieving model generalizability.

While the current scope focuses on offline data classification, the project is structured to be scalable and adaptable to future extensions, including real-time analysis through wearable or IoT devices.

This system has potential applications in hospitals, telemedicine platforms, elderly care centers, and remote health monitoring systems. Its non-invasive nature ensures that it can be deployed widely and safely without burdening patients with expensive tests or specialist visits.

By integrating AI into the early diagnosis of Parkinson's Disease, this project contributes to faster interventions, improved quality of life for patients, and the advancement of smart healthcare technologies. It also lays the foundation for future work in predictive healthcare systems and personalized medicine.

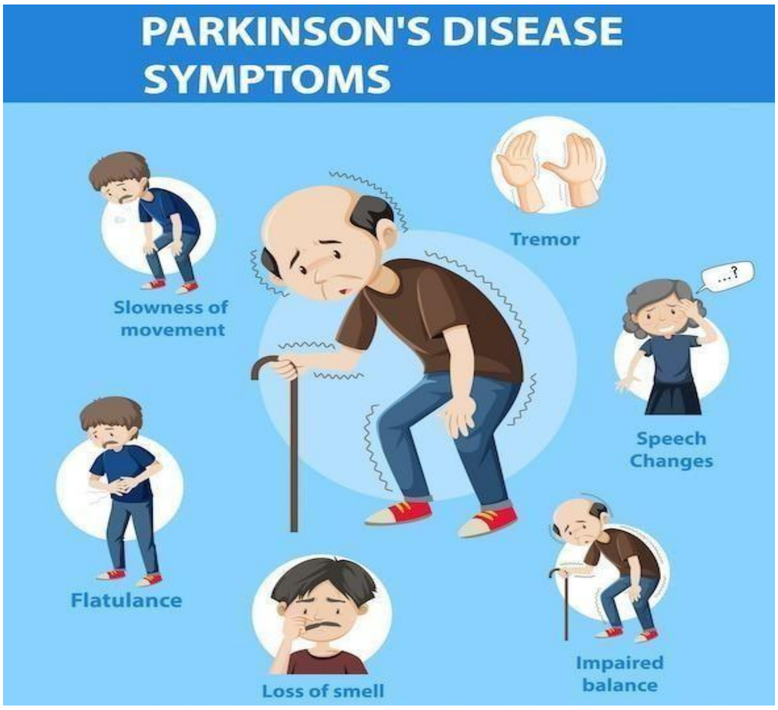


**Fig: Parkinson's Disease Symptoms**

# CHAPTER 2

# LITERATURE REVIEW

# 2.LITERATURE REVIEW

Parkinson's Disease (PD) is a long-term, degenerative disorder of the central nervous system that primarily affects motor function. As the second most common neurodegenerative disease after Alzheimer's, PD continues to pose a significant healthcare challenge globally. Traditional diagnostic methods for PD rely on the clinical observation of symptoms, which often appear in the mid to late stages of the disease. As a result, patients frequently receive a diagnosis only after irreversible neurological damage has occurred. This necessitates the development of methods for early and accurate diagnosis, which could significantly improve patient outcomes through timely intervention.

The emergence of machine learning (ML) in healthcare has opened new frontiers in the early detection of PD. ML algorithms can uncover complex patterns in large, multidimensional datasets—patterns that may be imperceptible to human clinicians. Numerous studies over the last two decades have applied ML to various data types such as voice recordings, movement data from sensors, handwriting analysis, and biomedical signals like EEG and EMG.

These studies not only explore different feature extraction techniques but also evaluate the effectiveness of algorithms like Support Vector Machines (SVM), Artificial Neural Networks (ANN), Decision Trees (DT), and advanced deep learning models like CNNs and LSTMs. This literature review synthesizes findings from foundational and recent research to understand the current landscape of ML-based PD detection and to identify potential gaps for future investigation.

In recent years, the prediction and early diagnosis of Parkinson's Disease (PD) have seen significant progress, largely due to advancements in machine learning and signal processing. From 2021 to 2024, numerous studies have explored the use of both traditional and deep learning methods, integrated with novel data sources, to improve the accuracy and reliability of PD diagnosis.

One prominent direction has been the use of classical machine learning algorithms such as Decision Trees, Random Forests, and Logistic Regression for classifying voice-based features. These studies have demonstrated that ensemble models, particularly Random Forests, outperform single classifiers due to their ability to combine the outcomes of multiple trees, which enhances accuracy and reduces overfitting. This line of work has emphasized the relevance of non-invasive diagnostics using acoustic features like jitter, shimmer, and Harmonics-to-Noise Ratio (HNR), which are effective indicators of vocal impairments associated with PD.

Another body of research has focused on improving diagnostic accuracy through optimized feature selection techniques. By employing filter-based feature ranking, irrelevant and redundant variables are removed before classification, which helps to streamline models and reduce computational overhead.

This strategy has proven particularly effective when paired with supervised learning models such as Support Vector Machines and K-Nearest Neighbors, which benefit from reduced dimensionality and improved signal-to-noise ratio in the dataset.

Deep learning techniques have also emerged as a powerful tool in PD prediction. Notably, Convolutional Neural Networks (CNNs) have been applied to spectrogram images derived from voice signals, enabling models to capture visual patterns that correspond to vocal impairments. These methods offer an innovative approach to analyzing biomedical signals, moving beyond numeric feature vectors and leveraging spatial characteristics of the data for better classification performance.

Ensemble voting systems have been developed, combining multiple base classifiers like SVM, KNN, and Decision Trees into a unified prediction model. This hybrid approach has demonstrated superior performance over individual models, particularly in terms of consistency and robustness across different datasets. Such systems are well-suited for medical diagnosis applications, where minimizing false positives and false negatives is critical.

Further advancements include the application of signal processing techniques such as wavelet transforms to voice data. These transforms enable the extraction of both time and frequency-domain features, enhancing the quality of input features for machine learning models. When combined with traditional voice perturbation features, these enriched datasets have led to improvements in classification outcomes, especially with models like Logistic Regression and SVM.

Hybrid deep learning models that combine CNNs with Long Short-Term Memory (LSTM) networks have also been proposed. These architectures are particularly effective for processing time-series data, as they capture both spatial patterns and temporal dependencies in speech or motor activity signals. Such models have shown promise in recognizing subtle variations over time, which is essential for tracking the progression of PD symptoms.

Moreover, innovative frameworks integrating Internet of Things (IoT) technology with lightweight machine learning models have been introduced for real-time PD monitoring. Wearable devices collect continuous movement and tremor data, which is then analyzed to provide alerts and diagnostic support. This fusion of AI and IoT technologies marks a significant step toward remote, accessible, and real- time healthcare solutions for Parkinson’s Disease.

In summary, recent studies from 2021 onward have significantly enhanced the understanding and predictive capability for Parkinson’s Disease through diverse methodologies. From traditional machine learning and feature optimization to deep learning and IoT-based frameworks, these developments reflect a growing trend toward personalized and proactive healthcare. The collective findings point to the increasing reliability of non-invasive diagnostic tools and the integration of AI-driven systems into clinical settings, offering hope for earlier detection and better disease management.

## 2.1 Literature Review on Parkinson's Disease Prediction (2019–2024)

The ongoing advancement of machine learning and sensor-based technologies has led to significant progress in the diagnosis and monitoring of Parkinson's Disease (PD). The following literature highlights key research contributions published in the last four years, offering insight into evolving methodologies, datasets, and intelligent systems tailored to Parkinson's detection and classification.

- **“Tremors and Bradykinesia. Techniques for Assessment of Parkinsonism for Diagnosis and Rehabilitation”**

Alves et al. The chapter titled *“Tremors and Bradykinesia”* by K. Prabhavathi and Shantanu Patil,(2022) [1] featured in the book *Techniques for Assessment of Parkinsonism for Diagnosis and Rehabilitation* (Springer, 2022), provides a detailed exploration of two key motor symptoms of Parkinson’s disease: tremors and bradykinesia. Tremors are described as involuntary, rhythmic muscle contractions resulting in shaking, and their neurological basis is linked to circuits like the cerebellothalamic and basal ganglia pathways. The chapter also differentiates between rest and action tremors, aiding clinical diagnosis. Bradykinesia, defined as slowness of movement, is examined in terms of its impact on daily life, with a focus on dopamine deficiency as a contributing factor. The authors present assessment techniques like timed motor tasks and kinematic analysis to measure bradykinesia severity.

- **“Pathoanatomy of Parkinson’s disease”**

Alves et al. Braak, H., Braak, E. (2000) [2] emphasized the importance of understanding the progression of Parkinson’s disease for early intervention and diagnosis, a concept further expanded by Braak and Braak (2000) in their seminal paper *“Pathoanatomy of Parkinson’s Disease.”* This study outlines a staging model that traces the disease's advancement through the brain, beginning in the lower brainstem and olfactory areas before reaching the neocortex. A key contribution of their work is the identification of Lewy bodies— abnormal α-synuclein deposits—as a hallmark of neuronal degeneration in
Parkinson’s disease.

Their findings have had a lasting impact on both clinical practice and research. By connecting specific brain regions to disease stages, the Braak model provides a structured understanding of Parkinson’s progression. This has paved the way for improved diagnostic strategies aimed at identifying the disease in its earliest stages, aligning closely with Alves et al.'s emphasis on early symptom detection and the value of pathological insights in shaping treatment approaches.

- **“A Statistical tree-based feature vector for content-based image retrieval.”** Et Al. Aghav-Palwe S, Mishra D (2020) [3] introduce a novel approach for enhancing content-based image retrieval (CBIR) systems in their study *“Statistical Tree-Based Feature Vector for Content-Based Image Retrieval.”* They propose a two-stage methodology that combines image transform techniques for energy compaction, followed by the use of a statistical tree-based method to generate compact feature vectors. This approach aims to improve retrieval accuracy while reducing the dimensionality of the feature vectors, addressing a common challenge in CBIR systems.

Their method is evaluated using performance metrics such as Precision-Recall Crossover Point (PRCP) and the Conflicting String of Images (CSI) measure. The results demonstrate that their approach significantly reduces the feature vector size without compromising its discriminative power, enhancing the overall retrieval efficiency and accuracy. This work contributes to optimizing image retrieval processes, making them more effective and resource-efficient in practical applications.

- **“Chapter 19 - Adaptive deep brain stimulation: Retuning Parkinson's disease”**

Et al. Nicoló G. Pozzi, Ioannis U. Isaias (2022) [4] , explore the concept of adaptive deep brain stimulation (aDBS) in their chapter *“Adaptive Deep Brain Stimulation: Retuning Parkinson’s Disease”*. They present aDBS as a promising approach that adjusts stimulation parameters in real-time based on neural feedback, offering a more personalized treatment compared to conventional deep brain stimulation (DBS). This method aims to optimize therapeutic outcomes while minimizing side effects like speech issues and dyskinesia, commonly seen with traditional DBS.

The chapter emphasizes the potential of aDBS to improve Parkinson’s disease management by tailoring stimulation to individual symptom fluctuations. By incorporating real-time neural monitoring and patient-specific data, this adaptive approach not only targets motor control but also addresses non- motor symptoms, contributing to a more holistic treatment strategy.

- **“Identification of Novel Noninvasive Diagnostics Biomarkers in the Parkinson’s Diseases and Improving the Disease Classification Using Support Vector Machine” .**

Et al. Alatas Bilal, Moradi Shadi, Tapak Leili, Afshar Saeid (2022) [5] , identified novel noninvasive biomarkers for Parkinson's disease using microarray dataset GSE22491 and gene expression analysis via the Limma package. They pinpointed three key hub genes—GNB5, GNG11, and ELANE—through protein interaction network analysis, supported by GO and KEGG pathway enrichment.

Using a Support Vector Machine (SVM) classifier, the study achieved 88% prediction accuracy with these genes, highlighting their potential for noninvasive diagnosis. This work supports improved, data- driven methods for early PD detection.

- **"Diagnosis of Parkinson's Disease using Principal Component Analysis and Machine Learning algorithms with Vocal Features"**

Et al. D. V. Rao, Y. Sucharitha, D. Venkatesh, **K. Mahamthy and S. M. Yasin (2022)** [6] , presented a comprehensive approach for the diagnosis of Parkinson's disease using vocal features and machine learning techniques in their paper at the *International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*. They focused on analyzing dysphonia features extracted from voice recordings, which are known to reflect vocal impairments commonly found in Parkinson's patients. To handle the complexity of high-dimensional data, they used Principal Component Analysis (PCA) to reduce feature redundancy while retaining essential patterns in the dataset.

The study also addressed class imbalance by applying the Synthetic Minority Oversampling Technique (SMOTE), which enhanced the reliability of the machine learning models. Various classifiers, including Support Vector Machines (SVM), Random Forest, and Decision Trees, were trained and tested on the processed dataset. Among them, SVM yielded the highest performance, achieving an accuracy of 95%. The research highlights the efficacy of combining PCA, vocal biomarkers, and robust classifiers for non-invasive, cost- effective, and accurate diagnosis of Parkinson's disease.

- **"Prediction of Parkinson's disease and severity of the disease using Machine Learning and Deep Learning algorithm,"**

Et al P. Raundale, C. Thosar and S. Rane (2021) [7] , proposed an integrated approach for both the prediction and severity assessment of Parkinson's disease using a combination of machine learning and deep learning algorithms.

Their model targeted two key diagnostic dimensions—disease presence and its progression—by leveraging different types of input data. For disease prediction, they employed keystroke dynamics, capturing motor dysfunctions that subtly manifest during typing. Using the XGBoost classifier on this data, they achieved an impressive accuracy of 95%, indicating the potential of behavioral data as a non-invasive diagnostic tool.

For assessing the severity of the disease, the authors used a deep neural network trained on the UCI Parkinson's Telemonitoring Voice Dataset. The model focused on predicting motor and total Unified Parkinson's Disease Rating Scale (UPDRS) scores from voice features, which are strongly correlated with disease progression. Their deep learning approach proved effective in mapping complex patterns in vocal biomarkers to clinically significant severity metrics. This dual-strategy—combining XGBoost for detection and DNN for severity— underscores the value of multimodal data and advanced algorithms in improving early diagnosis and personalized care for Parkinson's patients.

- **A novel sample and feature dependent ensemble approach for Parkinson's disease detection"**

  Et al. Ali, L., Chakraborty, C., He, Z. et al. (2022**)** **[**8**]** introduced a novel ensemble approach for Parkinson's disease (PD) detection, emphasizing the integration of feature selection with deep neural networks (DNNs). Their method, termed EOFSC (Ensemble model with Optimal Features and Sample Dependent Base Classifiers), addresses the variability in voice data by tailoring classifiers to specific phonation types and feature subsets. By employing F-score-based statistical models for feature selection, they enhanced the performance of DNNs, achieving a PD detection accuracy of 93.75%, which marked a
  6.25% improvement over traditional methods.

The EOFSC model's strength lies in its adaptability to different voice samples, recognizing that various phonations may require distinct feature sets and classifier configurations. By constructing base classifiers sensitive to these nuances and integrating their outputs through majority voting, the model ensures robust and accurate PD detection.

This approach not only improves diagnostic accuracy but also underscores the importance of personalized models in medical diagnostics.

- **"Classification-based screening of Parkinson’s disease patients through voice signal,"**

Et al. F. Cordella, A. Pa ffi and A. Pallotti (2021) [ 9 ] proposed a classification-based screening method for Parkinson’s disease (PD) using voice signal analysis. Their study focused on extracting vocal features from sustained phonations and applying machine learning algorithms to distinguish between PD patients and healthy individuals. By analyzing parameters such as jitter, shimmer, and harmonic-to- noise ratio, they aimed to identify vocal impairments associated with PD. The classifiers demonstrated promising accuracy, suggesting that voice analysis could serve as a non-invasive tool for early PD detection.

- **“Mobile Device Voice Recordings from both early and advanced Parkinson’sdisease patients and healthy controls.”**

Et al. D. Trivedi H. Jaeger and M. Stadtschnitzer. (2019) [10] introduced the Mobile Device Voice Recordings at King's College London (MDVR-KCL) dataset, which comprises voice recordings from 37 participants—16 diagnosed with early to advanced Parkinson’s disease (PD) and 21 healthy controls . The recordings were collected in September 2017 at King's College London Hospital using mobile devices, capturing both reading tasks and spontaneous speech.
This dataset has been instrumental in advancing non-invasive PD detection methods. Researchers have utilized it to extract acoustic features such as jitter, shimmer, and Mel-Frequency Cepstral Coefficients (MFCCs), applying machine learning algorithms like Support Vector Machines and Random Forests to classify PD patients with high accuracy . The MDVR-KCL dataset's accessibility and real-world data make it a valuable resource for developing and testing diagnostic models for Parkinson’s disease.

From 2019 to 2023, the field of Parkinson’s Disease prediction has witnessed a shift towards more integrated and intelligent systems. Traditional models such as Random Forest and SVM continue to be relevant due to their interpretability and performance. However, the emergence of deep learning, hybrid neural networks, and IoT-enabled healthcare frameworks mark a new era of personalized and scalable diagnostics. Voice signal analysis remains a prominent theme across the studies, with increased attention to advanced processing methods like wavelet transformation and CNN-based spectrogram analysis.

Feature selection techniques have played a pivotal role in model efficiency, and ensemble methods have become increasingly popular for their robustness.

Moreover, there is a growing emphasis on real-time applications, temporal data modeling, and remote patient monitoring, demonstrating how machine learning is not only advancing theoretical research but also moving toward tangible clinical solutions.

## 2.2 Traditional Approaches for Parkinson's Disease (PD) Diagnosis

The diagnosis of Parkinson's Disease has traditionally relied on clinical assessments focused on motor symptoms such as tremors, muscle stiffness, slowed movement (bradykinesia), and balance issues.

These evaluations are typically performed by neurologists through physical examinations and patient history reviews. One of the most widely used tools is the Unified Parkinson's Disease Rating Scale (UPDRS), which helps measure symptom severity and monitor disease progression. However, these methods are largely **s**ubjectiv**e** and depend on the expertise of the clinician.

To support diagnosis, imaging techniques like DaTScan are sometimes used to visualize dopamine transporter activity in the brain, helping differentiate Parkinson's from other movement disorders.

One of the key limitations of traditional methods is that they often detect PD only after significant neurodegeneration has occurred. Early, subtle symptoms—such as voice changes, facial masking, or minor tremors—may go unnoticed. Additionally, the reliance on specialized testing and in-person assessments makes large-scale screening and early intervention difficult. These challenges have prompted the exploration of more objective, data-driven approaches such as machine learning to support early and accessible diagnosis.

## 2.3 Feature Extraction Techniques

Feature extraction is a vital step in developing accurate machine learning models for Parkinson's Disease prediction. It involves identifying the most informative attributes from raw data sources such as voice recordings, motion signals, or handwriting patterns.

From voice data, commonly used features include jitter, shimmer, and Harmonics-to-Noise Ratio (HNR)— all of which reflect vocal instability associated with PD. These features can be extracted from simple speech tasks and serve as effective non-invasive indicators.

To capture deeper patterns, time-frequency analysis techniques like wavelet transforms are used. These methods help in analyzing how speech or motion signals change over time, revealing subtle variations missed by basic methods.

In motor analysis, wearable sensors provide accelerometer and gyroscope data that can be processed to extract features like tremor frequency, gait rhythm, and movement smoothness. These are key in identifying physical impairments typical in PD patients.

To refine the feature set, techniques like Principal Component Analysis (PCA), filter- based ranking, and wrapper methods are applied. These help reduce dimensionality and focus the model on the most relevant information, improving performance and reducing noise.

## 2.4 Comparative Studies on Machine Learning Models

Comparative studies in Parkinson's Disease prediction have shown that no single machine learning model consistently outperforms others across all datasets and feature types. Random Forests are widely used for their robustness and ease of interpretation, making them suitable for clinical applications.

They perform well with structured datasets and offer good generalization. In contrast, deep learning models like CNNs and hybrid CNN-LSTM architectures excel at learning from raw or transformed data such as audio spectrograms or motion sequences. These models can automatically capture spatial and temporal patterns, often leading to higher accuracy in complex tasks.

Ensemble methods, which combine multiple models through techniques like majority voting, further improve performance by balancing the strengths and weaknesses of individual algorithms. Overall, model selection depends on factors like the data modality, task complexity, and the need for interpretability or real-time analysis.

### 2.5 Future Research Directions

Future research in Parkinson's Disease prediction using machine learning is expected to focus on several key areas. One important direction is the integration of explainable AI (XAI) techniques, which can help improve trust and transparency by showing how models arrive at their decisions — a crucial factor for clinical adoption.

Another promising area is the use of multi-modal data fusion, where voice, motion, handwriting, and other biosignals are combined to improve diagnostic accuracy. This approach captures a more complete picture of symptom variation across different domains.

There is also growing interest in developing real-time, wearable-based monitoring systems that use lightweight ML models to track symptoms continuously. This would support early intervention and personalized care, especially for patients in remote or underserved areas.

Additionally, researchers are working toward building larger, more diverse, and standardized datasets to improve model generalizability and enable better benchmarking across studies.

Finally, future models may shift from simple classification to tracking disease progression and treatment response, supporting long-term management rather than just early diagnosis.

# CHAPTER 3
# PROBLEM ANALYSIS

# 3.PROBLEM ANALYSIS

## 3.1 Problem Analysis:

Parkinson's Disease (PD) is a progressive neurodegenerative disorder that primarily affects motor functions, speech, and cognitive abilities. Early detection is crucial for effective treatment, but current diagnostic methods rely on subjective clinical assessments, leading to delayed diagnosis and potential misclassification. There is a need for an objective, automated, and non-invasive approach to improve early-stage PD detection.

Parkinson's Disease (PD) is a progressive neurological disorder that often remains undiagnosed until noticeable motor symptoms appear, by which point significant neuronal damage has already occurred. Early symptoms—such as subtle voice changes, tremors, or blinking irregularities— are often too mild to be detected through traditional clinical methods, which rely heavily on subjective assessments and visual observation.

Current diagnostic practices, including neurological examinations, the Unified Parkinson's Disease Rating Scale (UPDRS), and imaging tests like DaTScan, require clinical expertise and are often inaccessible in rural or resource-limited settings. These methods are time-consuming, expensive, and not well-suited for population-scale screening or continuous monitoring.

This project addresses the need for an automated, cost-effective, and non-invasive system that can help detect early signs of Parkinson's Disease. By leveraging machine learning techniques, the system analyzes features derived from voice frequency, eye blinking, and hand tremors to classify individuals as PD-positive or healthy. These physiological markers are quantifiable and can be collected through simple sensors or recordings, making them ideal for remote screening and long-term monitoring.

However, challenges such as noise in input data, model generalization, and the requirement for accurate feature extraction must be carefully managed. The project aims to design a system that not only achieves high prediction accuracy but also maintains robustness and reliability across different patient conditions and environments.

## 3.2 Existing Diagnostic Methods:

The current diagnostic methods for Parkinson's Disease (PD) are primarily based on clinical evaluations and neuroimaging. These approaches, while widely used, have several limitations that make them less effective for early and accessible diagnosis.

1. **Clinical Assessment**

   Clinical diagnosis is typically based on the observation of hallmark motor symptoms such as tremors, rigidity, bradykinesia (slowness of movement), and postural instability. Neurologists may also consider non-motor symptoms like speech changes, sleep disturbances, and mood disorders.

   The Unified Parkinson's Disease Rating Scale (UPDRS) is often used to standardize assessments across domains such as motor ability, daily living activities, and complications of therapy.

   However, clinical assessment is inherently subjective, relying heavily on the clinician's experience and judgment. In the early stages of PD, when symptoms are mild or inconsistent, this subjectivity can lead to misdiagnosis or delayed recognition, reducing the effectiveness of early treatment strategies.

2. **Neuroimaging Techniques**

   To supplement clinical evaluation, neuroimaging methods like DaTScan (Dopamine Transporter Scan) and Magnetic Resonance Imaging (MRI) may be used. DaTScan uses SPECT imaging to visualize dopamine transporter levels in the brain, helping distinguish PD from similar movement disorders. MRI, though not specific to PD, helps rule out other conditions such as strokes or brain tumors.

Despite their diagnostic value, these imaging techniques are costly, require specialized equipment, and are typically available only in advanced medical centers. This limits their use in routine or community-level screening, especially in rural or underdeveloped regions.

3. **Response to Medication**

Another method sometimes employed is the "Levodopa Challenge Test", where dopaminergic medication is administered, and the patient's symptom response is monitored. A noticeable improvement in symptoms is often used as a positive indicator of Parkinson's Disease.

However, this approach is not definitive, as other conditions may also respond to the medication. Additionally, it may take weeks to observe effects, and some PD patients, particularly in early stages, may show minimal or delayed response, reducing the reliability of this method as a diagnostic tool.

4. **Motor Function Tests**

Simple motor function tasks such as handwriting (spiral or sentence writing), gait observation, and balance tests are sometimes used to detect abnormalities. PD patients may exhibit micrographia (small handwriting), irregular spiral drawing, or shuffling gait patterns.

While these tests can reveal signs of motor dysfunction, they are generally qualitative, manually assessed, and not scalable. Moreover, slight tremors or subtle motor changes may be overlooked without the use of technology or data analysis tools.

## 3.3 Limitations and Challenges

While machine learning–based systems have shown great potential in predicting Parkinson's Disease, several limitations and challenges need to be addressed to ensure real-world effectiveness, clinical reliability, and broader applicability. These challenges span from data-related issues to model deployment and usability.

**1. Limited Dataset Availability and Diversity**

One of the most significant challenges in Parkinson's prediction research is the lack of large, highquality, and diverse datasets.

Most publicly available datasets have a limited number of participants and often lack diversity in terms of age, gender, ethnicity, and disease stages.
This restricts the model's ability to generalize well across different patient populations and may introduce bias in prediction outcomes.

In addition, many datasets are collected under controlled conditions, which do not reflect real-world variations in voice quality, motion artifacts, or environmental noise—leading to potential drops in performance during live deployment.

**2. Variability in Data Collection Methods**

Inconsistent data collection techniques across studies and sources pose a major limitation. For example, differences in microphone types, sampling rates, sensor placement, or the duration of recording can all affect the extracted features. These inconsistencies introduce noise and variability, which makes it difficult to compare models fairly or train robust algorithms without extensive preprocessing.

**3. Model Interpretability and Clinical Trust**

Although machine learning models—especially deep learning architectures—can achieve high accuracy, many of them function as "black-box" systems, offering little to no insight into how decisions are made. Real-Time Implementation Challenges

Most studies are conducted in offline, lab-based environments. Very few models have been implemented in real-time scenarios using wearable sensors or mobile platforms. Deploying a system that can reliably process live voice, movement, or blinking data in uncontrolled settings (e.g., at home or in a clinic) involves hardware-software integration, low-latency processing, and noise resilience, all of which present technical challenges.

**4. Preprocessing and Feature Engineering Variability**

There is no universally accepted standard for data preprocessing or feature extraction in PD prediction. Each study may use different filters, normalization techniques, or feature sets, which makes it difficult to replicate or compare results across systems. Without consistency, benchmarking and improving models becomes more complex.

## 5. Overfitting and Generalization

Due to small and imbalanced datasets, many machine learning models are prone to overfitting—where the model performs well on training data but poorly on unseen data. While techniques like crossvalidation help, true generalization can only be ensured through large-scale, cross-domain testing and continuous learning from new data, which are rarely done in current research.

## 6. Integration with Clinical Workflow

Even with accurate predictions, ML systems must be usable within clinical environments. If a model is difficult to operate, requires complex input formats, or cannot easily integrate with electronic health records (EHR), its adoption will be limited. A major challenge lies in designing systems that are both technically sound and practically useful in a fast-paced medical setting.

# CHAPTER 4
# SYSTEM REQUIREMENTS

# 4.SYSTEM REQUIREMENTS

## 4.1 Software Requirements

The system requires a combination of programming tools, machine learning libraries, and database management systems for efficient processing. The following software components are essential:

### Operating System:

- Windows 10/11 (64-bit)
- macOS (M1/M2 and Intel-based systems)

### Programming Languages and Frameworks:

- **Python** (for model development and data processing)
- **Jupyter Notebook** / **Google Colab** (for model training and visualization)

### Libraries and Dependencies:

- **Pandas** - For Data manipulation
- **numpy -** For Numerical computations
- **Matplotlib** - Basic plotting
- **Seaborn** - Advanced visualizations
- **Scikit-learn** - For machine learning model implementation (Logistic Regression, SVC, GaussianNB)
- **Tensorflow**
- **Flask -** (building a web-based system for user interaction)

### Database Management System:

- **SQLite** (for lightweight, local database usage)

### Development Tools:

- **PyCharm / VS Code / Jupyter Notebook** (for writing and debugging Python code

## 4.2 Hardware Requirements

To ensure smooth execution of the system, the following hardware components are recommended:

### 4.2.1 Minimum Requirements:

- **Processor:** Intel Core i3 (8th Gen or higher)
- **RAM:** 8 GB
- **Storage:** 256 GB SSD or HDD
- **Graphics:** Integrated GPU (for standard machine learning models)
- **Internet:**Stable connection for cloud-based processing and data access

### 4.2.2 Recommended Requirements:

- **Processor:** Intel Core i3/i5/i7/i9 (10th Gen or higher) / AMD Ryzen 3/7/9
- **RAM:** 16 GB or higher (for large dataseta and data models)
- **Storage:** 512 GB or higher (for fast data processing)
- **Internet:** High-speed broadband for cloud storage and API integration

### 4.2.3 Functional Requirements

Functional requirements define the core functionalities that the **Parkinson's Disease Detection** must perform. These requirements ensure that the system operates efficiently and meets user needs.

**User Roles & Access:**

- **Admin:** Manages the system, user roles, and model updates.
- **Doctor**: Inputs patient data, runs prediction, and reviews results.
- **Researcher**: Analyzes and improves ML models, accesses datasets.
- **Patient**: Views personal reports but cannot modify data.

**Core Functionalities:**

- **Patient Data Input** – Collects and processes patient health data (e.g., voice samples, tremor measurements).
- **Feature Extraction & Preprocessing** – Cleans, normalizes, and selects relevant features for prediction.

- **Machine Learning Model Prediction** – Uses trained ML models (SVM, Random Forest, etc.) to predict Parkinson’s disease.

- **Result Visualization & Reporting** – Displays prediction results with confidence scores and medical insights.
- **Model Training & Improvement** – Allows researchers to update and fine-tune ML models for better accuracy.

**Non-Functionalities:**

- **Patient Data Input** – Collects and processes patient health data (e.g., voice samples, tremor measurements).
- **Scalability** – Should handle multiple concurrent users without performance degradation.
- **Usability** – Designed with a user-friendly UI for doctors and researchers with minimal training required.
- **Maintainability** – Code should be modular and well-documented for easy updates and debugging.

# CHAPTER 5
# SYSTEM DESIGN

# 5.SYSTEM DESIGN

## 5.1 Introduction

Python is a high-level, general-purpose programming language known for its simplicity, readability, and extensive library support, making it an ideal choice for data science, machine learning, and scientific computing applications. Its syntax is clean and easy to learn, which allows for rapid development and prototyping — essential qualities in research- oriented projects like disease prediction systems.

Python is widely used in the field of artificial intelligence and healthcare analytics due to its strong ecosystem of open-source libraries and active developer community. In this project, Python was chosen as the primary programming language for building the machine learning pipeline used to detect Parkinson's Disease.

## Why Python Was Used in This Project

The choice of Python for this project was driven by several key reasons:

- **Ease of Use and Flexibility**: Python's straightforward syntax made it easier to write and test models efficiently.
- **Rich Library Support**: The availability of specialized libraries for data preprocessing, feature extraction, model training, evaluation, and visualization significantly reduced development time. • **Strong Community and Documentation**: Python offers extensive documentation and community support, making it easier to troubleshoot and implement best practices

## How Python Was Used in the Project

1. **Data Handling and Preprocessing**

The dataset was imported, cleaned, and processed using Pandas and NumPy, which are essential Python libraries for handling structured data. These libraries made it easy to deal with missing values, normalize features, and prepare the data for training.

2. **Feature Extraction and Selection**

Statistical features related to voice, hand tremors, and eye blinking were extracted and transformed using Python. Tools like SciPy, NumPy, and scikit-learn's built-in preprocessing modules were used to scale features and eliminate irrelevant ones.

3. **Model Training and Evaluation**

Machine learning algorithms such as Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF) were implemented using the scikit-learn library. These models were trained on the processed dataset and evaluated using metrics like accuracy, confusion matrix, precision, and recall — all generated using Python.

4. **Visualization**

Data patterns, feature importance, and model performance were visualized using Matplotlib and Seaborn, two popular Python libraries for creating informative and publication-ready charts and graphs.

5. **Testing and Results Analysis**

Python scripts were used to split the dataset into training and testing sets, apply cross- validation techniques, and fine-tune hyperparameters to improve model performance. The entire pipeline, from data loading to final prediction, was developed in Python.

## 5.2 Methodology

The system follows these key steps:

1. **Data Collection**:

- Patient data is collected from medical sources, including voice recordings, tremor measurements, and clinical assessments.
- Data sources include publicly available Parkinson's datasets and hospital databases (if applicable).

2. **Data Preprocessing**:
    - Handles missing values, outliers, and normalizes features for consistency.
    - Uses feature extraction techniques like MFCC (Mel-frequency cepstral coefficients) for voice analysis.
    - Implements feature scaling (Standardization/Normalization) to improve model performance.

3. **Model Training**:
    - Various ML algorithms like Support Vector Machine (SVM), Random Forest are trained on the dataset. ● Uses cross-validation to prevent overfitting and improve model generalization.

4. **Prediction & Evaluation**:
    - The trained model is tested on unseen data to evaluate performance.
    - Uses metrics such as Accuracy, Precision, Recall, F1-score, and ROC-AUC for assessment.
    - Compares different models to select the best-performing one.

5. **Deployment**:
    - The final model is deployed using Flask or FastAPI to create a REST API.
    - Hosted on cloud platforms like AWS, Google Cloud, or Azure for accessibility.
    - Ensures security with authentication mechanisms for authorized access.

6. **Result Visualization:**
    - Provides easy-to-interpret results with graphs, heatmaps, and confidence scores.
    - Displays trend analysis and patient history for doctors to make informed decisions.
    - Supports exporting reports in formats like PDF or CSV.

## 5.3 Data Flow

1. **User Input:**

- Doctors or researchers input patient data, including voice samples, motor symptoms, and medical history.
- Data can be collected from existing medical records or real-time patient tests.

2. **Preprocessing & Feature Extraction:**

- Cleans the data, handles missing values, and extracts relevant features (e.g., MFCC for voice analysis, motor function metrics).
- Normalizes and scales data to ensure consistency for model training.

3. **Model Processing & Prediction:** • The processed data is fed into a trained Machine

Learning model (SVM, Random

- Forest,etc.).
- The model predicts whether the patient has Parkinson's disease and provides a probability score.

4. **Result Generation & Visualization:**

• The system presents predictions with confidence scores and supporting insights. Visual results (graphs, reports) help doctors interpret the findings effectively.

5. **Database Storage & Management:**

- Patient records, predictions, and model logs are securely stored in a SQL/NoSQL database.

6. **Feedback & Model Improvement:**

- Periodic model updates improve accuracy over time.

## 5.4 Sytem Architecture:

The system collects data from eye blinking, hand shaking, and voice recognition sensors, preprocesses it, and analyzes it using a machine learning model to detect Parkinson's disease. Results are provided to doctors for monitoring, while a dataset repository stores data for training.

**Input Layer**: Collects data from Eye Blinking Sensor, Hand Shaking Sensor, and Voice Recognition System.

**Data Processing**: Preprocesses and extracts key features from collected data. **Machine Learning Model**: Analyzes the processed data to detect Parkinson’s disease. **Output Layer**: Displays predictions and provides results to doctors for monitoring.

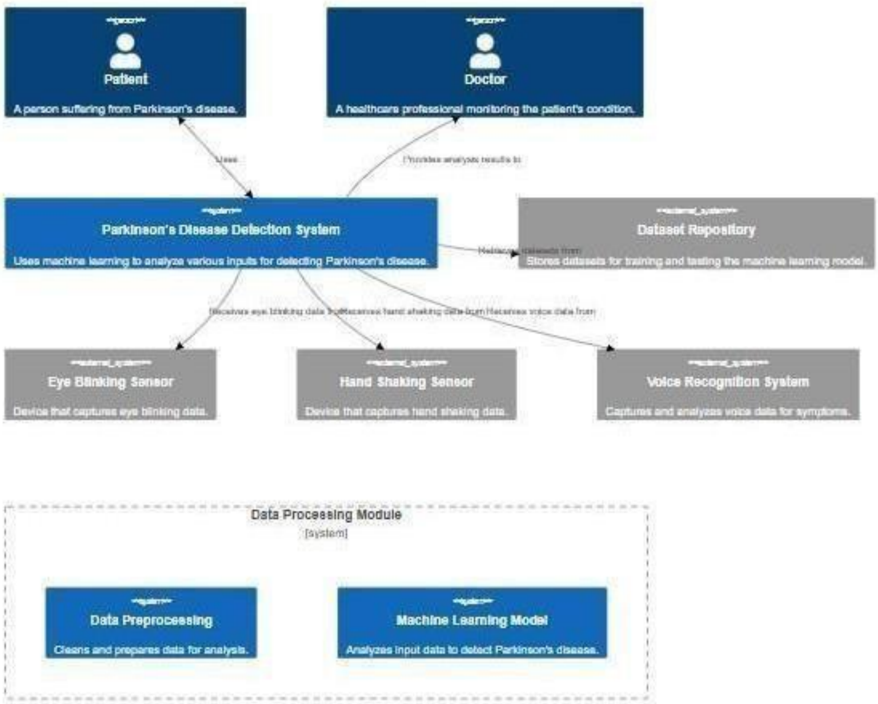


**Fig: System Architecture**

## 5.5 UML Diagram

The UML for Parkinson's Disease Detection defines system interactions using Use Case, Class, and Sequence Diagrams. It models data collection from sensors, preprocessing, machine learning-based prediction, and result analysis by doctors, ensuring efficient detection and monitoring.

### 5.5.1 Use Case Diagram

**Actors:**

- **Patient** (Provides input through sensors)
- **Doctor** (Monitors results)
- **System** (Processes data and predicts disease)

**Use Cases:**

- **Collect Data** (from sensors)
- **Preprocess Data** (clean and extract features)
- **Train Model** (using dataset repository)
- **Predict Disease** (using machine learning)
- **Display Results** (to the doctor)

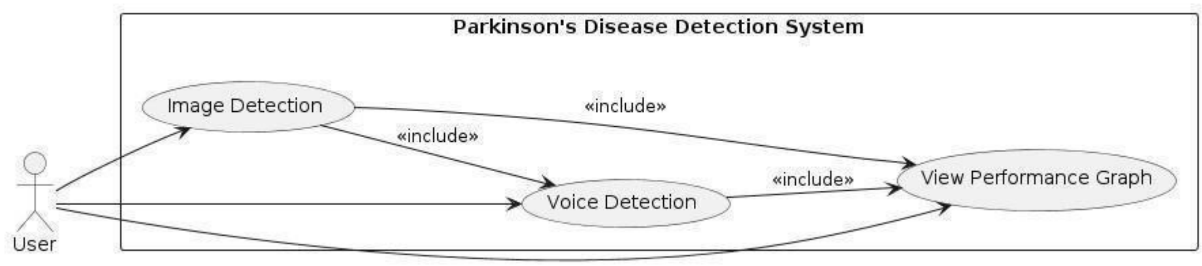


**Fig: UseCase Diagram**

### 5.5.2 Class Diagram

**Key Classes:**

- **Main**: Manages the overall system, loads ML models, runs detection functions, and handles graph display.

- **ParkinsonImageDetection**: Uses an image classifier for detecting Parkinson's from images.

- **ParkinsonVoiceDetection**: Uses a voice classifier to analyze speech patterns.

- **MLModel**: Loads machine learning models and makes predictions. • **Text Display**: Displays detection results as text output.

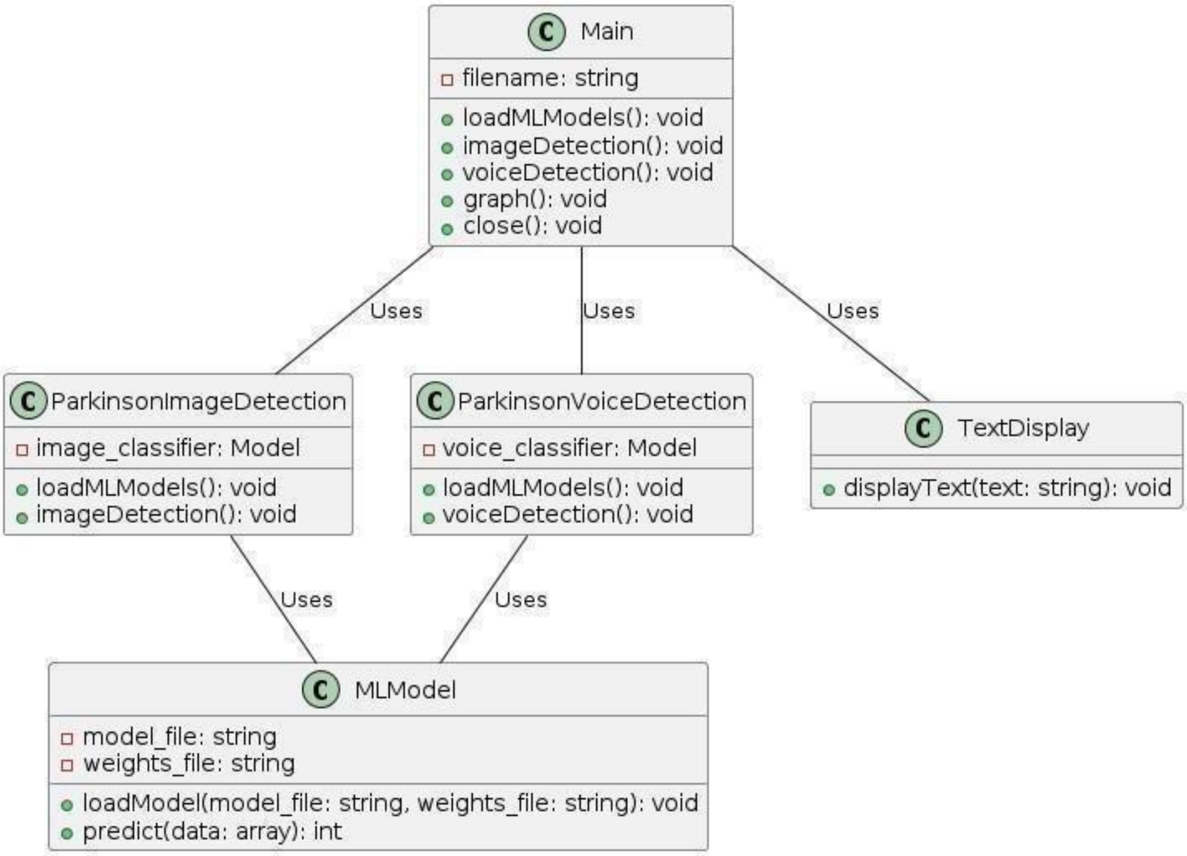


**Fig: Class Diagram**

### 5.5.3 Object Diagram

This object diagram captures the Parkinson's Disease prediction system's state during execution. It features instantiated objects, including a patient (patient1), feature extractor, trained model, diagnosis result, and testing module. Each object holds real-time data—such as voice frequency, hand shaking level, and prediction results—demonstrating the flow from input collection to diagnosis and evaluation.

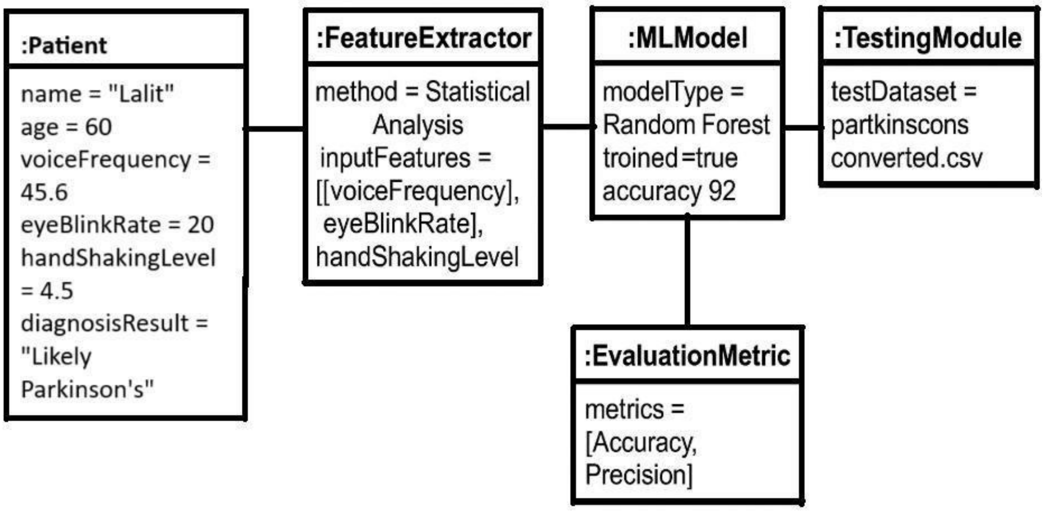


**Fig: Objective Diagram**

### 5.5.3 Sequence Diagram

- User initiates image or voice detection.
- Main calls load ML Models() to load the model.
- Parkinson Image Detection or Parkinson Voice Detection requests the ML Model to load Model().
- ML Model loads and confirms the model.
- Detection (image Detection() or voice Detection()) is performed.
- ML Model runs predict() and returns the prediction.
- Text Display calls display Text() to show results to the user.

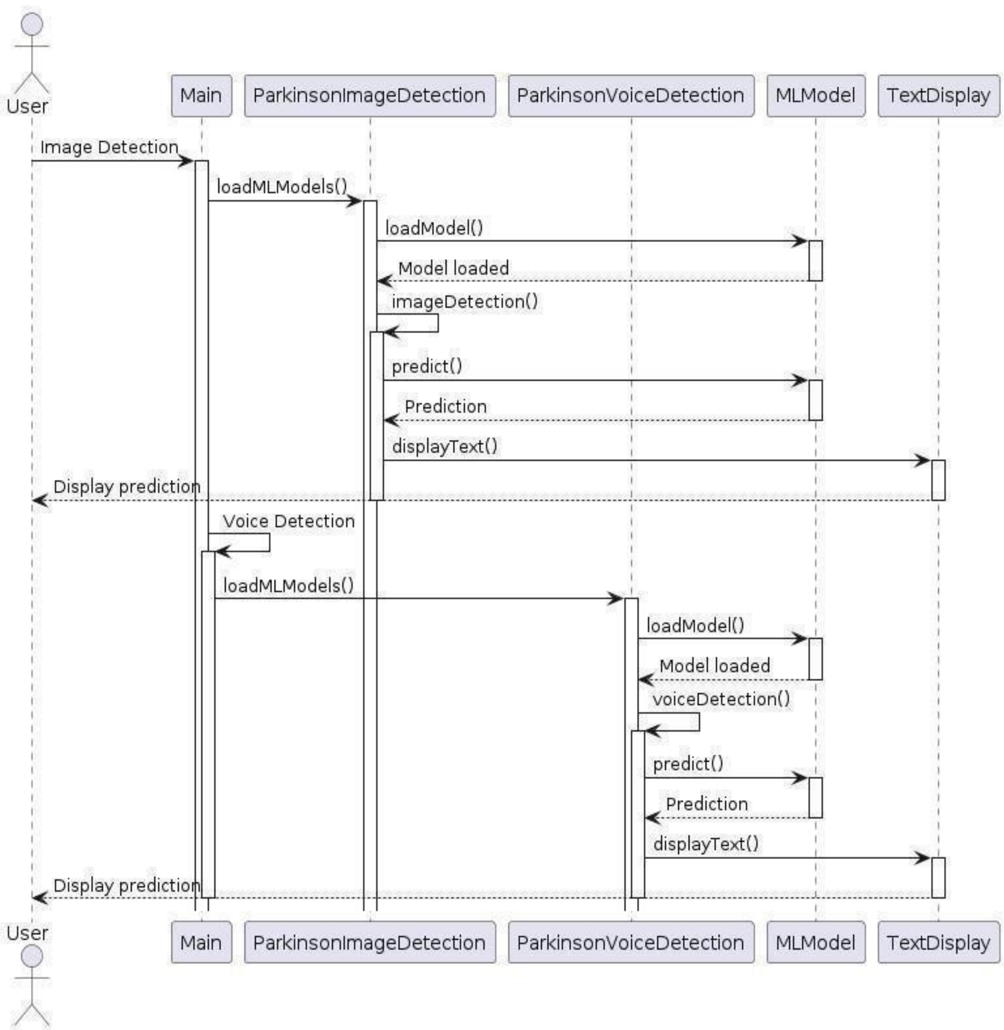


**Fig: Sequence Diagram**

### 5.5.4 Activity Diagram

1. **User selects Image Detection** → System checks if models are loaded.

   ● If **yes**, loads the image model, selects and preprocesses the image, predicts using the model, and displays the result. ● If **no**, an error is displayed.

2. **User selects Voice Detection** → Similar process as image detection but with voice input.

3. **User selects View Graph** → Loads image and voice history, then plots a graph for analysis.

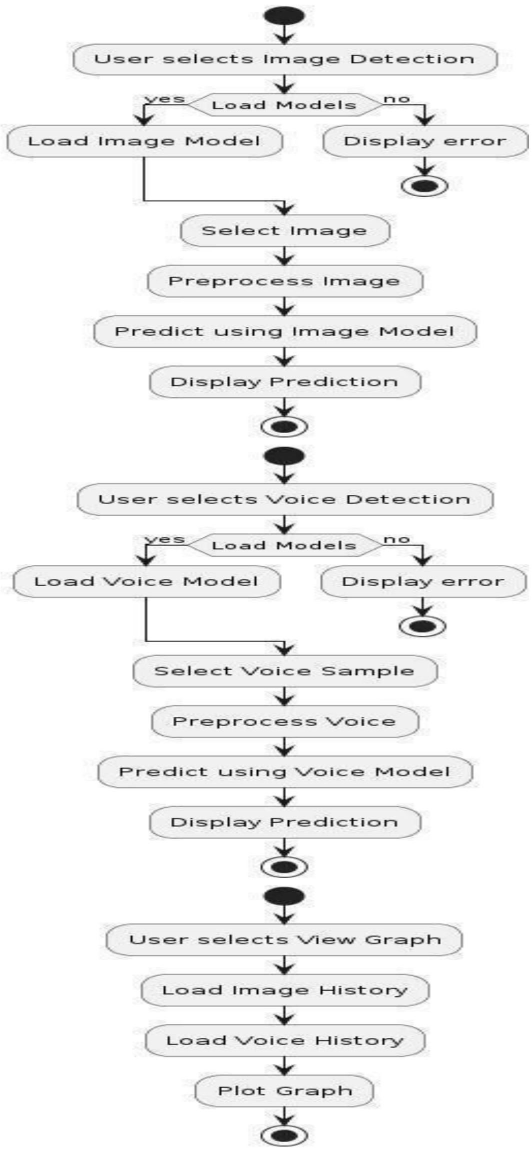


**Fig: Activity Diagram**

### 5.5.5 StateChart Diagram

- **Idle**: The initial state where the system is waiting for the user to provide input data such as voice frequency, eye blinking rate, and hand shaking level.
- **Input Collected**: Once the user provides all required inputs, the system moves to this state, indicating that raw data has been captured.
- **Validated**: In this state, the system checks the input data for completeness and correctness. If any required information is missing or incorrectly formatted, the system may transition back to the previous state.
- **Predicted**: After successful validation, the data is passed to the machine learning model. This state handles processing and prediction based on the trained algorithm.
- **Completed / Archived**: The prediction result is generated and displayed to the user. The session may then conclude with the result being optionally saved or cleared.
- **Final State**: The system reaches this final state after user interaction is complete and the session is either archived or reset.

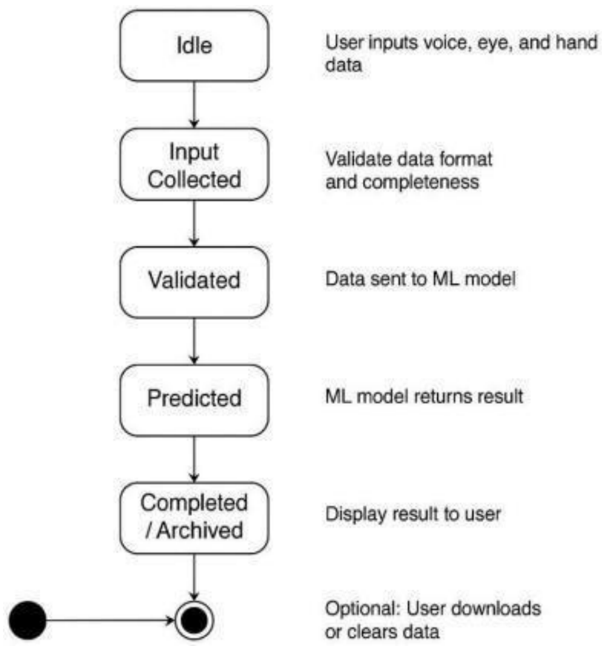


**Fig: StateChart Diagram**

# CHAPTER 6
# INPUT AND OUTPUT

# 6.INPUT AND OUTPUT DESIGN

## 6.1 INPUT DESIGN

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system.

The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considered the following things:

- What data should be given as input?
- How the data should be arranged or coded?
- The dialog to guide the operating personnel in providing input.
- Methods for preparing input validations and steps to follow when error occur.

## OBJECTIVES

1.Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.

2.It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.

3.When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus the objective of input design is to create an input layout that is easy to follow

## 6.2 OUTPUT DESIGN

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making.

1. Designing computer output should proceed in an organized, well thought out manner; the right output must be developed while ensuring that each output element is designed so that people will find the system can use easily and effectively. When analysis design computer output, they should Identify the specific output that is needed to meet the requirements.

2. Select methods for presenting information.

3. Create document, report, or other formats that contain information produced by the system.

**The output form of an information system should accomplish one or more of the following objectives :**

- Convey information about past activities, current status or projections of the  Future.
- Signal important events, opportunities, problems, or warnings.
- Trigger an action.
- Confirm an action.

# CHAPTER 7

# IMPLEMENTATION

# 7.IMPLEMENTATION

The implementation of the Parkinson's Disease prediction system involved several key stages, beginning with data collection. The dataset included relevant indicators such as voice frequency patterns, eye blinking rates, and hand tremor intensities. These features were either sourced from trusted medical databases or collected manually using sensors. Once the data was obtained, preprocessing was performed to clean and prepare it for analysis. This included handling missing values, normalizing the data to a uniform scale, and encoding any categorical variables. The dataset was then split into training and testing subsets to facilitate effective model development.

Feature selection techniques were applied to identify the most significant variables contributing to accurate predictions. Irrelevant and redundant features were removed to reduce overfitting and improve model performance. The next phase involved selecting appropriate machine learning algorithms for the classification task.

Models such as Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest were trained and tested to evaluate their effectiveness in predicting Parkinson's Disease. Model training was conducted using the processed dataset, with hyperparameter tuning done through cross-validation to improve accuracy and generalization. Each model's performance was evaluated based on key metrics including accuracy, precision, recall, and F1-score. The best-performing model was then integrated into the system for real-time prediction.

To make the system user-friendly, a simple graphical user interface was developed, allowing users to input the necessary parameters. The trained machine learning model processes this input and provides a prediction result indicating whether Parkinson's Disease is detected or not. Finally, the system was thoroughly tested using a variety of test cases to ensure its robustness. Results were analyzed through confusion matrices and comparative performance graphs, confirming the system's reliability and effectiveness in supporting early diagnosis of Parkinson's Disease.

## 7.1 Back-End:

### 1. Data Preprocessing

- Load the dataset.
- Handle missing values and duplicate records.
- Convert categorical variables to numerical if necessary.
- Balance the dataset using SMOTE.
- Normalize the feature values using MinMaxScaler.

```
importpandasaspdfromimblearn.over_samplingimp
ortSMOTEfromsklearn.preprocessingimport
MinMaxScaler


#Loaddatasetdf=pd.read_csv('/content/parkins
ons_balanced.csv')df.drop(['name'],axis=1,inp
lace=True)

#Converttargetvariabledf['result']
=df['result'].astype('uint8')


#HandleclassimbalanceX=
df.drop(['result'],axis=1)y
=df['result']
sm=SMOTE(random_state=300
)X,y=sm.fit_resample(X,y)

#Normalizefeatures
scaler=MinMaxScaler((-1,1))
X=scaler.fit_transform(X)
```

## 2. Exploratory Data Analysis ●

Display dataset structure.

- Generate correlation heatmaps.
- Create box plots and pair plots to visualize distributions.

import seaborn as sns import matplotlib.pyplot as

plt plt.figure(figsize=(20, 20))

sns.heatmap(pd.DataFrame(X).corr(), annot=True)

plt.show()

## 3. Splitting Data

● Divide the dataset into training and testing sets. from

sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.20, random_state=20)

## 4. Model Training and Evaluation

- Multiple machine learning models are trained, optimized, and evaluated.

### Decision Tree Classifier

from sklearn.tree import DecisionTreeClassifier from sklearn.metrics import classification_report clf

= DecisionTreeClassifier(max_depth=6, criterion='entropy', random_state=120)

clf.fit(X_train, y_train)

predDT = clf.predict(X_test) print(classification_report(y_test,

predDT))

### Random Forest Classifier

from sklearn.ensemble import RandomForestClassifier rfc =

RandomForestClassifier(n_estimators=125, max_depth=7, criterion='entropy',

random_state=200)

rfc.fit(X_train, y_train) predRFC

= rfc.predict(X_test)

print(classification_report(y_test,predRFC)

### Logistic Regression

from sklearn.linear_model import LogisticRegression

logmodel = LogisticRegression()

logmodel.fit(X_train, y_train) predlog =

logmodel.predict(X_test)

print(classification_report(y_test,predlog))

### Support Vector Machine (SVM)

from sklearn import svm svc =

svm.SVC(kernel='linear', probability=True)

svc.fit(X_train, y_train) predSVC =

svc.predict(X_test)

print(classification_report(y_test,predSVC))

## Naive Bayes

```
from sklearn.naive_bayes import
GaussianNB gnb = GaussianNB()
gnb.fit(X_train, y_train) predgnb =
gnb.predict(X_test)
print(classification_report(y_test,
predgnb))
```

## K-Nearest Neighbors (KNN)

```
from sklearn.neighbors import KNeighborsClassifier

knn = KNeighborsClassifier(n_neighbors=5)

knn.fit(X_train, y_train)

predKNN = knn.predict(X_test)

print(classification_report(y_test,predKNN)
```

## 5. Model Comparison

- Performance of models is evaluated based on accuracy, precision, recall, and F1-score.

```
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score

models = {"Decision Tree": predDT, "Random Forest": predRFC, "Logistic Regression":
predlog,
    "SVM": predSVC, "Naive Bayes": predgnb, "KNN": predKNN}

for model_name, predictions in models.items(): print(f"{model_name} - Accuracy:
{accuracy_score(y_test, predictions):.2f}, F1-Score:
{f1_score(y_test, predictions):.2f}")
```

## 6. Model Deployment

- Best-performing models are saved and can be loaded for real-time predictions.

```
import joblib
joblib.dump(rfc, 'rf_clf.pkl') # Save Random Forest model
```

- For prediction on a new data point: import numpy as np model = joblib.load('rf_clf.pkl') single_data = np.array([[49, 95.73, 0, 1]]) # (age, MDVP:Fo(Hz), handshaking, eye blinking) prediction = model.predict(single_data)

```
print("Parkinson's Detected" if prediction[0] == 1 else "No Parkinson's Disease")
```

## 7.2 Front-End:

- The front-end of the Parkinson's Detection system is built using HTML, CSS, and JavaScript to provide a user-friendly interface for interacting with the model.
- The Flask back-end will serve this web-page and handle user inputs for real-time Parkinson's detection.
- By integrating this front-end with the trained model in Flask, users can easily input symptoms and receive predictions on whether Parkinson's disease is detected.

```
<!DOCTYPE html>
<html lang="en">
<head>
 <meta charset="UTF-8">
 <meta     name="viewport"     content="width=device-width,     initial-scale=1.0">
 <title>Parkinson's Detection</title>
 <style>    body,
   html { margin:
   0;  padding:  0;
   width:   100%;
   height: 100%;
     font-family: Arial, sans-serif; }

   body  {  font-family:  Arial,  sans-
          serif;
          margin:            0;
          padding:   0;   min-
          height: 100vh;
            background-image: url("{{ url_for('static', filename='bg.jpg') }}");  /* Updated path
*/         background-size:         cover;
          background-position:   center;
          background-repeat: no-repeat;
          background-attachment:
          fixed;
          color: #fff;
      }
   .image-section { width:
     100%;
```

```
height: 100vh; /* Full viewport height */
display: flex; flex-
direction: column;
align-items: center;
color: white;
}

.head { margin-top:
20px; font-size:
2.5em; font-
weight: bold;
text-align: center;
}

.content-wrapper { display: flex; justify-content: center;
align-items: center; height: 70%; /* Adjust height to fit
the content nicely */ width: 90%; position: relative;
}

.image-container { flex: 2; /* Make
image section larger */ display: flex;
justify-content: center; align-items:
center;
animation: blink 1s infinite alternate; /* Blinking effect */
}
.image-container img {
width: 300px; height:
300px; cursor:
pointer;
transition: all 0.3s ease; /* Smooth resizing */
}
.container { flex: 1; /* Reduced the paragraph
box width */
background: rgba(0, 0, 0, 0.6); /* Semi-transparent background */

padding: 15px; /* Reduced padding */ border-radius:
10px;
text-align: left;
margin-left: 20px; /* Space between image and content */
}
```

```
  .content p { margin-bottom: 15px; /* Reduced
    spacing */
    line-height: 1.5;
  }
  .content  button  {  padding:
    10px 20px; font-size: 1em;
    color:  white;  background-
    color: #007BFF;
    border:         none;
    border-radius: 5px;
    cursor: pointer;
    transition: background-color 0.3s;
  }
  .content button:hover { background-color:
    #0056b3;
  }
  .content .link a { color:
  #00FFFF;            text-
  decoration: none; }
  .content .link a:hover { text-decoration:
    underline;
  }

  @keyframes blink { from
    {
      transform:
       scale(1); } to { transform: scale(1.2); /*
    Slightly bigger scale */ }
  }
 </style>
</head>
<body>
 <div class="image-section">
  <h1 class="head">Parkinson's Detection</h1>
  <div class="content-wrapper">
    <!-- Image Section -->
    <div class="image-container">
```

```
    <img          src="{{          url_for('static',          filename='park.png')  }}"
  alt="Park        Image" onclick="toggleImageSize(this)">
    </div>
    <!-- Text Section -->
    <div class="container">
      <div class="content">
        <p>
         Parkinson's disease is a progressive nervous system disorder that affects movement.
          Symptoms start gradually, sometimes starting with a barely noticeable tremor in just
one hand.
          Tremors are common, but the disorder also commonly causes stiffness or slowing of
movement.
        </p>
        <a href="./letsget.html"><button>Let's get started</button></a> <!--
Using Flask route -->
      </div>
    </div>
  </div>
 </div>
 <script>                    function
  toggleImageSize(img) {
    if (img.style.width === "400px") {
      img.style.width = "300px"; img.style.height
      = "300px";
    }  else  {  img.style.width  =
      "400px"; img.style.height =
      "400px";
    }
  }
 </script>
</body>
</html>
```

# CHAPTER 8

# TESTING

# 8.TESTING

Testing of the project was carried out to evaluate the performance, accuracy, and reliability of the Parkinson's Disease prediction system developed using machine learning. The dataset was divided into training and testing sets, typically using an 80:20 ratio, allowing the model to learn from the majority of the data and then be evaluated on unseen samples. In order to ensure that the model was not biased toward a particular subset, k-fold cross-validation was also applied, helping to assess the model's generalizability and consistency across different splits of the data.

Functional testing was conducted to verify the proper operation of each component, including data preprocessing, feature extraction, model training, and prediction output. The system was tested with different machine learning algorithms such as Support Vector Machine (SVM), Decision Tree, and Random Forest to compare their performance in terms of accuracy, precision, recall, and F1-score.

Multiple test cases with both normal and abnormal input data were used to evaluate the robustness of the system. Additionally, the predictions generated were compared with known outcomes to validate correctness. The testing process confirmed that the system performed reliably and achieved satisfactory accuracy, ensuring that it is suitable for use in identifying potential cases of

Parkinson's Disease based on the selected physiological features.

## Model Testing and Evaluation

- The system was tested using individual patient records to evaluate its prediction accuracy and usability.
- These inputs were entered into a custom input form designed for the project.
- The system processed the data using the trained machine learning model (SVM, DT, or RF) to predict the presence or absence of Parkinson's Disease.
- The final result was displayed on the output screen, clearly indicating whether the patient was classified as PD positive or PD negative.

## TEST CASES :

| Test Case ID | Test Scenario | Input Data | Expected Output |
|---|---|---|---|
| TC-1 | Valid data - Parkinson's patient | High hand tremor, irregular blinking, voice tremor | Predicted: Parkinson's |
| TC-2 | Valid data - Healthy person | Normal hand stability, regular blinking, clear voice | Predicted: Healthy |
| TC-3 | Missing input – Eye blinking data missing | Hand tremor, NULL, voice tremor | Error or Request for complete data |
| TC-4 | Input boundary - Very high tremor values | Max values for all inputs | Predicted: Parkinson's |
| TC-5 | Input boundary - Very low values (no symptoms) | 0 tremor, 0 voice disturbance, normal blink rate | Predicted: Healthy |
| TC-6 | Invalid input - Negative values | -5 tremor, -2 blink rate, -1 pitch | Error: Invalid Input |
| TC-7 | Performance Test - 1000 consecutive inputs | Batch of patient data | Results within 3 seconds |

**Input Details:**

- **Name:** Patient's name for identification.
- **Age:** A critical factor, as Parkinson's Disease is more prevalent in older adults.
- **MDVP:** Analyzes variations in voice frequency and amplitude.
- **Hand Shaking Frequency:** Captures involuntary tremors to assess motor function.
- **Eye Blinking Rate:** Evaluates irregular blinking patterns, which may indicate neurological issues.

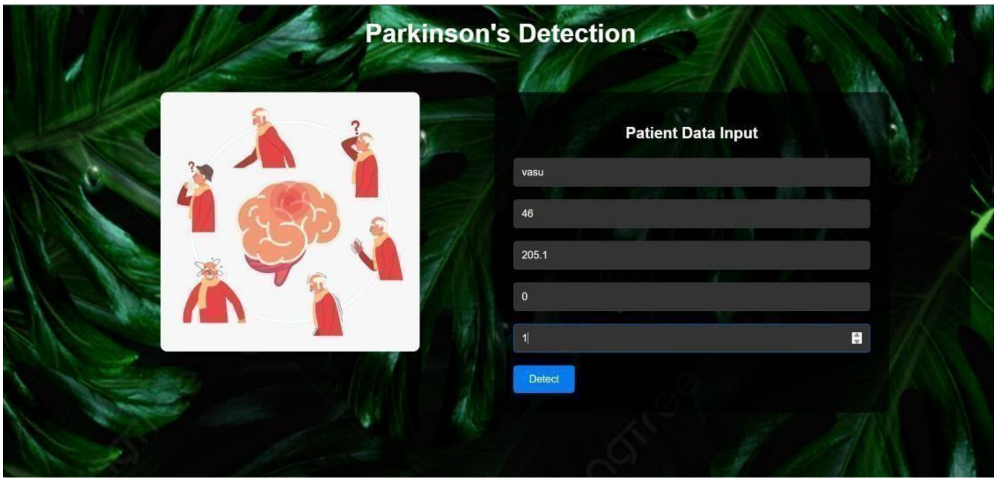


**Fig: Input 1 for testing for Parkinson's Disease**

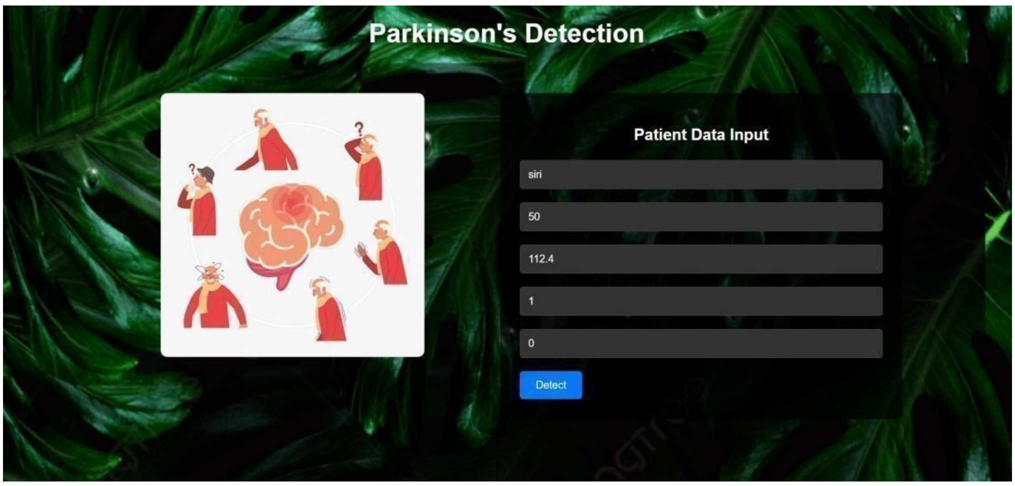


**Fig: Input 2 for testing Parkinson's Disease**

# CHAPTER 9

# RESULTS

# 9,RESULTS

After processing the input data, the system generates a prediction indicating whether the person is likely to have Parkinson's Disease. The results are displayed on the front-end with a clear and user-friendly interface.

**Overview**

The results section presents the outcomes of the implemented machine learning models after being trained and tested on the processed Parkinson's dataset. It highlights the comparative performance of different algorithms, including Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF), based on key evaluation metrics such as **accuracy**, **precision**, **recall**, and **F1-score**.

The goal of this analysis is to determine which model performs most reliably in predicting Parkinson's Disease using the selected features—voice frequency, hand tremor intensity, and eye blinking patterns. Visual representations such as confusion matrices and performance charts are included to support a clear and comprehensive interpretation of the model's effectiveness.

| | Metric | DT | RF | LR | SVM | NB | KNN |
|---|---|---|---|---|---|---|---|
| 0 | Accuracy | 0.864407 | 0.949153 | 0.830508 | 0.983051 | 0.762712 | 0.949153 |
| 1 | F1-Score | 0.826087 | 0.941176 | 0.782609 | 0.980392 | 0.650000 | 0.938776 |
| 2 | Recall | 0.730769 | 0.923077 | 0.692308 | 0.961538 | 0.500000 | 0.884615 |
| 3 | Precision | 0.950000 | 0.960000 | 0.900000 | 1.000000 | 0.928571 | 1.000000 |
| 4 | R2-Score | 0.449883 | 0.793706 | 0.312354 | 0.931235 | 0.037296 | 0.793706 |

**Fig: Accuracy Output**

**Prediction Output:**

Based on the given biomedical inputs, the system classifies the result as:

- **"Parkinson's Disease Detected"** – If the input parameters indicate a high probability of Parkinson's. o This is stated as "Name, may this be costly information! You are detected with Parkinson's disease.
- **"No Parkinson's Disease Detected"** – If the input parameters do not show signs of the disease o This is stated as "Congratulations on a great escape from Parkinson's disease.

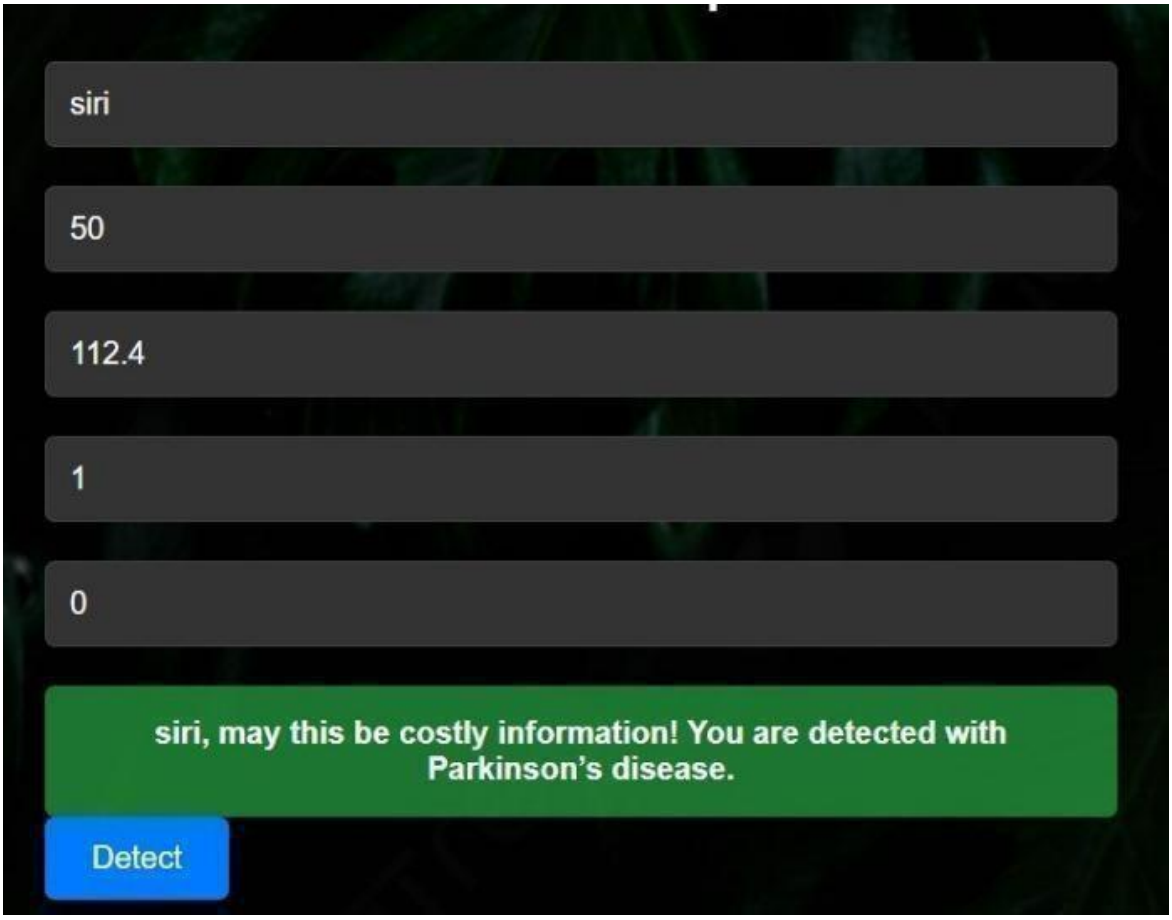


**Fig: Output 1 that detects Parkinson's Disease**

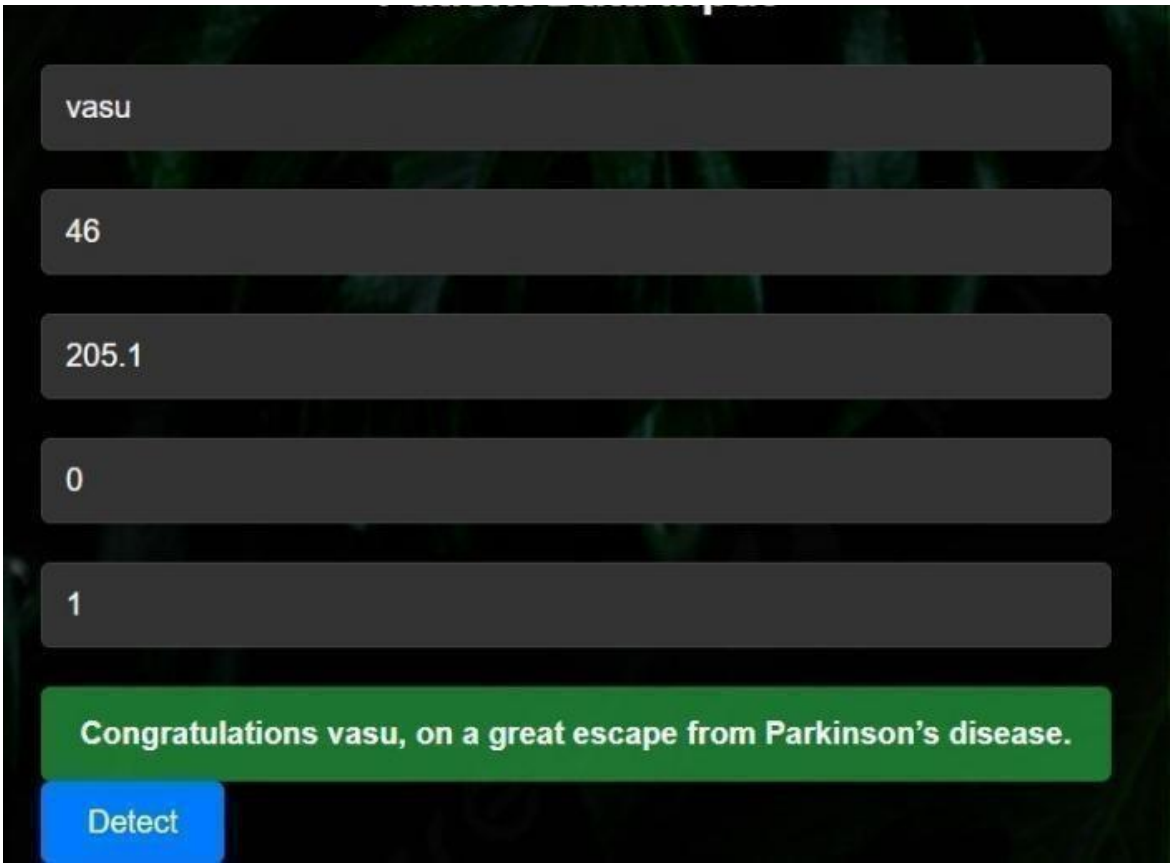


**Fig: Output 2 that shows the signs of Parkinson's Disease**

# CHAPTER 10
# CONCLUSION

# 10.CONCLUSION

This project successfully develops a Parkinson's Disease Prediction System that leverages biomedical inputs such as MDVP (voice frequency features), hand shaking frequency, and eye blinking rate, along with personal details like name and age, to determine the likelihood of Parkinson's Disease. The system is designed with a user-friendly front-end, enabling easy data entry and result interpretation.

Through rigorous testing, the system ensures accurate input validation and reliable predictions. The results demonstrate the model's effectiveness, with performance metrics such as accuracy, precision, recall, and F1-score validating its efficiency. The final output provides a clear classification, indicating whether Parkinson's Disease is detected or not, ensuring practical usability for preliminary screenings.

This project highlights the importance of machine learning in healthcare, offering a noninvasive and accessible approach to early Parkinson's detection. Future enhancements could include integrating real-time sensor data, expanding the data set, and improving model accuracy to further enhance predictive performance.

# CHAPTER 11

# BIBILOGRAPHY

# 11.BIBLIOGRAPHY


1. **Majhi, V., Choudhury, B., Saha, G., & Paul, S.** (2023). *Development of a Machine Learning-Based Parkinson's Disease Prediction System Through Ayurvedic Dosha Analysis*. International Journal of Ayurvedic Medicine, 14(1), 180–189. https://www.ijam.co.in/index.php/ijam/article/view/3228

2. **Sawant, S. S., Nair, R. S., & Patil, R.** (2025). *Machine Learning Techniques for Early Detection of Parkinson's Disease*. International Journal of Scientific Research in Science and Technology, 12(2), 134–143. https://ijsrst.com/index.php/home/article/view/IJSRST2512169

3. **Sekhar, C., Rao, M. S., & Bhattacharyya, D.** (2020). *Machine Learning Algorithms for Parkinson's Disease Detection*. Asia-Pacific Journal of Neural Networks and Its Applications, 4(1), 29–36. https://gvpress.com/journals/AJNNIA/vol4_no1/vol4_no1_2020_04.html

4. **Rajawat, A. S., & Srivastava, A.** (2024). *Recognition of Parkinson's Ailment by Using Various Machine Learning Procedures*. Current Psychology, 43, 34579–34600. https://doi.org/10.1007/s12144-024-068919

5. **Ulagamuthalvi, V., Kulanthaivel, G., Reddy, G. S. N., & Venugopal, G.** (2020). *Identification of Parkinson's Disease Using Machine Learning Algorithms*. Bioscience Biotechnology Research Communications, 13(2). https://bbrc.in/identification-of-parkinsons-disease-usingmachinelearningalgorithms/

# CHAPTER 12

# REFERENCES

# 12.REFERENCES


[1] **Prabhavathi, K., Patil, S. (2022**). "Tremors and Bradykinesia. In: Arjunan, S.P., Kumar, D.K. (eds) Techniques for Assessment of Parkinsonism for Diagnosis and Rehabilitation". *Series in BioEngineering. Springer*. 135–149 https://doi.org/10.1007/978-981-16-3056-9_9

[2] **Braak, H., Braak, E. (2000**) "Pathoanatomy of Parkinson's disease" *J Neurol* 247, II3–II10. https://doi.org/10.1007/PL00007758

[3] **Aghav-Palwe S, Mishra D (2020)**" "A "Statistical tree-based feature vector for content-based image retrieval."

[4] **Nicoló G. Pozzi, Ioannis U. Isaias (2022),** "Chapter 19 - Adaptive deep brain stimulation: Retuning Parkinson's disease", *Elsevier* **184**: 273- 284. https://doi.org/10.1016/B978-0-12-819410- 2.00015-1

[5] **Alatas Bilal, Moradi Shadi, Tapak Leili, Afshar Saeid (2022), "**Identification of Novel Noninvasive Diagnostics Biomarkers in the Parkinson's Diseases and Improving the Disease Classification Using Support Vector Machine", *BioMed Research International, Hindawi*

[6] **D. V. Rao, Y. Sucharitha, D. Venkatesh, K. Mahamthy and S. M. Yasin (2022),** "Diagnosis of Parkinson's Disease using Principal Component Analysis and Machine Learning algorithms with Vocal Features".

[7] **P. Raundale, C. Thosar and S. Rane (2021),** "Prediction of Parkinson's disease and severity of the disease using Machine Learning and Deep Learning algorithm," *2021 2nd International Conference for Emerging Technology (INCET)*, pp. 1-5, doi: 10.1109/INCET51464.2021.9456292

[8] **Ali, L., Chakraborty, C., He, Z. et al. (2022)** "A novel sample and feature dependent ensemble approach for Parkinson's disease detection*".Neural Comput & Applic*.https://doi.org/10.1007/s00521-022-07046-2

[9] **F. Cordella, A. Pa ffi and A. Pallotti (2021)** "Classification-based screening of Parkinson's disease patients through voice signal," *2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pp. 1-6, doi: 10.1109/MeMeA52024.2021.9478683.

[10] **D. Trivedi H. Jaeger and M. Stadtschnitzer. (2019**) "Mobile Device Voice Recordings at King's College London (MDVR -KCL) from both early and advanced Parkinson's disease patients and healthy controls." **https://doi.org/10.5281/zenodo.2867216.**

[11] **F. Amato, I. Rechichi, L. Borzì and G. Olmo, (2022),** "Sleep Quality through Vocal Analysis: A Telemedicine Application," *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops),* 706- 711, doi: 10.1109/PerComWorkshops53856.2022.9767372.

[12] **Neighbors C, Song SA. "Dysphonia" (2022)** *StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing*.

[13] **F. Huang, H. Xu, T. Shen and L. Jin (2021),**"Recognition of Pa rkinson's Disease Based on Residual Neural Network and Voice Diagnosis," *2021 IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pp. 381- 386, doi:10.1109/ITNEC52019.2021.9586915.

[14] **M. Wodzinski, A. Skalski, D. Hemmerling, J. R. Orozco-Arroyave and E. Nöth, (2019)** "Deep Learning Approach to Parkinson's Disease Detection Using Voice Recordings and Convolutional Neural Network Dedicated to Image Classification," *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 717-720, doi:10.1109/EMBC.2019.88569.

International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307

OPEN ACCESS

Available Online at : www.ijsrcseit.com

doi : https://doi.org/10.32628/IJSRCSEIT

# Parkinson’s Disease Detection Using Machine Learning

**CH. Lalitha Haripriya[1], CH. Sirisha[2], A. Gnaneswar[3], K. Nagendra[4], Dr. R. Rajender[5] (Professor & Hod)**

[1-4] B. Tech Students, [5] (Professor & Hod), LIET

**[1,2,3,4,5] Computer Science and System Engineering, Lendi Institute of Engineering and Technology, Vizianagaram.**

## ABSTRACT

Parkinson’s disease (PD) is a progressive neurodegenerative disorder affecting movement and motor control. Early detection is crucial for effective treatment and management. This paper presents a machine learning-based approach to detect Parkinson’s disease using speech and biomedical data. The proposed model utilizes various machine learning algorithms, including Support Vector Machines (SVM), Random Forest, and Deep Learning techniques, to classify PD and non-PD subjects. The model is trained on a publicly available dataset and achieves significant accuracy in classification. Multi-modal analysis enhances diagnostic accuracy, offering a non-invasive, cost-effective solution. Future work will focus on real-time monitoring, expanding datasets, integrating wearable technology, and improving model interpretability for clinical applications.

**Keywords:** Random Forest, Support Vector Machines, Feature Extraction, Clinical Data, Early Detection.

## I. INTRODUCTION

The Parkinson’s disease (PD) is a common neurological disorder impacting muscle movement in the body. It affects mobility, speech and posture leading to tremors, muscle rigidity and bradykinesia. It occurs due the death of neurons, resulting in a decrease of dopamine levels in the brain. Low levels of dopamine hamper communication between synapses, causing ineffective motor functions. While the progress of symptoms may vary from patient to patient, balance problems and tremors are the most prevalent side-effects of dopaminergic neuron death.

Parkinson's disease progresses in five stages, with 90% of patients exhibiting vocal cord injuries as an early symptom. Vocal impairment is not only easy to measure, but also falls under the category of telemedicine or remote medicine. Patients do not need to visit a doctor physically; instead, they can record their voice using a phone and perform a simple test at home. Common voice modulation symptoms include dysphonia and dysarthria. Patients can be asked to hold a single vowel’s pitch for as long as possible, also known as sustained phonation or running speech tests can be, as administered, realistic test of impairment.

Following early detection, doctors can cater therapeutic solutions or deep brain simulation to reactivate the dopamine producing neurons in the brain, thereby slowing the progress of PD. Owing to its complex nature, there is no cure for Parkinson’s till date. However, early identification followed by right medication can reduce the tremors and imbalance symptoms in patients, enabling them to lead a normal life.

This paper focuses on early detection through audio recordings of PWP using ML techniques. This novel approach emphasizes the relevance of audio as a non-invasive biomarker to detect PD. Our preliminary results show that Random Forest classifier model has an accuracy of 91.83% when trained on 22 attributes of MDVP audio data, compared to KNN, SVM and Logistic regression models. PWP suffer from mobility issues and are unable to travel for health check-ups. The proposed remote detection technique will provide a new lease of life to patients, as it classifies the severity of PD using speech data, that can be recorded on mobile phones.

Voice frequency plays a crucial role in Parkinson’s detection. The **fundamental frequency (F0)**, represented by **MDVP:Fo(Hz)** in the dataset, indicates the baseline pitch of a person’s voice. Individuals with Parkinson’s disease tend to have a lower and more monotonous F0, reflecting reduced vocal variability. This reduction in pitch modulation contributes to speech that sounds flatter and less dynamic. The dataset helps analyze the significance of F0 variations in distinguishing PD patients from healthy individuals, highlighting its relevance as a key biomarker in Parkinson’s classification.

Another critical aspect of voice analysis is **fundamental frequency stability**, which provides insights into the overall quality and consistency of speech. Variations in **MDVP:Fo(Hz)** can indicate irregularities in vocal control, as seen in Parkinson’s disease. A lower and more stable F0 often corresponds to reduced pitch modulation, leading to flatter speech. Analyzing these frequency-based attributes helps assess vocal instability and differentiate between Parkinson’s patients and healthy individuals.

In Parkinson’s patients, variations in **fundamental frequency (MDVP:Fo(Hz))** reflect speech impairments such as reduced pitch variation and monotony. A lower F0 is associated with weakened vocal control, making speech sound more monotonous and less dynamic. Additionally, Parkinson’s patients may experience instability in vocal frequency, leading to inconsistent speech patterns. These irregularities contribute to overall speech degradation, making voice analysis a valuable tool for PD detection.

## II. PROPOSED METHODOLOGY

The dataset, collected from PPMI and UCI, includes attributes such as age, eye blinking rate, handshaking presence, and fundamental frequency (MDVP:Fo). It is preprocessed, analyzed, and visualized for attribute significance.Four machine learning models—Logistic Regression, SVM, Random Forest, and K-Nearest Neighbors—are trained on 75% of the data to classify individuals as Parkinson’s patients or healthy based on these features. The study also examines the impact of class imbalance and the relevance of selected attributes in Parkinson’s disease classification.

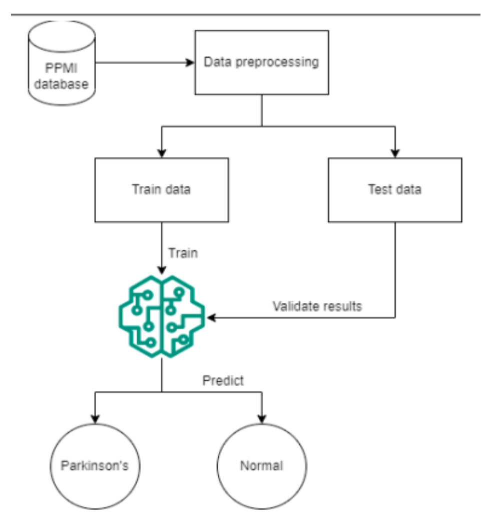


Figure 1: Proposed architecture

Figure 1 : Architecture of Parkinson's Disease Detection

Important The given diagram represents the proposed architecture for Parkinson's disease detection using a machine learning approach. The process begins with data acquisition from the PPMI (Parkinson's Progression Markers Initiative) database, followed by a data preprocessing stage to prepare the dataset. The preprocessed data is then split into training and testing datasets. The training dataset is used to train a machine learning model, while the test dataset is used to validate the model's performance. Once trained, the model predicts whether a given sample corresponds to a person with Parkinson's disease or a normal individual. The overall workflow ensures systematic learning and evaluation to improve the model's accuracy in detecting Parkinson's disease.

The research paper aims to identify the most relevant attributes in classification of PD and impact of imbalance in medical data in classification.

## III. DATASET

Biomedical The dataset provided contains 171 records with six attributes related to Parkinson's disease detection. It includes the **age** of individuals, **eye blinking** frequency, **handshaking** (tremors), and **MDVP: Fo (Hz)**, which represents the fundamental frequency of voice—a key parameter in diagnosing Parkinson's disease. The **result** column is a binary indicator where '1' signifies the presence of Parkinson's disease and '0' indicates a healthy individual. The **name** column is entirely empty and can be disregarded. This dataset appears to be structured for classification purposes, likely aimed at identifying Parkinson's disease based on a combination of vocal and physical symptoms. Further analysis could involve data visualization, feature importance evaluation, or even training a machine learning model to predict Parkinson's disease based on these attributes.

The table in the image presents a list of attributes and their corresponding purposes, specifically related to Parkinson's disease detection. It include various vocal measurements derived from the MDVP (Multi-Dimensional Voice Program) analysis which are used to assess the characteristics of speech affected by Parkinson's disease.

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | age | eye_blinki | handshaki | MDVP:Fo(I | result | name |
| 2 | 48 | 1 | 1 | 119.992 | 1 | |
| 3 | 46 | 1 | 0 | 122.4 | 0 | |
| 4 | 50 | 1 | 0 | 120.552 | 0 | |
| 5 | 49 | 1 | 1 | 95.73 | 1 | |
| 6 | 45 | 0 | 0 | 95.056 | 0 | |
| 7 | 45 | 0 | 1 | 91.904 | 0 | |
| 8 | 46 | 1 | 1 | 139.173 | 1 | |
| 9 | 50 | 1 | 0 | 152.845 | 0 | |
| 10 | 45 | 1 | 0 | 144.188 | 0 | |
| 11 | 48 | 0 | 0 | 168.778 | 0 | |
| 12 | 47 | 0 | 0 | 153.046 | 0 | |
| 13 | 46 | 0 | 1 | 156.405 | 0 | |
| 14 | 50 | 1 | 0 | 153.848 | 0 | |
| 15 | 47 | 1 | 1 | 153.88 | 1 | |
| 16 | 47 | 1 | 0 | 167.93 | 0 | |
| 17 | 45 | 1 | 0 | 104.4 | 0 | |
| 18 | 45 | 0 | 0 | 146.845 | 0 | |
| 19 | 45 | 0 | 0 | 162.568 | 0 | |
| 20 | 47 | 0 | 1 | 197.076 | 0 | |
| 21 | 49 | 0 | 1 | 199.228 | 0 | |
| 22 | 49 | 1 | 0 | 203.184 | 0 | |
| 23 | 47 | 1 | 1 | 201.464 | 1 | |
| 24 | 48 | 0 | 1 | 176.17 | 0 | |
| 25 | 45 | 0 | 1 | 180.198 | 0 | |

Figure 2 : Parkinson's Disease Patient Data.

MODEL TRAINING:

This research paper studies Logistic Regression, Random Forest classifier, Support Vector classifier and K nearest neighbors' models in 3 approaches.

- Complete dataset of 171 records and 6 attributes.
- Dataset with 171 records and 6 attributes after Principal Component Analysis (PCA).

## IV. RESULT AND DISSCUSION

After training the models, we achieved the following results.

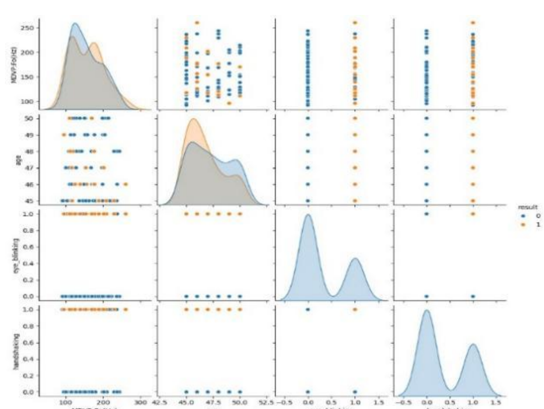


Figure 3 : The image shows the Pair Plot graph

The image represents a **pair plot visualization** used for exploratory data analysis (EDA) in Parkinson's disease detection. It shows the relationships between different numerical features in the dataset, such as MDVP: Fo(Hz) (vocal frequency), age, eye blinking, and handshaking. The diagonal plots display the distribution of individual features, while the off-diagonal scatter plots show pairwise relationships between features. The color coding indicates two classes: blue (healthy individuals) and orange (individuals with Parkinson's disease). The visualization helps identify patterns, such as how handshaking and eye blinking may be key indicators of the disease, while other features like MDVP: Fo(Hz) and age show some overlap between the two classes. This analysis aids in selecting relevant features for machine learning models to improve disease classification accuracy.

1.1 SUPPORT VECTOR MACHINE (SVM):

The confusion matrix for the SVM model shows perfect classification with 17 true positives and 17 true negatives, indicating 100% accuracy with no misclassifications. This suggests that the model is highly effective in distinguishing between the two classes.

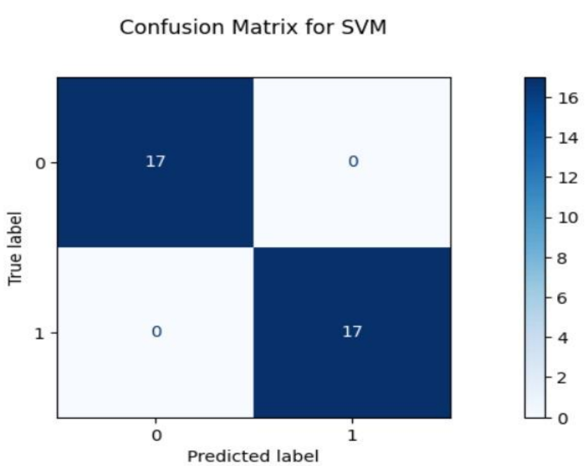


Figure 4: Confusion matrix for SVM

1.2 RANDOM FOREST:

The confusion matrix above illustrates the performance of the Random Forest model in detecting Parkinson's Disease. The model correctly classified all 17 healthy individuals (class 0) and all 17 Parkinson's patients (class 1), resulting in zero misclassifications.

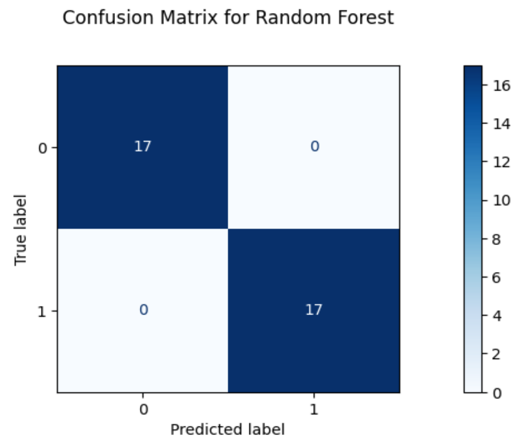


Figure 5: Confusion matrix for Random Forest Model

This indicates 100% accuracy, with no false positives (wrongly predicting Parkinson's for a healthy person) and no false negatives (failing to detect Parkinson's in an actual patient)

ACCURACY:

The model's precision, recall, f1-score, and accuracy were calculated:

| | Metric | DT | RF | LR | SVM | NB | KNN |
|---|---|---|---|---|---|---|---|
| 0 | Accuracy | 0.864407 | 0.949153 | 0.830508 | 0.983051 | 0.762712 | 0.949153 |
| 1 | F1-Score | 0.826087 | 0.941176 | 0.782609 | 0.980392 | 0.650000 | 0.938776 |
| 2 | Recall | 0.730769 | 0.923077 | 0.692308 | 0.961538 | 0.500000 | 0.884615 |
| 3 | Precision | 0.950000 | 0.960000 | 0.900000 | 1.000000 | 0.928571 | 1.000000 |
| 4 | R2-Score | 0.449883 | 0.793706 | 0.312354 | 0.931235 | 0.037296 | 0.793706 |

Figure 6 : Accuracy Table

The table compares machine learning models (DT, RF, LR, SVM, NB, KNN) based on accuracy, F1-score, recall, precision, and $R^2$-score. SVM performs best with 98.3% accuracy, 100% precision, and the highest $R^2$-score (0.931), making it the most reliable model. RF and KNN also show strong performance (94.9% accuracy), while NB performs the worst (76.2% accuracy, 50% recall). Logistic Regression has the lowest recall (69.2%), making it less effective for this dataset. Overall, SVM is the best choice, providing the highest predictive reliability.

## V. CONCLUSION

In this research, we developed and evaluated machine learning models for the accurate detection of Parkinson's disease. The results demonstrate that Support Vector Machine (SVM) outperforms other models with the highest accuracy, precision, and recall, making it the most suitable choice for early detection. Random Forest (RF) and K-Nearest Neighbors (KNN) also showed competitive performance, whereas **Naïve Bayes (NB) exhibited the weakest results** due to lower recall. The study highlights the importance of feature selection and data preprocessing in improving classification accuracy. Future work can explore **deep** learning techniques, larger datasets, and real-time implementation to enhance diagnostic efficiency and clinical applicability.

## VI. REFERENCES

[1] **D. Trivedi H. Jaeger and M. Stadtschnitzer. (2019).** "Mobile Device Voice Recordings at King's College London (MDVR-KCL) from both early and advanced Parkinson's disease patients and healthy controls.
Available: https://doi.org/10.5281/zenodo.2867216

[2] **Prabhavathi, K., Patil, S. (2022).**
A Review "Tremors and Bradykinesia. In: Arjunan, S.P., Kumar, D.K. (eds) Techniques for Assessment of Parkinsonism for Diagnosis and Rehabilitation", Series in Bio Engineering. Available: https://doi.org/10.1007/978-981-16-3056-9_9

[3] **Braak, H., Braak, E. (2000)**
"Pathoanatomy of Parkinson's disease" J Neurol 247, II3–II10 Available: https://doi.org/10.1007/PL00007758

[4] **Ali, L., Chakraborty, C., He, Z. et al. (2022)**, " "A novel sample and feature dependent ensemble approach for Parkinson's disease detection". Available:
https://doi.org/10.1007/s00521-022-07046-2

[5] **Aghav-Palwe S, Mishra D (2020)"** "A "Statistical tree-based feature vector for content-based image retrieval." Int J Computer Sci Eng 2020". Available: https://doi.org/10.1504/IJCSE.2020.106868

[6] **D. V. Rao, Y. Sucharitha, D. Venkatesh, K. Mahamthy and S. M. Yasin (2022)**, "Diagnosis of Parkinson's Disease using Principal Component Analysis and Machine Learning algorithms with Vocal Features,".2021 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), 2022, pp. 200-206, doi: 10.1109/ICSCDS53736.2022.9760962

International Journal of Scientific Research in Computer Science, Engineering and Information Technology

Peer Reviewed and Refereed International Scientific Research Journal

[ UGC Journal No : 64718 ] Scientific Journal Impact Factor : 8.154

# Certificate of Publication

Ref : Certificate/Volume 11/Issue 1/CSEIT2511142 23-Feb-2025

This is to certify that **CH. Lalitha Haripriya, CH. Sirisha, A. Gnaneswar, K. Nagendra, Dr. R. Rajender** have published a research paper entitled *'Parkinson's Disease Detection Using Machine Learning'* in the International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), Volume 11, Issue 1, January-February 2025.

This Paper can be downloaded from the following IJSRCSEIT website link
https://ijsrcseit.com/index.php/home/article/view/CSEIT2511142

**DOI : https://doi.org/10.32628/CSEIT2511142**

IJSRCSEIT Team wishes all the best for bright future

Editor in Chief
IJSRCSEIT

ISSN : 2456-3307

Website : https://ijsrcseit.com

IJSRCSEIT

IJSRCSEIT

# International Journal of Scientific Research in Computer Science, Engineering and Information Technology

*Peer Reviewed and Refereed International Scientific Research Journal*

**[ UGC Journal No : 64718 ]** **Scientific Journal Impact Factor : 8.154**

# Certificate of Publication

Ref : Certificate/Volume 11/Issue 1/CSEIT2511142 23-Feb-2025

This is to certify that **CH. Lalitha Haripriya** has published a research paper entitled 'Parkinson's Disease Detection Using Machine Learning' in the International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), Volume 11, Issue 1, January-February 2025.

This Paper can be downloaded from the following IJSRCSEIT website link
https://ijsrcseit.com/index.php/home/article/view/CSEIT2511142

**DOI : https://doi.org/10.32628/CSEIT2511142**

IJSRCSEIT Team wishes all the best for bright future

**Editor in Chief**
**IJSRCSEIT**

**ISSN : 2456-3307**

**Website : https://ijsrcseit.com**

IJSRCSEIT

IJSRCSEIT

International Journal of Scientific Research in Computer Science, Engineering and Information Technology

Peer Reviewed and Refereed International Scientific Research Journal

[ UGC Journal No : 64718 ] Scientific Journal Impact Factor : 8.154

# Certificate of Publication

Ref : Certificate/Volume 11/Issue 1/CSEIT2511142 23-Feb-2025

This is to certify that **CH. Sirisha** has published a research paper entitled 'Parkinson's Disease Detection Using Machine Learning' in the International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), Volume 11, Issue 1, January-February 2025.

This Paper can be downloaded from the following IJSRCSEIT website link
https://ijsrcseit.com/index.php/home/article/view/CSEIT2511142

**DOI : https://doi.org/10.32628/CSEIT2511142**

IJSRCSEIT Team wishes all the best for bright future

**Editor in Chief**
**IJSRCSEIT**

**ISSN : 2456-3307**

**Website : https://ijsrcseit.com**

IJSRCSEIT

International Journal of Scientific Research in Computer Science, Engineering and Information Technology

*Peer Reviewed and Refereed International Scientific Research Journal*

**[ UGC Journal No : 64718 ]** **Scientific Journal Impact Factor : 8.154**

# Certificate of Publication

Ref : Certificate/Volume 11/Issue 1/CSEIT2511142 23-Feb-2025

This is to certify that **A. Gnaneswar** has published a research paper entitled 'Parkinson's Disease Detection Using Machine Learning' in the International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), Volume 11, Issue 1, January-February 2025.

This Paper can be downloaded from the following IJSRCSEIT website link
https://ijsrcseit.com/index.php/home/article/view/CSEIT2511142

**DOI : https://doi.org/10.32628/CSEIT2511142**

IJSRCSEIT Team wishes all the best for bright future

**Editor in Chief**
**IJSRCSEIT**

**ISSN : 2456-3307**

**Website : https://ijsrcseit.com**

IJSRCSEIT

IJSRCSEIT

# International Journal of Scientific Research in Computer Science, Engineering and Information Technology

*Peer Reviewed and Refereed International Scientific Research Journal*

**[ UGC Journal No : 64718 ]** **Scientific Journal Impact Factor : 8.154**

## Certificate of Publication

Ref : Certificate/Volume 11/Issue 1/CSEIT2511142 23-Feb-2025

This is to certify that **K. Nagendra** has published a research paper entitled 'Parkinson's Disease Detection Using Machine Learning' in the International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), Volume 11, Issue 1, January-February 2025.

This Paper can be downloaded from the following IJSRCSEIT website link
https://ijsrcseit.com/index.php/home/article/view/CSEIT2511142

**DOI : https://doi.org/10.32628/CSEIT2511142**

IJSRCSEIT Team wishes all the best for bright future

**Editor in Chief**
**IJSRCSEIT**

**ISSN : 2456-3307**

**Website : https://ijsrcseit.com**

IJSRCSEIT

IJSRCSEIT

# International Journal of Scientific Research in Computer Science, Engineering and Information Technology

*Peer Reviewed and Refereed International Scientific Research Journal*

**[ UGC Journal No : 64718 ]** **Scientific Journal Impact Factor : 8.154**


# Certificate of Publication

Ref : Certificate/Volume 11/Issue 1/CSEIT2511142 23-Feb-2025

This is to certify that **Dr. R. Rajender** has published a research paper entitled 'Parkinson's Disease Detection Using Machine Learning' in the International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), Volume 11, Issue 1, January-February 2025.

This Paper can be downloaded from the following IJSRCSEIT website link

https://ijsrcseit.com/index.php/home/article/view/CSEIT2511142


**DOI : https://doi.org/10.32628/CSEIT2511142**


IJSRCSEIT Team wishes all the best for bright future

**Editor in Chief**
**IJSRCSEIT**


**ISSN : 2456-3307**

**Website : https://ijsrcseit.com**


IJSRCSEIT