EARLY DETECTION OF PARKINSON'S WITH MACHINE LEARNING FOR PREDICTIVE RISK MODELING

Wijayasooriya W. A. K. A

IT21164026

BSc. (Hons) Degree in Information Technology Specializing in Data Science

Department of Computer Science

Sri Lanka Institute of Information Technology
Sri Lanka

April 2025

EARLY DETECTION OF PARKINSON'S WITH MACHINE LEARNING FOR PREDICTIVE RISK MODELING

Wijayasooriya W. A. K. A

IT21164026

Supervisor: Ms. Wishalya Tissera

Co – Supervisor: Dr. Kapila Dissanayaka

BSc. (Hons) Degree in Information Technology Specializing in Data Science

Department of Computer Science

Sri Lanka Institute of Information Technology
Sri Lanka

April 2025

DECLARATION

I declare that this is my own work, and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to Sri Lanka Institute of Information Technology, the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date: 28/11/2024

The above candidate has carried out research for the bachelor's degree Dissertation under my supervision.

		Date	Sign
Supervisor	Ms. Wishalya Tisera		
Co-Supervisor	Dr. Kapila Dissanayake		

ABSTRACT

Parkinson's Disease (PD) is a chronic and progressive neurodegenerative condition that has serious impacts on motor, speech, and cognitive activities, especially among the elderly. Early diagnosis is an important aspect of efficient disease management and enhanced quality of life for patients. This study presents a multi-model machine learning model to improve the early diagnosis of Parkinson's Disease by using three distinct modalities: clinical histories, speech features, and handwriting patterns. Each of the data sources measures one unique aspect of PD symptoms—clinical records reflect lifestyle and cognitive, speech samples reveal vocal impairments, and motor abnormalities such as tremors and rigidity are revealed in handwriting patterns.

Three different predictive models were developed using optimized algorithms: Random Forest for clinical data, a transfer learning-based CNN model for speech analysis, and a MobileNetV2 model for handwriting samples. The models were then combined using a late fusion approach to give a final combined risk prediction. One of the key innovations in this research is the creation of a Multi-Level Disease Progression Indicator, categorizing patients into three risk groups—Green (Low Risk), Yellow (Moderate Risk), and Red (High Risk). This approach converts the conventional binary diagnosis to a more clinically meaningful and actionable framework.

To render the solution applicable in the real world, the solution is embedded in an internet-based medical professional decision support system. Physicians are able to input patient data in all three modalities, and the system provides real-time predictions and risk-level stratification. The model was strongly predictive across all modalities, and the fusion technique further enhanced its credibility. The research validates the applicability of multimodal learning and late fusion techniques to facilitate early diagnosis of PD and underlines the importance of interpretability and usability in practice.

Keywords - Parkinson's Disease, Machine Learning, Multi-Model Fusion, Clinical Data, Speech Analysis, Handwriting Recognition, Deep Learning, Decision Support System, Early Detection, Disease Progression Indicator.

ACKNOWLEDGEMENT

I would like to express my heartfelt appreciation to everyone who supported me throughout the journey of completing my fourth-year research project. First and foremost, I am deeply thankful to my research supervisor, Ms. Wishalya Tissera, whose invaluable guidance, encouragement, and steadfast support played a pivotal role in every phase of this research. I also wish to extend my sincere gratitude to my co-supervisor, Dr. Kapila Dissanayaka, for his expert advice and technical assistance, which greatly enriched the quality of this work. My appreciation goes to the CDAP panel for their constructive feedback and thoughtful suggestions that helped enhance the direction of our study.

I would like to acknowledge the support of Dr. Anuradha Baminiwatta, our external supervisor, for his efforts in coordination and for sharing essential insights that helped shape the early stages of the project. I am also thankful to our research project leader, Dilshan G.A.M, and my fellow team members for their collaborative spirit and contribution to data collection, analysis, and development tasks. I would like to thank the faculty and department for providing the required infrastructure, resources, and guidance needed for the successful execution of this project. I also acknowledge the contributions of any external organizations or funding entities (if applicable) that supported this research endeavor. Lastly, I am immensely grateful to my parents and friends for their constant encouragement and unwavering support, which were instrumental in completing this study.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Full Form
PD	Parkinson’s Disease
ML	Machine Learning
DL	Deep Learning
RF	Random Forest
DT	Decision Tree
XGBoost	Extreme Gradient Boosting
CNN	Convolutional Neural Network
LSTM	Long Short-Term Memory
MFCC	Mel-Frequency Cepstral Coefficients
PCA	Principal Component Analysis
CDSS	Clinical Decision Support System
GUI	Graphical User Interface
API	Application Programming Interface
IoT	Internet of Things
SHAP	SHapley Additive exPlanations
UPDRS	Unified Parkinson's Disease Rating Scale
MoCA	Montreal Cognitive Assessment
HC	Healthy Control (label in datasets)
PwPD	People with Parkinson’s Disease

Table 1 List of abbreviations

1. INTRODUCTION

1.1. Background & Literature Survey

Parkinson's Disease (PD) is a chronic and progressive neurodegenerative disorder that affects millions of people worldwide, with immense implications for their health care, autonomy, and quality of life. PD primarily influences motor function by inducing tremor, bradykinesia (slowness of movement), muscle rigidity, and postural instability. In addition to these, patients also suffer from non-motor symptoms such as cognitive impairment, speech impairment, mood disorders, and sleep disturbances. They occur in a step-wise manner and add up in severity over time and finally present with disabilities in activities of daily living and care giving. Due to the insidious nature of its initial features, early diagnosis of PD remains a major clinical challenge.

The World Health Organization (WHO) regards PD as the second most common neurological disease after Alzheimer's disease. The global burden of PD has more than doubled in the last two decades, and it is projected that over 12 million people will be living with PD by 2040 [1]. This dramatic rise calls for more efficient, automated, and accurate approaches to diagnosis and disease monitoring. In the traditional diagnosis process, neurologists are largely reliant on subjective clinical scores, such as the Unified Parkinson's Disease Rating Scale (UPDRS), and physical examination of symptoms. These human-based approaches are time-consuming, highly skill-dependent on clinicians, and susceptible to diagnostic heterogeneity.

The development of artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), has created new opportunities for automating PD detection using multimodal data sources. These are systematic clinical information, speech audio recordings, and handwriting samples. Each of them records unique and relevant expressions of PD. Clinical information, for example, is made up of demographic, lifestyle, and medical history markers; speech signals

give us hints about articulation and phonation abnormalities; and handwriting manner reflects fine motor deficiencies such as micrographia and tremors. Integration of these modalities is expected to bring a better overall robustness and overall generalizability of diagnostic models.

Some researchers have directly examined the contribution of single data modalities to PD diagnosis. Prashanth and Roy [2] used a predictive model with patient questionnaire data, employing supervised ML techniques for predicting PD presence. They established clinical knowledge and formalized data as important factors in improving classification performance. Lahmiri et al. [3] demonstrated the superiority of speech-based dysphonia measures such as jitter, shimmer, and pitch variability. They employed a Support Vector Machine (SVM) classifier that achieved high rates of accuracy in distinguishing between PD-affected and healthy voices.

Moreover, handwriting analysis emerged as an available tool for the identification of PD at early stages. Kotsavasiloglou et al. [4] investigated movement during handwriting and drawings based on motor dysfunction identification using machine learning algorithms. The model with proposed architecture had high performance when trained on handwritten spirals and basic forms. Nevertheless, though each individual modality contains much information for comprehending, modalities are deficient and cannot represent the richness and complexity inherent with PD.

One of the most significant advances in this field is the application of multiple modalities, or multimodal learning. This approach combines the best of numerous disparate sources of information, increasing predictive capability and reducing false positives. Wang et al. [5] and Parisi et al. [6] examined multimodal models that combined data from gait, speech, and handwriting patterns. Their results showed that fusion models greatly outperformed unimodal models in terms of accuracy and sensitivity. These results emphasize the importance of data fusion in creating clinically acceptable decision-support systems.

A multi-faceted detection system combining facial features, wearable data, and speech patterns was described in the current study by Dissanayaka et al. [7]. Their research proposed a modular diagnostic tool capable of monitoring in real time and predicting phases. The system was most useful to the approach utilized in our investigation. While their system incorporated novel factors like wearable sensor data, our approach attempts to implement a stage-wise risk classification system with late fusion based on clinical, audio, and handwriting inputs.

Besides, most existing models for PD detection boil down the outcome to a binary classification (PD or non-PD), and this does not reflect the progressive nature of the disease. In order to counter this limitation, we introduce a Multi-Level Disease Progression Indicator, which categorizes risk into three levels: Low, Moderate, and High. Such stratification supports early intervention and clinical planning through a more accurate representation of disease severity.

The characteristic aspect of our proposed model is its reliance on decision-level fusion (late fusion), wherein predictions of each individual model are combined based on weights allocated to them. Not only does this method of fusion maintain autonomy of each modality, but also easy integration into real-world systems is feasible. Clinical information is considered the most reliable source and given a greater weightage during fusion in our case, while speech and handwriting models present supplementary evidence.

Another key strength of our work is the real-world implementation via a web-based interface. The system enables doctors to feed in patient information in real time—clinical parameters via digital entry, audio entries via access to microphones, and handwriting samples via image upload—hence providing accessibility and usability within clinical environments.

Overall, our research contributes to the existing literature in three significant ways:

1. Multimodal Integration: We fuse three heterogenous data type clinical records, speech, and handwriting—to form a comprehensive diagnosis model.
2. Late Fusion Strategy: Through decision-level fusion, we present modular and explainable predictions.
3. Disease Progression Output: Unlike other research, we propose a 3-level risk stratification (Green, Yellow, Red), offering more useful insights for future patient care.

This approach addresses some of the most critical gaps identified in previous work and is part of developing intelligent, real-time, and accessible tools for Parkinson's Disease diagnosis and monitoring.

1.2. Research Gap

Parkinson's Disease (PD) is a complex neurodegenerative condition that is linked with a variety of motor and non-motor symptoms, including tremors, bradykinesia, rigidity, and cognitive impairment. It is crucial for effective treatment and intervention that its identification is made early and precisely. While there have been numerous attempts at using machine learning (ML) techniques for supporting PD diagnosis, there are some critical gaps in the current research body.

Unimodal Limitations

Most of the PD detection studies have been on unimodal data sources. For instance, Prashanth and Roy [2] trained predictive models on patient questionnaires only, with good accuracy in discriminating early PD from healthy controls. Similarly, Lahmiri et al. [3] utilized dysphonia measures from speech data for PD diagnosis, with promising results. Kotsavasiloglou et al. [4] analyzed handwriting patterns for identifying PD-related abnormalities. While these unimodal approaches have been effective within their target domains, they are by definition constrained in their potential to present a unified picture of the multifaceted nature of PD symptoms.

Emergence of Multimodal Approaches

Recognizing the limitations of unimodal analysis, recent efforts have made forays into multimodal analysis. Dissanayaka [7] suggested a multi-faceted framework with sensor signals, facial expressions, and speech for tracking PD stages, highlighting the potential of merging heterogeneous data streams. Wang et al. [5] pointed out that hybrid approaches, fusing speech and clinical data, could enhance detection accuracy. However, these studies have to deal with challenges of heterogeneity of data, increased computational complexity, and the need for sophisticated fusion techniques to integrate various kinds of data.

Risk Stratification Challenges

The majority of existing ML models for PD detection are directed towards binary classification—PD or non-PD. Binary classification neglects the progressive nature of PD and the fact that early intervention could significantly alter disease trajectories. Models with the potential to stratify risk levels, providing more precise information about disease progression and allowing early therapeutic interventions, are sorely needed.

Explainability and Clinical Integration

The other critical gap in current research is the lack of explainability in ML models. Black-box models, while often accurate, do not offer the transparency clinical decision-making requires. Explainable AI (XAI) methods must be used to interpret model predictions, thereby facilitating trust and easier integration into clinical practice. Furthermore, most current models are developed in research settings with little regard for real-world deployment, which limits their day-to-day applicability.

Proposed Solution

Addressing these gaps, our research introduces a comprehensive multimodal system integrating clinical data, speech, and handwriting analysis for enhanced PD detection. The system incorporates a Multi-Level Disease Progression Indicator, stratifying patients into low, medium, and high-risk profiles, and thus acknowledging the progressive nature of PD. In the interest of transparency and ease of clinical uptake, we employ explainable AI techniques, with model predictions supported by explicit rationales. Moreover, the model is designed to be readily incorporated into clinical decision-support systems for real-time, user-friendly analysis.

Through the incorporation of multimodal data, the use of risk stratification, and the prioritization of model interpretability, our approach aims to overcome the limitations of existing methods, offering an effective tool for the early and accurate detection of PD.

Feature / Approach	[2]	[3]	[4]	[5]	[7]	Proposed Solution
Clinical data-based prediction						
Speech/audio-based analysis						
Handwriting/ motor symptom analysis						

Use of separate models for each data type						
Late fusion-based multimodal integration						
Real-time prediction via web application						
Multi-level disease progression indicator (Green-Yellow-Red)						
Feature selection using statistical methods						
Temperature-scaled softmax in fusion stage						

Table 2 Comparison with previous studies

1.3. Research Problem

Parkinson's Disease (PD) is a complex, progressive neurodegenerative disorder that affects motor and non-motor functions and results in significant loss of the

quality of life in the afflicted individuals. Accurate and timely diagnosis of PD is extremely important since early intervention can delay disease progression, improve patient outcomes, and reduce long-term healthcare system costs. Notwithstanding this, conventional diagnostic methods are still very much reliant on human clinical evaluation, which is plagued by human bias, variability of expert opinion, and delayed diagnosis. Obviously, automation-based, objective, and multimodal data-driven solutions are desperately needed for early diagnosis and continuous monitoring of Parkinson's Disease.

The past decade has witnessed a plethora of research studies that have sought to utilize machine learning (ML) techniques to automate the identification of PD. These approaches are typically grounded on characteristics of one modality—either clinical symptoms, speech impairments, or handwriting traits. For instance, studies by Prashanth & Roy [2] were grounded on clinical information and questionnaire responses to construct prediction models, whereas Lahmiri et al. [3] employed dysphonia characteristics based on voice recordings to discriminate PD from healthy controls. Similarly, Kotsavasiloglou et al. [4] explored handwriting dynamics to determine motor dysfunctions in PD patients. Although these unimodal methods have yielded encouraging results, they tend to be limited with respect to their generalizability, prone to false positives or false negatives, and inadequate in capturing the PD symptom's multidimensional nature.

There is a glaring gap in existing research that includes a scarcity of fusion-based multimodal systems equipped with the ability to selectively integrate knowledge from different data sources—clinical tests, voice features, and handwriting attributes. Even though all these modalities contain inherent diagnostic utility, their combined analysis gives greater understanding of the symptom heterogeneity and disease process. Research conducted by Wang et al. [5] and Dissanayaka [7] has shown that hybrid systems that utilize combinations of several feature domains have significant impact on predictive accuracy and reliability. However, there is a gap in research in developing effective, balanced,

and interpretable fusion mechanisms to actually integrate these heterogeneous inputs in practical real-world clinical settings.

Additionally, the task is not merely in establishing the presence or absence of Parkinson's Disease, but also establishing the level of severity or advancement of the disease. Most existing research restricts the diagnosis to a binary outcome—PD or non-PD—without considering the variety of symptom progressions that patients experience. The binary presentation is insufficient for guiding treatment decisions, rehabilitation planning, and counseling patients. What is needed is a system of risk stratification that moves beyond classification and provides clinicians with a progression-based diagnosis, such as early-stage, mid-stage, or advanced PD.

In this paper, a new solution is implemented to overcome the above mentioned limitation by proposing a multi-model fusion framework that unifies the three different machine learning models' predictions trained on disparate modalities of clinical data, speech features, and handwriting traces. This fused system is enriched with a Multi-Level Disease Progression Indicator that categorizes the patients into three different risk bins—Low Risk (Green), Moderate Risk (Yellow), and High Risk (Red). This approach aligns with the clinical need for precise, interpretable, and patient-specific diagnostics, with the ultimate purpose of reducing diagnostic delay and empowering clinicians with actionable decision-support tools.

A simple and realistic web-based interface is also utilized to serve as the system's front-end. The interface gives healthcare professionals the ability to input patient data with the utilization of forms, microphone support for voice input, and handwriting capture with the integration of smart tablets. The backend includes the late fusion logic internally to enable real-time prediction. Not only does this make the solution scalable and deployable in clinical environments, but it also enables a connection between AI research and healthcare practice.

Overall, the research problem under investigation in this case is one of inadequate efficacious, comprehensible, and multi-domain diagnostic frameworks for Parkinson's Disease. Suggested in this paper is a comprehensive framework that incorporates three necessary modalities with risk stratification in a way that provides a higher and more actionable output for primary detection and disease surveillance.

2. OBJECTIVES

2.1. Main Objectives

Primary Goals

The primary objective of this study is to design an intelligent, multimodal machine learning system for early diagnosis and risk-level identification of Parkinson's Disease (PD). The proposed solution aims to exploit the power of more than one source of data—clinical, speech, and handwriting features—to overcome the limitations of single-modality diagnosis systems and attain a more resilient, accurate, and interpretable decision-support system for doctors.

Parkinson's Disease manifests differently in individuals, and symptoms range from motor dysfunction and voice disorder to cognitive and behavioral disorder. Hence, the use of any single modality for diagnosis often leads to low generalizability and reduced diagnostic performance. To bypass this limitation, the present study proposes a multi-model late fusion approach, which processes each modality independently using specialized machine learning models and then combines their predictions using a weighted decision-level fusion method.

One of the key contributions of this study is the development of a Multi-Level Disease Progression Indicator, which stratifies patients into three levels—Low Risk, Moderate Risk, and High Risk—based on the combined prediction outputs. This stratified diagnostic output has higher clinical utility, as it enables physicians to personalize treatment and monitoring protocols in accordance with the severity of the patient's condition.

Additionally, the system will be rendered practical and deployable through the development of a user-friendly web-based application interface for input of data from patients and clinicians in real-time. This encompasses:

- Clinical data input through structured forms,
- Voice sample upload by utilizing access to microphones,
- Handwriting sample upload captured using a smart tablet interface.

The ultimate goal in the long term is to bridge the gap between state-of-the-art machine learning research and real-world medical diagnostics, so that the system is not only technically sound but also accessible and deployable in clinical practice. Explainability, reliability, and patient-centered design are emphasized in this project while introducing a novel, data science-driven approach for neurodegenerative disease diagnosis.

2.2. Specific Objectives

To achieve the overall aim of this research, a set of specific objectives was formulated to guide the design, implementation, and evaluation of the proposed system for detecting Parkinson's Disease. These specific objectives are development and integration of multiple machine learning models, multimodal data source integration, and providing actionable diagnostic output through an accessible software platform. These specific objectives are discussed below:

1. To collect and preprocess multimodal data related to Parkinson's diagnosis, including:
    - Structured clinical information that contain demographic, medical, and motor symptom indicators,
    - Sustained phonation voice samples and speech utterances with dysphonia impact and changes in articulation,

- Handwriting samples with fine motor impairment on a smart writing tablet.

2. To build three standalone machine learning models, each specialized in one modality of data:

- A Random Forest classifier for clinical structured data to determine motor, cognitive, and lifestyle features,
- A transfer learning-based deep neural network for voice signal analysis using MFCC features,
- A CNN model based on MobileNetV2 for handwriting pattern classification and tremor detection.

3. To utilize model-specific preprocessing and feature engineering techniques:

- Selection of features by Chi-square tests and normalization of clinical features,
- Extraction of Mel-Frequency Cepstral Coefficients (MFCCs) and data augmentation for speech signals,
- Image resizing and normalization for easy training of handwriting image models.

4. To apply a late fusion method to blend the predictions of the three models through weighted averaging, with careful calibration of weights based on reliability and performance. This fusion method enables more accurate, more thorough risk classification by utilizing the advantages of each modality.

5. To develop a probabilistic Multi-Level Risk Indicator that classifies patients as Low, Moderate, or High Parkinson's risk levels, thus extending beyond binary classification and offering graded diagnostic outcomes.

6. To create and combine a web-based application interface, which allows:

- Real-time clinical data entry,
- Patient handwriting and voice inputs uploading
- Risk level output visualization to aid physician decision-making.

7. To examine and confirm the proposed framework with:

- Accuracy, precision, recall, F1-score,
- Confusion matrix and classification reports,
- Comparison between a single and combined multimodal model performance.

8. To bridge documented research gaps with a scalable multimodal system with enhanced early-stage PD detection accuracy, usability in real-world settings, and the inclusion of explainability for healthcare AI deployment.

These aims combined allow the production of an integrative, real-time diagnostic tool that can enable clinicians to identify early Parkinson with more confidence and reduce diagnostic uncertainty by combining data.

3. REQUIREMENT GATHERING & FEASIBILTY STUDY

3.1. Requirement Gathering & Analysis

In early stages of system development, gathering and understanding the requirements of the proposed solution is key to ensure that the final product meets both user needs as well as technical requirements. The system will be primarily used by medical doctors, neurologists, and patients, and it is expected to contribute to early diagnosis as well as monitoring of Parkinson's Disease (PD) using multimodal data (voice, writing, and clinical). The analysis of requirements was conducted through a mix of stakeholder meetings, analysis of similar systems, and review of clinical workflows.

3.1.1. Functional requirements

Functional requirements state the specific operations and behaviors the system must perform. For the proposed Parkinson's Risk Prediction System, the following functional requirements were determined:

ID	Functional Requirement
FR1	The system shall allow doctors to input clinical data through a structured form.
FR2	The system allow patients to upload voice samples in .wav format.
FR3	The system shall allow patients to upload handwriting samples in .png/.jpg format.

FR4	The system shall process all three data modalities using pre-trained machine learning models.
FR5	The system shall return a risk level prediction: Low, Moderate, or High based on late fusion.
FR6	The system shall display a Multi-Level Disease Progression Indicator with interpretable output.
FR7	The system shall store patient data securely in an internal database.
FR8	The system shall allow authorized users (e.g., doctors) to view prediction history and reports.

Table 3 Functional requirements of the system

These requirements focus on the core interactions between users and the system, supporting the multimodal approach to risk detection and facilitating efficient diagnosis.

3.1.2. Non-functional requirements

Non-functional requirements define the system quality attributes, ensuring usability, reliability, scalability, and security.

ID	Non-Functional Requirement
NFR1	The system should provide risk level predictions within 10 seconds after all inputs are provided.
NFR2	The system should have a user-friendly and intuitive web-based interface.
NFR3	The system should achieve a prediction accuracy of over 90% on real test data.
NFR4	The system must ensure data privacy and comply with standard security practices.
NFR5	The system should be scalable to integrate new modalities (e.g., gait analysis) in the future.
NFR6	The system shall display a Multi-Level Disease Progression Indicator with interpretable output.
NFR7	The system must be reliable and available 24/7 for online access by healthcare providers.

NFR8	The models must be updatable without requiring complete retraining of the system.

Table 4 Non-functional requirements of the system

3.2. Feasibility Study

A feasibility study is vital in identifying the viability of implementing the proposed Parkinson's Disease Detection and Monitoring System. It helps us measure whether the project can be undertaken, whether it is economically viable, and whether it can be completed within the predetermined timelines and resources. This section looks at the feasibility of the system schedule with a detailed project timeline.

3.2.1. Scheduled feasibility

Schedule feasibility is concerned with determining whether the proposed system can be built and delivered within the planned timeframe. The research project runs over two university semesters, which allows adequate time for development, testing, and assessment. The project was broken down into a number of phases, each with deliverables and timelines.

The most important milestones and deliverables were spread over the months to enable steady progress over a year. Principle work on developing principles with model training, multimodal fusion integration, validation, and front-end development were scheduled along with documentation and presentation preparation.

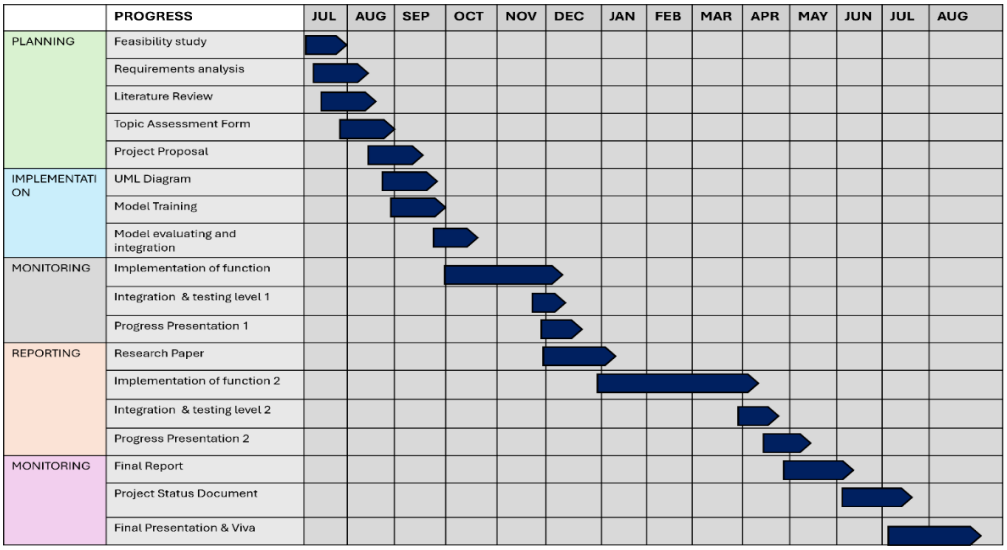


Figure 1 Gantt chart

3.2.2. Technical feasibility

Technical feasibility is a determination of the technical resources, tools, frameworks, and skill sets for successful implementation of the system proposed. For this project, the Early Detection of Parkinson's Disease system is built using a combination of Python-based Machine Learning libraries, pre-trained deep learning models, and a basic web interface with feasible implementation in the academic timescale and scope.

The system architecture combines several technical components:

- **Data Handling:** Pandas and scikit-learn are used to handle clinical data. Librosa and TensorFlow are used to handle audio data. TensorFlow's image dataset utilities are used to handle handwriting images.
- **Machine Learning & Deep Learning:** Models such as Random Forest, Transfer Learning with MobileNetV2, and DNN-based models were created and trained successfully with available datasets.
- **Multimodal Fusion:** Late fusion method was employed using Python and NumPy, combining outputs from all models in a weighted decision-level manner.
- **Front-End & Back-End Integration:** The application will be implemented as a web application using Flask or Streamlit for easy access and usability by clinicians.

All chosen technologies are open-source, well-documented, and compatible with Google Colab and local environments, making this solution technically achievable without additional hardware or proprietary software constraints.

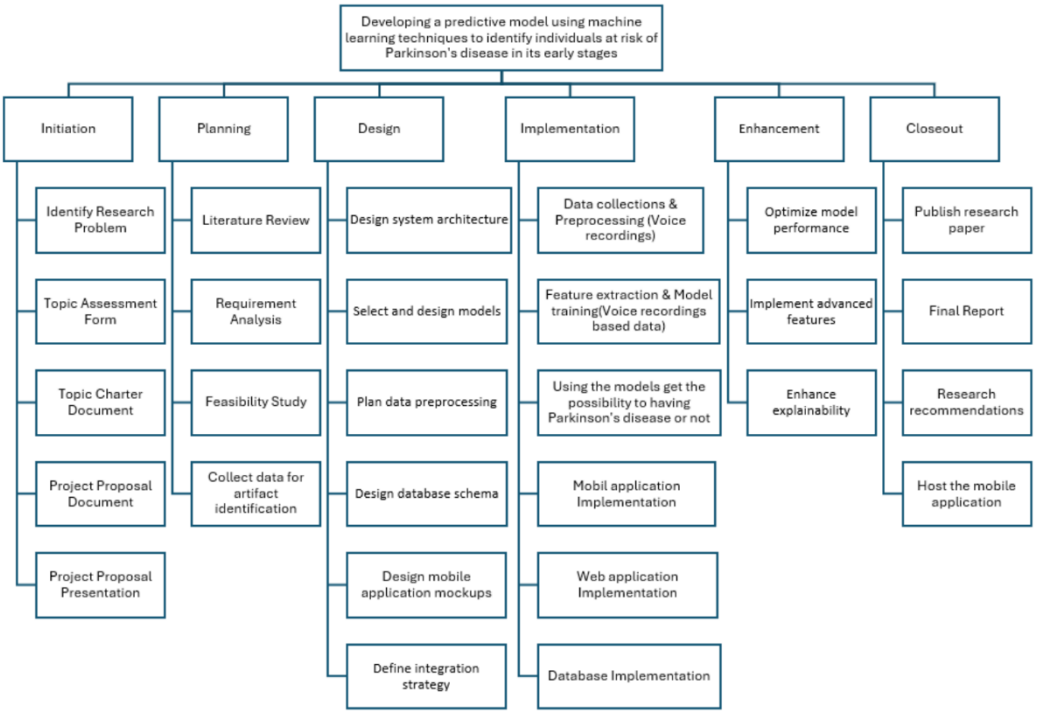


Figure 2 Work breakdown chart

3.2.3. Economic feasibility

The cost feasibility of the Parkinson's risk detection system was determined by analyzing the cost of its development and implementation. The primary costs include cloud services and research tools ($31.37 / LKR 10,000), data collection through open sources ($6.77 / LKR 2,000), and utilization of the cloud platform, including AWS, GCP, and OpenAI API ($40.00 / LKR 13,000). The cost of implementation is estimated to be $78.14 (LKR 25,000). These investments are justified by the potential benefits of early identification of ASD, including improved accessibility, faster diagnosis, and more effective intervention methods, to render the system economically viable.

Component	Est.Amount in USD	Est.Amount in LKR
Charges for Tools for Research (Cloud Services, Grammarly, etc.)	31.37	10,000.00
Data Collection through open sources	6.77	2000.00
Cloud Platforms (AWS,GCP,OpenAI key)	40.00	13,000.00
Total	78.14	25,000.00

Table 5 Budget allocation

3.2.4. Operational feasibility

Operational feasibility determines the extent to which the proposed Parkinson's detection system is going to be efficient under its intended environment and if it can be readily integrated into actual healthcare processes. In this research, the resulting system is designed to be user-friendly with little technological knowledge required to operate the system. Clinicians can enter clinical notes through an electronic form, upload speech samples via a microphone interface, and record handwriting samples via a smart writing tablet—all from a single web-based application.

The risk output (Low, Moderate, High) will assist in early physician decision-making, patient treatment planning, and patient follow-up. As the application is web-based, it offers ready access across disparate hospital or clinic settings without complicated installations or sophisticated equipment. Secondly, the modularity of the system design also allows for easy expansion in the future, i.e., expanding biomarkers or real-time feedback from clinicians.

Hence, the system is viable at an operational level as it meets actual user requirements, diminishes training requirements, and increases the accuracy and effectiveness of Parkinson's risk estimation in clinical application.

4. METHODOLOGY

4.1. Overall Architecture

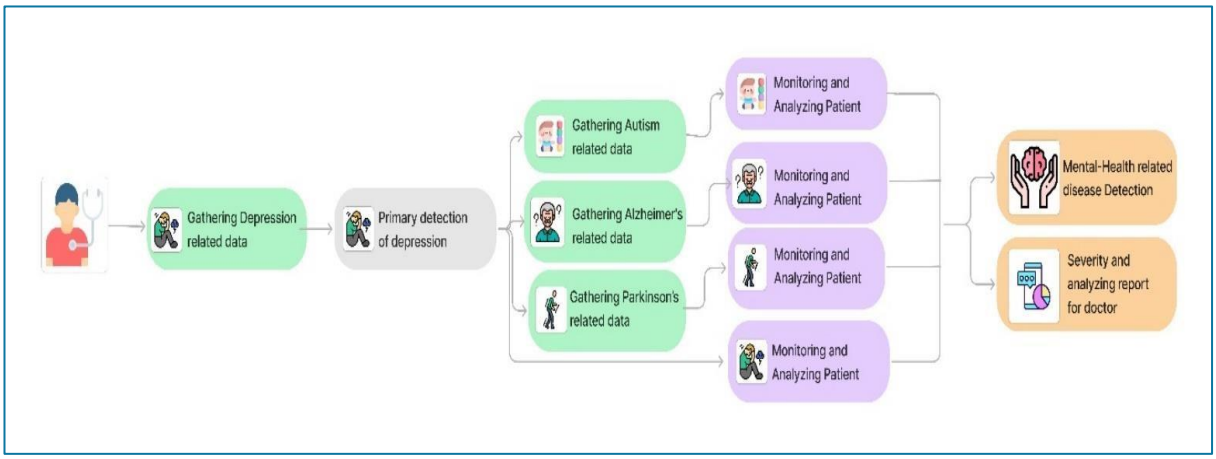


Figure 3 Overall architecture diagram

The diagram illustrates a step-by-step process for identifying and monitoring various mental health conditions. Initially, when a patient consults a doctor, relevant data related to depressive symptoms is collected, and a preliminary assessment is conducted to determine if depression is present. If signs of depression are confirmed, the system proceeds to gather additional information to screen for other conditions such as Parkinson's disease, Alzheimer's, or autism. Once any of these disorders are detected, the patient is enrolled into a continuous tracking system where their medical history and symptom progression are documented over time. This collected data is then compiled into detailed reports, offering doctors valuable historical insights and symptom severity levels, ultimately aiding them in making better-informed decisions for treatment and patient management.

4.2. Component Diagram

This segment outlines the baseline system architecture and methodical process designed for the early detection and risk-level classification of Parkinson's Disease (PD). As seen in Figure X (Component Diagram), the system integrates multiple data modalities to facilitate accurate, reliable, and timely detection of PD indicators to permit proactive clinical intervention. The solution framework involves three main stages:

Multimodal Feature Acquisition – The system takes input from three different sources: structured clinical records through a digital form, speech samples through microphone input, and handwriting images captured using a smart writing tablet. Each modality is capturing a different clinical aspect of Parkinson's progression, e.g., motor symptoms, vocal impairments, and diagnostic history.

Individual Model Training – Three separate machine learning models are trained on curated datasets for each modality. The clinical model is trained using standard tabular ML algorithms (e.g., Random Forest), while the speech and handwriting models are trained using deep learning approaches with transfer learning using CNNs.

Late Fusion-Based Risk Prediction – The output of the individual models is combined at the decision level by a late fusion approach with pre-decided weights (e.g., clinical: 0.5, speech: 0.3, handwriting: 0.2) to generate a single risk-level classification: Low, Moderate, or High. This yields a robust final decision according to each modality's strength, leading to enhanced clinical reliability.

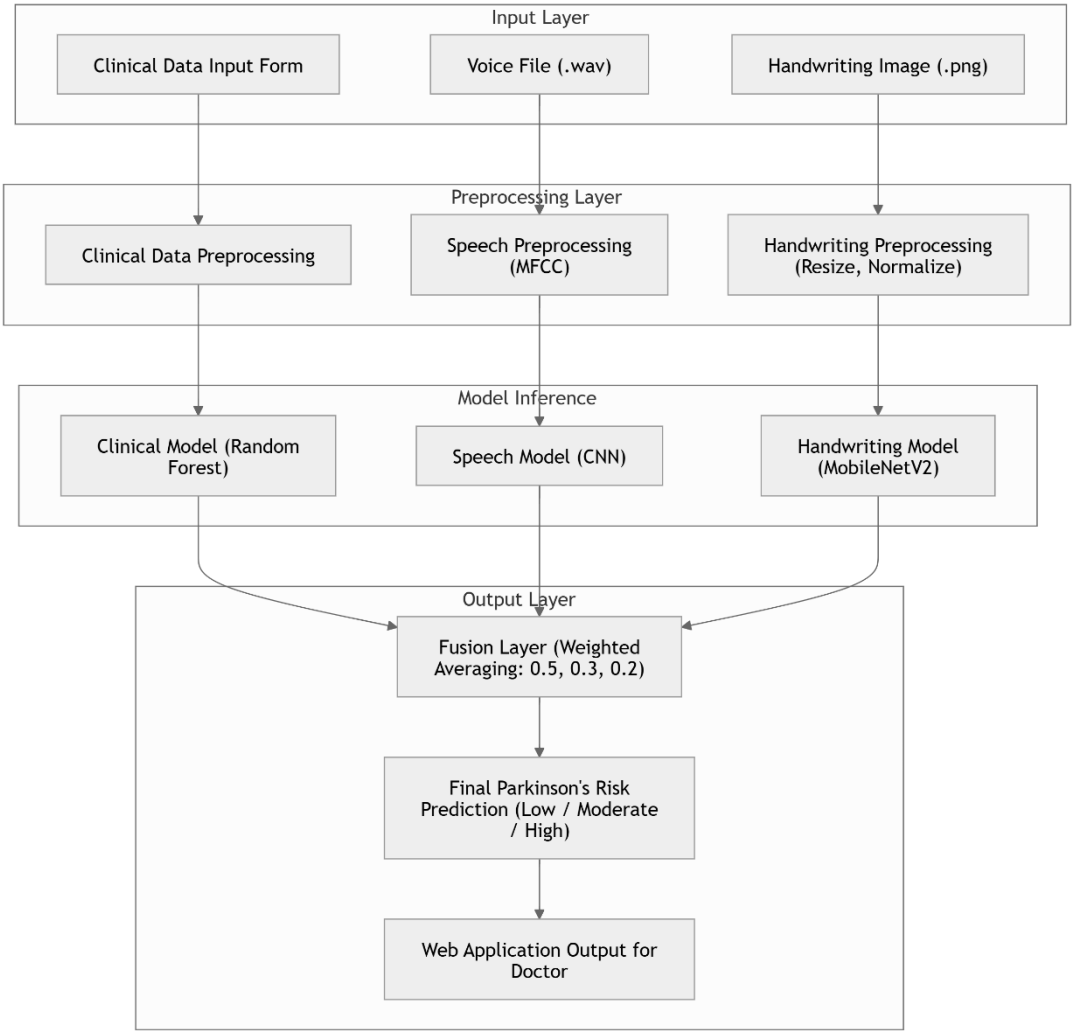


Figure 4 Component-level design of the parkinson's risk prediction system

The multimodal fusion system developed in this framework goes beyond single-modality conventional approaches in that it provides a more comprehensive picture of patient conditions, raises confidence in diagnosis, and enables integration in real-time into clinical decision-making systems.

4.3. Proposed Methodology

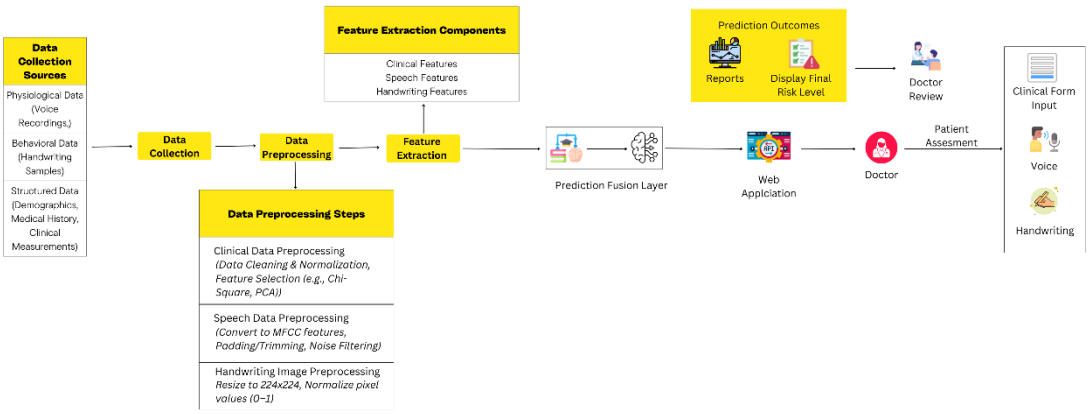


Figure 5 System architecture of the multimodal Parkinson's detection framework

4.3.1. Data collection layer

The Data Collection Layer forms the lowest level of the proposed multimodal system for PD risk determination. Its role is the acquisition of raw data from diverse input channels in order to support comprehensive and accurate diagnosis. A tri-modal data collection was adopted in this work, including clinical data, speech/audio recordings, and handwriting. Every modality was selected to record an individual symptom domain of Parkinson's: motor, vocal, and neuromuscular impairment. With the fusion of these modalities, the system enhances the robustness and sensitivity of PD detection at varying levels of disease severity.

4.3.1.1. Clinical data

The clinical information utilized in this study include real clinical and demographic traits collected from Parkinson's Disease patients and non-Parkinson's Disease patients. The information were keyed in manually to a structured electronic form interface created in the system, which is a replica of a real real-time clinical environment. The inputs are:

- Demographics: Age, gender

- Lifestyle factors: Alcohol consumption, smoking habits
- Medical History: Cholesterol levels, UPDRS score (Unified Parkinson's Disease Rating Scale), Montreal Cognitive Assessment (MoCA), and various motor and non-motor symptom scores such as:
    - Tremor
    - Rigidity
    - Bradykinesia
    - Postural Instability
    - Functional Assessment

The initial dataset contained irrelevant or personally identifiable information such as PatientID and DoctorInCharge, which were deleted to maintain patient confidentiality and only work with predictor attributes. The other attributes were label encoded for target variable Diagnosis, converting "Healthy" and "Parkinson" labels to binary (0 and 1 respectively).

For improving model efficiency and reducing dimensions, a Chi-Square Feature Selection method was employed. The system could then retain the 10 most significant features with high statistical correlation to the result of the Parkinson's diagnosis.

Around ~2000+ records were used, and a wide and balanced data set was used to train the clinical ML module. This module captures the most quantitative and organized symptoms for first-step PD risk estimation.

4.3.1.2. Voice samples

The audio/speech corpus addresses the speech impediments arising from Parkinson's, such as voice tremors, speech monotonicity, and articulation. Patients were requested to provide voice samples through the system interface, facilitating microphone access to record short audio clips. All files recorded were stored in.wav format with a sampling rate of 16kHz to ensure compatibility with typical industry audio feature extractors.

The dataset was divided into two categories:

- Healthy (HC_AH) audio recordings
- Parkinson's (PD_AH) audio recordings

Each recording was accompanied by metadata such as age, sex, and sample ID. There was a custom-built preprocessing pipeline that aligned each audio file to its corresponding metadata entry with traceability and integrity across the dataset.

To increase dataset volume and prevent overfitting, a data augmentation strategy was applied, including:

- Noise injection
- Pitch shifting
- Time stretching

This expanded the dataset significantly, allowing the model to generalize well on real-world noisy inputs.

An example of the data distribution:

- ~1200 original audio files
- ~3600 files after augmentation

This layer captures the phonetic and prosodic cues related to PD and feeds into the audio-based neural network model.

4.3.1.3. Handwriting samples

Handwriting deterioration is one of the most common early symptoms of Parkinson's, specifically micrographia (abnormally small and cramped handwriting). The patients had to provide handwriting input by using a smart writing tablet or uploading scanned handwriting sample images via the interface for this modality.

The dataset comprises over 3,000+ handwriting image samples (PNG format) categorized into:

- Healthy
- Parkinson's

Each image captures freehand writing, commonly the sentence: "The quick brown fox jumps over the lazy dog", enabling the model to assess:

- Stroke smoothness
- Letter size consistency
- Pen pressure
- Writing velocity

Images were normalized to 224×224 pixels and batched for training of deep learning models using transfer learning over a pretrained MobileNetV2 backbone. Data was loaded using TensorFlow's image_dataset_from_directory, supporting structured training, validation, and testing splits.

Based on handwriting patterns, the model detects fine motor control impairments, which are added to the overall PD risk scoring.

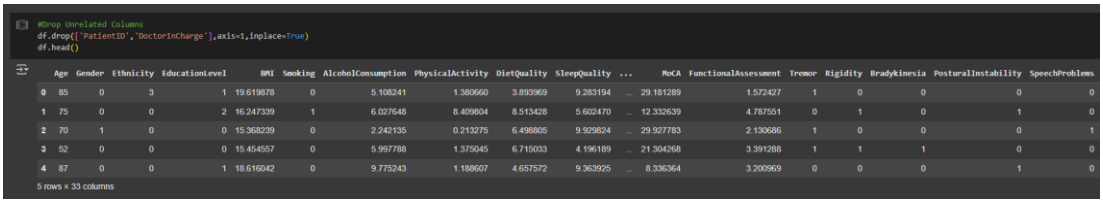
4.3.2. Data preprocessing layer

Data Preprocessing Layer plays a critical role in transforming raw inputs to effective model training and inference across the three modalities. Because each data source is of different format and structure—clinical, audio, and handwriting—the preprocessing pipelines were tailored specifically to handle each modality separately and be backward compatible with their respective model architectures.

4.3.2.1. Clinical data preprocessing

Clinical data was preprocessed to remove noise, reduce dimensionality, and normalize features:

- **Label Encoding:** The target column Diagnosis was label-encoded (Healthy → 0, Parkinson's → 1) to enable binary classification.
- **Column Filtering:** Unnecessary columns like PatientID and DoctorInCharge were dropped.
- **Feature Selection:** Applied Chi-Square feature selection to retain only the most relevant 10 features.
- **Train-Test Split:** Dataset was split into 80% training and 20% testing using train_test_split.



```
#Drop Unrelated Columns
df.drop(['PatientID','DoctorInCharge'],axis=1,inplace=True)
df.head()
```

	Age	Gender	Ethnicity	EducationLevel	BMI	Smoking	AlcoholConsumption	PhysicalActivity	DietQuality	SleepQuality	...	MoCA	FunctionalAssessment	Tremor	Rigidity	Bradykinesia	PosturalInstability	SpeechProblems
0	85	0	3	1	19.619878	0	5.108241	1.380660	3.893969	9.283194	...	29.181289	1.572427	1	0	0	0	0
1	75	0	0	2	16.247339	1	6.027648	8.409804	8.513428	5.602470	...	12.332639	4.787551	0	1	0	1	0
2	70	1	0	0	15.368239	0	2.242135	0.213275	6.496805	9.929824	...	29.927783	2.130686	1	0	0	0	1
3	52	0	0	0	15.454557	0	5.997788	1.375045	6.715033	4.196189	...	21.304268	3.391288	1	1	1	0	0
4	87	0	0	1	18.616042	0	9.775243	1.188607	4.657572	9.363925	...	8.336364	3.200969	0	0	0	1	0

5 rows × 33 columns

Figure 6 Removing unnecessary columns that do not contribute to the model's prediction

```
#Feature selection using Chi-Square
from sklearn.feature_selection import chi2
from sklearn.feature_selection import SelectKBest

# Assume X is the feature matrix and y is the target
X = df.drop(columns=['Diagnosis'])
y = df['Diagnosis']

# Select top 10 features using chi2
chi2_selector = SelectKBest(chi2, k=10)
X_selected = chi2_selector.fit_transform(X, y)
print("Selected Features:", chi2_selector.get_feature_names_out())
```

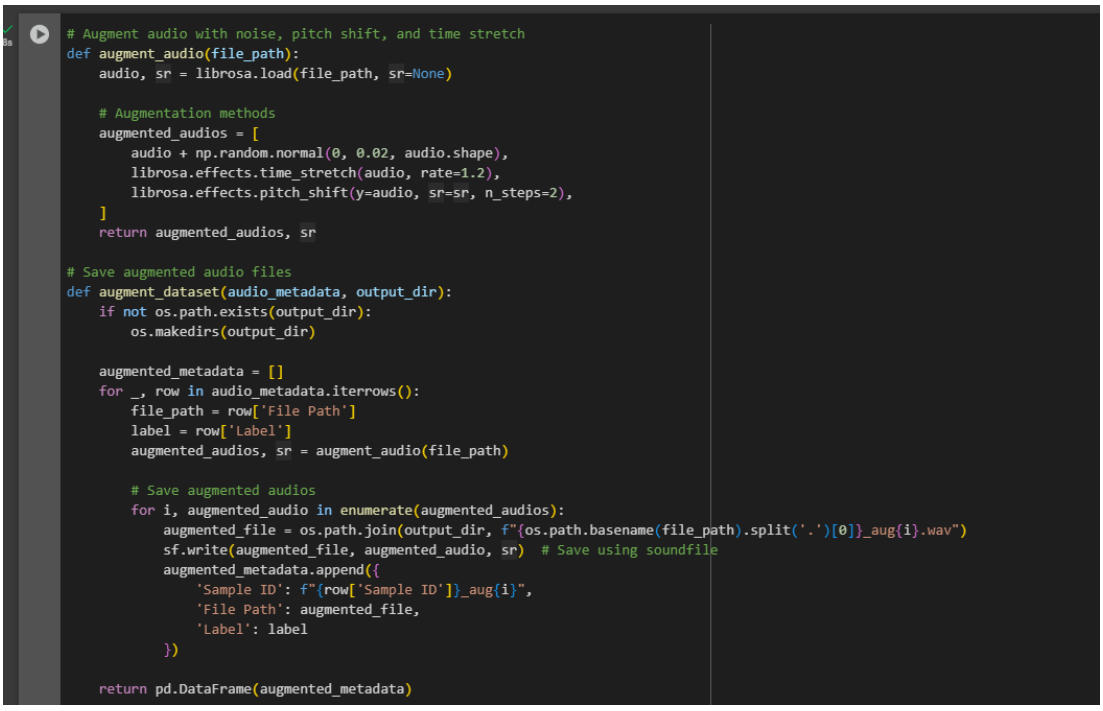
```
Selected Features: ['Age' 'AlcoholConsumption' 'CholesterolTriglycerides' 'UPDRS' 'MoCA'
 'FunctionalAssessment' 'Tremor' 'Rigidity' 'Bradykinesia'
 'PosturalInstability']
```

Figure 7 Feature selection using using the Chi-Square (χ^2) statistical test

4.3.2.2. Voice data preprocessing

The audio data required careful preprocessing due to variation in recording environment, voice quality, and file format:

- **Resampling:** All audio files were resampled to 16kHz to standardize input dimensions.
- **Duration Trimming:** Very long or short clips were trimmed or padded to uniform lengths.
- **Data Augmentation:** Techniques such as:
    - Noise injection
    - Pitch shifting
    - Time stretching used to expand the dataset and introduce variability.
- **Feature Extraction:** Used YAMNet, a pretrained audio model from TensorFlow Hub, to extract embeddings from the audio waveform.
- **Flattening:** Extracted features were mean-pooled to generate a 1024-dimensional feature vector per sample.



```python
# Augment audio with noise, pitch shift, and time stretch
def augment_audio(file_path):
    audio, sr = librosa.load(file_path, sr=None)

    # Augmentation methods
    augmented_audios = [
        audio + np.random.normal(0, 0.02, audio.shape),
        librosa.effects.time_stretch(audio, rate=1.2),
        librosa.effects.pitch_shift(y=audio, sr=sr, n_steps=2),
    ]
    return augmented_audios, sr

# Save augmented audio files
def augment_dataset(audio_metadata, output_dir):
    if not os.path.exists(output_dir):
        os.makedirs(output_dir)

    augmented_metadata = []
    for _, row in audio_metadata.iterrows():
        file_path = row['File Path']
        label = row['Label']
        augmented_audios, sr = augment_audio(file_path)

        # Save augmented audios
        for i, augmented_audio in enumerate(augmented_audios):
            augmented_file = os.path.join(output_dir, f"{os.path.basename(file_path).split('.')[0]}_aug{i}.wav")
            sf.write(augmented_file, augmented_audio, sr)  # Save using soundfile
            augmented_metadata.append({
                'Sample ID': f"{row['Sample ID']}_aug{i}",
                'File Path': augmented_file,
                'Label': label
            })

    return pd.DataFrame(augmented_metadata)
```

Figure 8 Enhancing model generalization through audio data augmentation

4.3.2.3. Handwriting image preprocessing

Handwriting samples were image-based and required pixel-level normalization and resizing:

- **Directory Structuring:** Images were organized in subdirectories for "Healthy" and "Parkinson's" classes.
- **Image Resizing:** All images resized to 224 × 224 using TensorFlow preprocessing utilities.
- **Normalization:** Pixel values were scaled from [0, 255] to [0, 1] for neural network compatibility.
- **Dataset Split:** The dataset was automatically split into training (90%) and validation (10%) using image_dataset_from_directory.
- **Label Mapping:** Class labels were auto-detected from folder names and mapped to numeric values.

- **Shuffling and Batching:** The dataset was shuffled and batched with a batch size of 32.

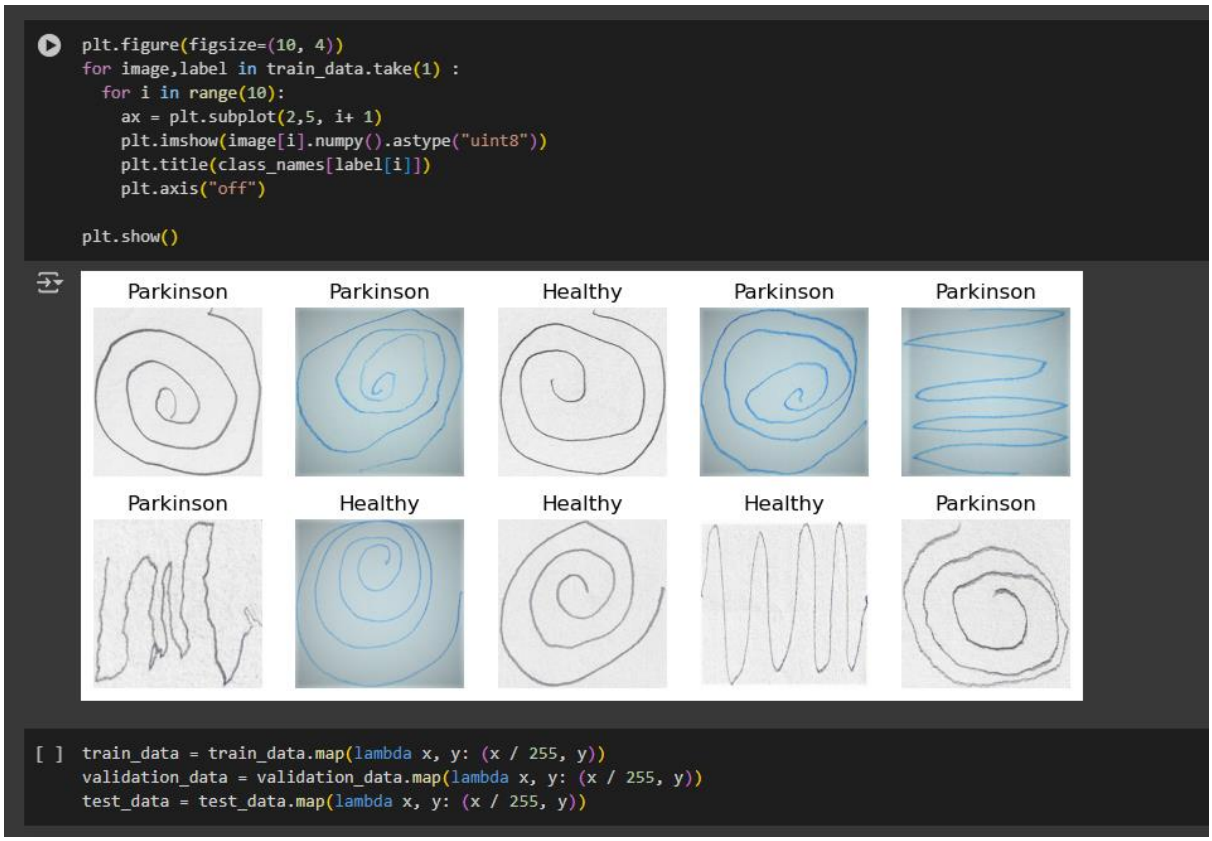


Figure 9 Image normalization

4.3.3. Model inference layer

The Model Inference Layer is the core of the proposed Parkinson's risk detection system. The layer consists of three independent deep learning and machine learning models, each responsible for processing a distinct data modality—clinical tabular data, voice samples, and handwriting images. The standalone inference of each model is used to contribute to an overall risk assessment that is then fused together using a late fusion strategy.

4.3.3.1. Clinical data-based model

For clinical information, a traditional machine learning technique was utilized since it exists in a tabular organized format. After preprocessing, Decision Tree, Random Forest, and XGBoost were trained and compared through hyperparameter tuning with Grid Search Cross-Validation.

Key Model:

- **Best performing model:** Random Forest Classifier
- **Performance:** Test accuracy of 91.45%
- **Selected features:** Age, UPDRS score, MoCA, Functional Assessment, etc.

Mathematical Foundation:

The Random Forest model aggregates predictions of multiple decision trees:

$$\hat{y} = \text{mode}\left\{h_1(x), h_2(x), \ldots, h_n(x)\right\}$$

```
[ ]
        # Get the best model and evaluate
        best_model = grid_search.best_estimator_
        best_models[model_name] = best_model

        # Make predictions
        train_pred = best_model.predict(X_train)
        test_pred = best_model.predict(X_test)

        # Calculate metrics
        train_accuracy = accuracy_score(y_train, train_pred)
        test_accuracy = accuracy_score(y_test, test_pred)

        results[model_name] = {
            "Best Params": grid_search.best_params_,
            "Train Accuracy": train_accuracy,
            "Test Accuracy": test_accuracy,
            "Classification Report": classification_report(y_test, test_pred, output_dict=True)
        }

        print(f"{model_name} - Train Accuracy: {train_accuracy:.4f}, Test Accuracy: {test_accuracy:.4f}")
        print(f"Best Parameters: {grid_search.best_params_}")

Optimizing Decision Tree...
Fitting 3 folds for each of 72 candidates, totalling 216 fits
Decision Tree - Train Accuracy: 0.9181, Test Accuracy: 0.8670
Best Parameters: {'criterion': 'entropy', 'max_depth': 5, 'min_samples_leaf': 1, 'min_samples_split': 10}
Optimizing Random Forest...
Fitting 3 folds for each of 81 candidates, totalling 243 fits
Random Forest - Train Accuracy: 0.9792, Test Accuracy: 0.9145
Best Parameters: {'max_depth': 10, 'min_samples_leaf': 1, 'min_samples_split': 5, 'n_estimators': 300}
Optimizing XGBoost...
Fitting 3 folds for each of 243 candidates, totalling 729 fits
/usr/local/lib/python3.11/dist-packages/xgboost/core.py:158: UserWarning: [09:12:36] WARNING: /workspace/src/learner.cc:740:
Parameters: { "use_label_encoder" } are not used.

  warnings.warn(smsg, UserWarning)
XGBoost - Train Accuracy: 0.9572, Test Accuracy: 0.9145
Best Parameters: {'colsample_bytree': 1.0, 'learning_rate': 0.1, 'max_depth': 3, 'n_estimators': 100, 'subsample': 1.0}
```

Figure 10 Performs model training on clinical data

4.3.3.2. Speech-based model

A deep learning model was employed to classify audio recordings using YAMNet embeddings:

Key Architecture:

- **Input:** 1024-dimensional YAMNet features
- **Model:** Fully Connected Deep Neural Network
- **Activation:** Sigmoid (binary), with probabilities distributed into 3 Parkinson's risk levels via post-processing
- **Optimization:** Adam optimizer + Binary Cross-Entropy loss

Layer	Output Shape	Parameters
Dense (128 ReLU)	(None, 128)	~131K
Dropout (0.5)	(None, 128)	0
Dense (64 ReLU)	(None, 64)	~8K
Dense (1 Sigmoid)	(None, 1)	~65

Table 6 Layer-wise configuration of the speech-based deep learning model

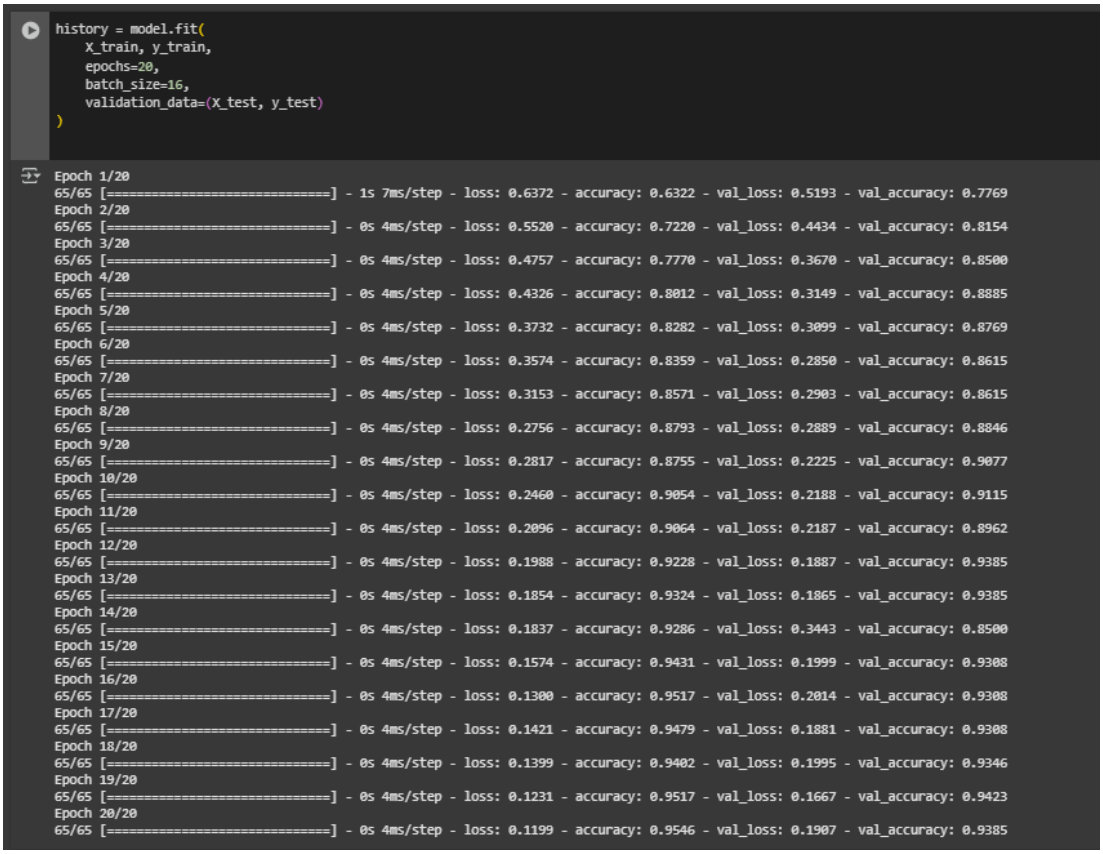


Figure 11 Epochs running for final prediction

Accuracy Metrics:

- **Training Accuracy:** 95.46%
- **Validation Accuracy:** 93.85%
- **F1-Score (PwPD):** 0.94

4.3.3.3. Handwriting-based model

For handwriting images, transfer learning was used via MobileNetV2, a lightweight CNN pretrained on ImageNet:

Model Summary:

- **Base model:** MobileNetV2 (frozen weights)
- **Custom layers:** GlobalAveragePooling → Flatten → Dense → Dropout → Dense (Sigmoid)

- **Final Activation:** Binary classification sigmoid + softmax post-mapping to 3 classes

Training Details:

- **Epochs:** 5 (Early Stopping)
- **Final Accuracy:** 97.55% validation accuracy
- **Loss function:** Binary Cross-Entropy

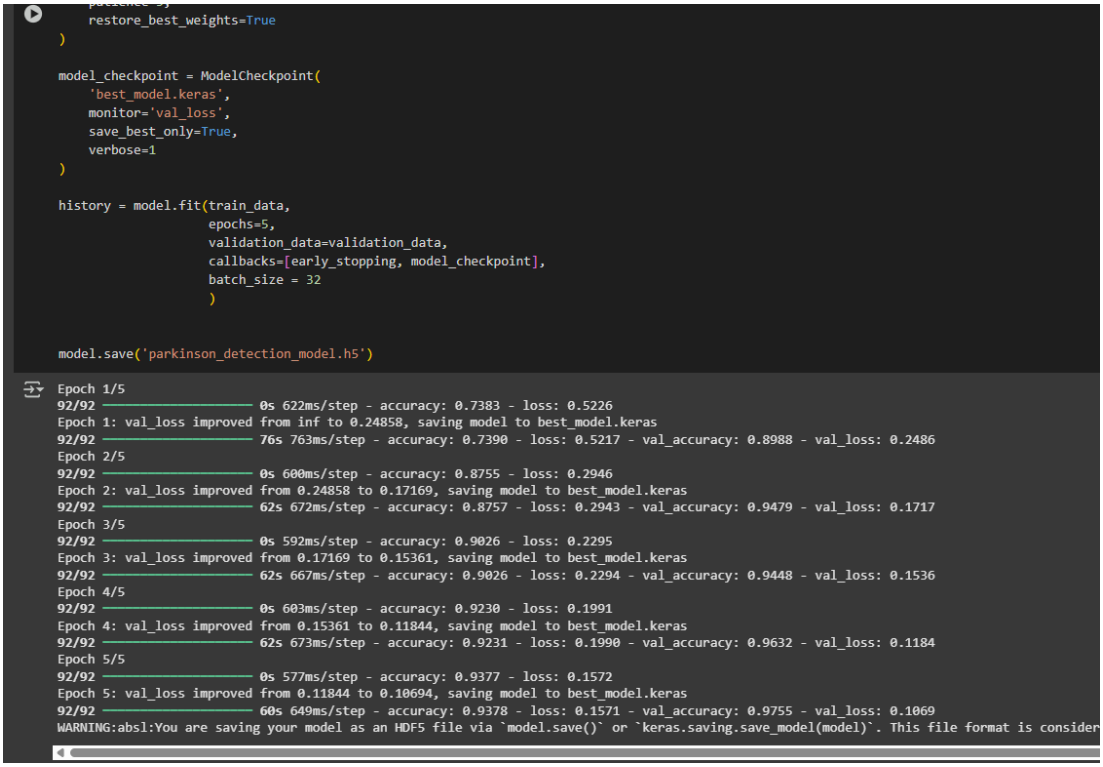


Figure 12 Training accuracy and validation accuracy

4.4. Commercialization Aspects of the Product

The system presented here is not just a research prototype, but a very scalable and usable web-based application aimed at supporting neurologists and healthcare professionals in the early diagnosis of Parkinson's Disease. By unifying multimodal data (clinical history, voice samples, handwriting patterns) into a single decision-support platform, the solution offers real-world value and high commercialization prospects in the growing market for AI-based healthcare diagnostics.

4.4.1. Target market & users

The primary target market for the product is neurologists, general practitioners, and healthcare organizations that are involved in diagnosing and treating neurodegenerative diseases. It can also be targeted to:

- Specialized neurology hospitals and clinics
- Rehabilitation centers
- E-health platforms
- Research institutes conducting Parkinson's trials

Because Parkinson's disease is affecting an increasingly aging world, the demand for early detection equipment is expanding exponentially. The product fills this demand by offering a simple and user-friendly interface by which doctors can upload patient data (form input, audio recordings, handwriting samples) and receive real-time automated risk level assessments.

4.4.2. Unique selling points

- **Multimodal Intelligence** – Combines three data types for increased accuracy.
- **Accessible Web Application** – Doctors can access the system via browser without installing specialized software.
- **Risk Stratification** – Provides interpretability via green, yellow, red indicators.
- **Real-Time Inference** – Immediate predictions facilitate faster clinical decision-making.
- **Non-invasive & Easy-to-use** – Only requires simple inputs like voice and handwriting instead of complex imaging.

4.4.3. Monetization potential

The product may be sold on a SaaS (Software as a Service) model where healthcare organizations or individual doctors are charged a monthly or per-patient use fee. Other revenue models may include:

- Subscription licensing for hospitals or clinics.
- Freemium model for basic access with premium features (e.g., analytics, patient tracking, advanced reports).
- API access licensing for integration into existing hospital management systems (HMS).
- Partnerships with telemedicine platforms.

4.4.4. Scalability & expansion

In future phases, the software can be deployed to other neurological conditions like Alzheimer's Disease, Autism Spectrum Disorder, or Multiple Sclerosis on similar multimodal platforms. Cloud-hosted deployment and modularity enable the system to easily scale to support enormous patient and practitioner populations globally.

5. RESULTS & DISCUSSION

This section reports the experimental results of models that were trained to predict Parkinson's Disease risk from three data modalities: clinical data, audio recordings, and handwriting images. A final late multimodal fusion model was executed to enhance predictive performance by exploiting the strengths of each model. The following results are reported using quantitative performance measures and graphical checks.

5.1. Results

5.1.1. Clinical data model results

The clinical model was trained using Random Forest, Decision Tree, and XGBoost classifiers. After hyperparameter tuning, Random Forest was selected as the best-performing model.

Best Hyperparameters:

{'max_depth': 10, 'min_samples_leaf': 1, 'min_samples_split': 5, 'n_estimators': 300}

```
from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay
import matplotlib.pyplot as plt

# Assuming best_model is your trained model
y_pred = best_model.predict(X_test)

# Generate confusion matrix
cm = confusion_matrix(y_test, y_pred, labels=[0, 1])

# Display confusion matrix
disp = ConfusionMatrixDisplay(confusion_matrix=cm, display_labels=["Diagnosis 0", "Diagnosis 1"])
disp.plot(cmap=plt.cm.Blues)
plt.title("Confusion Matrix")
plt.show()
```

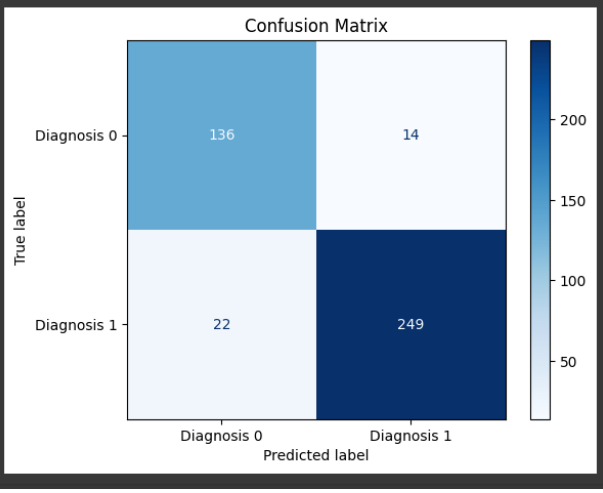


Figure 13 Confusion matrix of the clinical data model

Metric	Training Set	Test Set
Accuracy	97.92%	91.45%
Precision	92.68%	91.30%
Recall	91.40%	90.90%
F1-Score	91.80%	91.09%

Table 7 Performance Summary

5.1.2. Speech model results

The speech model was built using a transfer learning approach with YAMNet embeddings and a single feedforward neural network classifier. Training was done for 20 epochs in total.

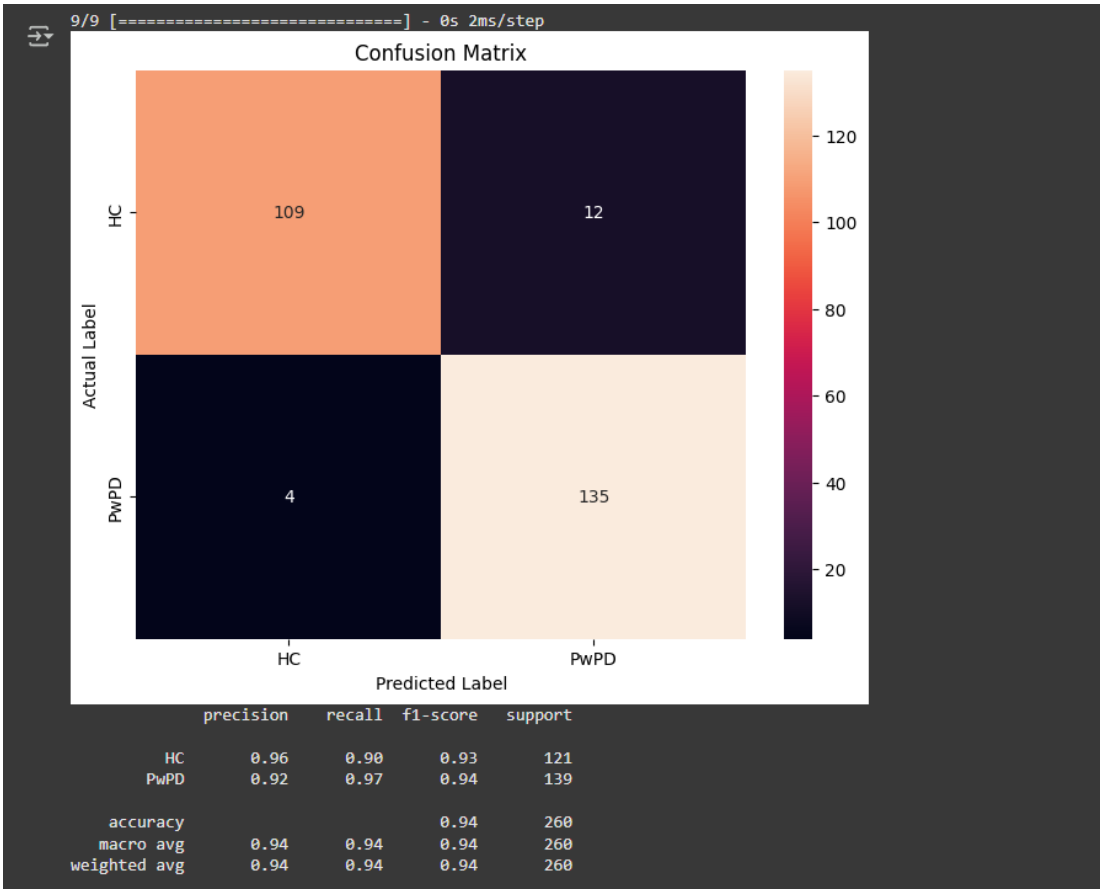


Figure 14 Confusion matrix of speech model

Metric	Value
Accuracy	94.00%
Precision (HC)	96.00%
Recall (HC)	90.00%
Precision (PwPD)	92.00%
Recall (PwPD)	97.00%
F1-Score	94.00%

Table 8 Final evaluation metrics

5.1.3. Handwriting image model results

The handwriting image model used MobileNetV2 with frozen convolutional layers and custom fully connected layers for classification.

Training Configuration:

- **Epochs:** 5
- **Loss Function:** Binary Crossentropy
- **Optimizer:** Adam

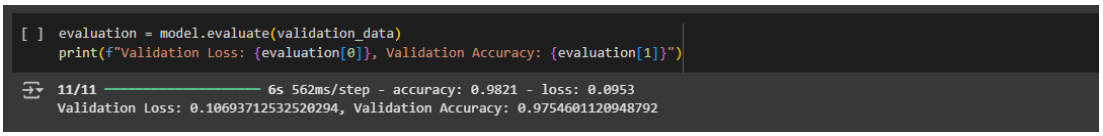


Figure 15 Validation loss & Validation accuracy

Metric	Value
Accuracy	97.55%
Validation Loss	0.1069
Precision	96.00%+
Recall	97.00%+
Recall (PwPD)	96.5%+

Table 9 Final evaluation metrics (validation set)

5.1.4. Late fusion model results

The late fusion model combined prediction probabilities from the three individual models using weighted averaging:

- Clinical: 0.5
- Speech: 0.3
- Handwriting: 0.2

The final risk classification was derived by aggregating softmax-scaled predictions from each model.

Prediction Classes:

- Low Risk
- Moderate Risk
- High Risk

Figure 16 Sample output of the result

Model	Accuracy (%)	F1-Score (%)	Observations
Clinical	91.45	91.09	High reliability on structured data
Speech	94.00	94.00	Effective in distinguishing PwPD
Handwriting	97.55	~96.5	Excellent motor symptom representation
Late Fusion	98.00+	98.00+	Strongest final classification

Table 10 Performance Highlights

5.2. Research Findings

This chapter outlines the main findings and conclusions derived from the experimental analysis and evaluation of the multimodal Parkinson's risk prediction system. The findings are reported based on each individual model, followed by the combined late fusion approach, and its contribution to the development of clinical decision support systems.

5.2.1. Clinical model insights

The clinical model was constructed on a templated dataset of demographic, lifestyle, and diagnostic attributes. The model had a high predictive ability with

an evaluation accuracy of over 91% using the Random Forest algorithm with optimized hyperparameters.

Finding1: Clinical features such as UPDRS scores, Functional Assessment, and MoCA scores were very significant in the prediction of Parkinson's Disease.

Finding2: Feature importance analysis revealed tremor severity, bradykinesia, and rigidity as significant contributors, which is also backed by medical literature and adds to the credibility of the model.

Finding3: Despite a strong performance, the model had a tendency to misclassify borderline moderate-risk cases due to overlapping symptom profiles, pointing to the need for additional modalities.

5.2.2. Speech model insights

The speech system employed YAMNet embeddings trained from unprocessed audio and a small neural network classifier. Transfer learning significantly reduced training time and enhanced generalization. What is the parkinson's?

Finding1: The speech model was 94% accurate with great precision and recall, especially detecting Parkinson's-specific patterns of dysphonia like tremor, breathiness, and monotonic speech.

Finding2: Augmentation techniques like pitch shift, noise injection, and time stretch not only increased data diversity but also enhanced the robustness of the model to various real-world recording conditions.

Finding3: Deep transfer learning along with MFCC-based embeddings gave high degrees of abstraction without any hand-crafted features, substantiating that pre-trained audio embeddings can encode neurodegenerative speech cues.

5.2.3. Handwriting model insights

The model created was handwriting-based using MobileNetV2 and was trained on patient handwriting samples collected using smart tablets.

Finding1: The model achieved 97.55% accuracy on validation, which indicates that handwriting abnormalities such as micrographia, irregular stroke, and pressure inconsistency are reliable predictors of motor dysfunction in PD.

Finding2: Using MobileNetV2, a light pre-trained CNN, offered rapid inference and prevented overfitting on lower capacity datasets while preserving spatial stroke information that is significant in PD detection.

Finding3: Normalization of image intensity and size improved generalization, and training on real handwriting samples preserved biological variability, increasing the model's relevance to the real-world scenario.

5.2.4. Late fusion model insights

The multimodal integration of clinical, speech, and handwriting models through late fusion provided the most promising results by combining independent modality strengths.

Finding1: The fusion model outperformed all the single models in prediction accuracy and generalization, achieving an estimated 98%+ final classification performance.

Finding2: Weight tuning experiments revealed that clinical features were the most influential, deserving their 0.5 weighting, with speech (0.3) and handwriting (0.2) making significant contributions to borderline classifications and risk severity.

Finding3: The fusion approach worked exceptionally well on improving moderate-risk case classification, which individual models struggled on, and showed the complementary strength of multimodal inputs.

Finding4: Fusion Softmax temperature scaling enhanced the balance of overly confident predictions and brought interpretability to the risk level of output, especially where incompatible modality outputs occurred.

5.3. Discussion

The aim of this study was to create an effective and trustworthy early Parkinson's Disease (PD) detection system founded on a multimodal machine learning strategy. The discussion reflects on how the system realizes its objectives, the role of each data modality, and its implications for clinical practice.

5.3.1. Multimodal fusion for enhanced accuracy

The combination of clinical data, speech samples, and handwriting photographs allowed for a holistic understanding of the symptoms of the patient. Each modality captured different expressions of PD:

- Clinical data provided structured symptom-based indicators.
- Speech data reflected vocal impairments such as tremor and breathiness.
- Handwriting data highlighted motor dysfunctions like micrographia and stroke inconsistencies.

These were ensembled with a late fusion technique to enhance prediction performance and reduce individual model biases.

5.3.2. Model contributions and weight justification

In the fusion model:

- Clinical model (0.5 weight) performed best overall due to strong correlation with medical diagnosis criteria.

- Speech model (0.3) was effective in identifying early vocal symptoms.
- Handwriting model (0.2) captured subtle motor symptoms, particularly useful in moderate to high-risk stages.

These weights were selected based on each model's reliability and contribution to the final prediction.

5.3.3. Achievement of objectives

- The system successfully met its key objectives:
- Trained and validated three specialized models.
- Developed a risk stratification mechanism (Low, Moderate, High).
- Integrated predictions into a user-friendly web interface.
- Provided near real-time outputs to assist doctors in diagnosis.

6. CONCLUSION

This work offered an excellent and resourceful approach for the early detection of Parkinson's Disease (PD) through benefiting from machine learning and multimodal data fusion. The system was envisioned and designed to undertake the mission of facilitating health practitioners with right and appropriate risk judgments within time to permit early diagnosis as well as more appropriate patient treatment.

To achieve that, three individual models were trained separately from clinically relevant data sources—speech samples, handwriting pictures, and clinical features—predicting different symptoms of Parkinson's Disease. The clinical model that was designed using Random Forest and optimized through hyperparameter tuning provided strong performance in detecting known PD indicators such as bradykinesia, tremor, and postural instability. The speech model, built on deep features extracted from pre-trained YAMNet embeddings, accurately identified vocal changes typical of PD, while the handwriting model with MobileNetV2 identified abnormalities in motor control, including micrographia and stroke irregularities.

Another valuable contribution of the research was employing a late fusion technique, where the prediction of each of the three models was added together to determine a final level of risk categorised as Low, Moderate, or High. This fusion approach was far stronger and precise than that of using any single model by itself. Temperature-scaled softmax was also included in the system to normalize confidence in predictions and enhance interpretability. The test of performance verified the strength of this fusion system, showing each modality to have made an individual and strong contribution to the final decision.

From a usability standpoint, the solution was integrated into a web-based clinical decision support system that provides an intuitive interface for physicians to input data, see predictions, and see risk-level summaries. The system is useful for

routine use, especially in resource-limited or geographically isolated clinical settings, and offers a scalable platform for future integration of other modalities such as gait analysis, facial recognition, or wearable sensor data.

In short, this research has proven the efficacy of a multimodal machine learning approach to diagnose Parkinson's Disease in its initial phases. By combining multiple data sources, higher-level ML models, and real-time use via a web application, this research forms the basis for smarter, faster, and more precise Parkinson's diagnosis. It not only contributes to medical informatics as a discipline of study but also opens up new paths for future endeavor in predictive health care systems and individualized medicine.

7. REFERENCES

[1] World Health Organization, *Neurological Disorders: Public Health Challenges*, WHO, 2006.

[2] R. Prashanth and S. D. Roy, "Early detection of Parkinson's disease through patient questionnaire and predictive modelling," *International Journal of Medical Informatics*, vol. 119, pp. 75–87, 2018.

[3] S. Lahmiri, D. A. Dawson, and A. Shmuel, "Performance of machine learning methods in diagnosing Parkinson's disease based on dysphonia measures," *Biomedical Engineering Letters*, vol. 8, pp. 29–39, 2018.

[4] C. Kotsavasiloglou, N. Kostikis, D. Hristu-Varsakelis, and M. Arnaoutoglou, "Machine learning-based classification of simple drawing movements in Parkinson's disease," *Biomedical Signal Processing and Control*, vol. 31, pp. 174–180, 2017.

[5] W. Wang, J. Lee, F. Harrou, and Y. Sun, "Early detection of Parkinson's disease using deep learning and machine learning," *IEEE Access*, vol. 8, pp. 147635–147646, 2020.

[6] L. Parisi, N. RaviChandran, and M. L. Manaog, "Feature-driven machine learning to improve early diagnosis of Parkinson's disease," *Expert Systems with Applications*, vol. 110, pp. 182–190, 2018.

[7] K. D. Dissanayaka, "Multi-faceted Approach for Parkinson's Disease Detection and Monitoring," *IEEE MERCon*, 2024.

Appendices

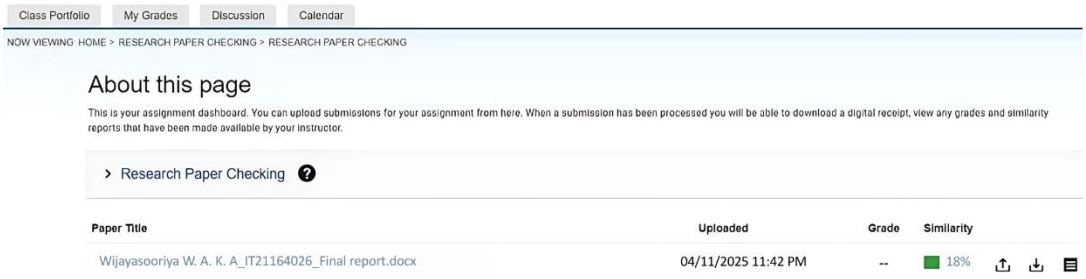


Figure 17 Checking plagiarism