A Framework for the Prediction of Parkinson's Disease Using Agentic Artificial Intelligence

Areej Fatima[1], Rabia Tehseen[2], Kashif Nasr[2], Uzma Omer[3], Muhammad Inam Ul Haq[4], Rubab Javaid[5]
[1]Department of Data Science, University of Central Punjab, Lahore, Pakistan
[2]Department of Computer Science, University of Central Punjab, Lahore, Pakistan
[3]Department of Computer Science, University of Education, Lahore, Pakistan
[4]Department of Computer Science and Bioinformatics, Khushal Khan Khattak University, Karak, Pakistan
[5]Department of Software Engineering, University of Central Punjab, Lahore, Pakistan
*Correspondence: rabia.tehseen@ucp.edu.pk

Citation | Fatima. A, Tehseen. R, Nasr. K, Omer. U, Haq. M. I. U, Javaid. R, "A Framework for the Prediction of Parkinson's Disease Using Agentic Artificial Intelligence", IJIST, Vol. 7, Issue. 4 pp 3033-3047, December 2025
Received| November 03, 2025 **Revised|** November 19, 2025 **Accepted|** November 24, 2025 **Published|** December 02, 2025.

Parkinson's disease (PD) is a progressive neurodegenerative disorder that is difficult to diagnose, particularly in its early stages. Subtle, slowly evolving symptoms often delay confirmation, reducing opportunities for timely intervention that could improve outcomes and quality of life. Conventional diagnosis relies largely on clinical observation, which can be subjective and insufficiently sensitive for early detection. This thesis proposes an Agentic Artificial Intelligence (AAI) framework for early PD detection and severity assessment using voice-based biomarkers. Biomedical voice parameters are leveraged because vocal changes can reflect early neurological impairment. Two publicly available Kaggle datasets containing voice recordings from individuals with PD and healthy controls are used to train and evaluate the models. For detection, an XGBoost classifier achieves 94.68% accuracy with strong discriminative performance. For severity estimation, XGBoost regression models predict motor and total Unified Parkinson's Disease Rating Scale (UPDRS) scores with high agreement to clinically reported measurements. A key contribution is an agentic decision-making layer that autonomously interprets model outputs, performs disease staging, and generates stage-dependent monitoring and treatment recommendations. Unlike conventional predictive pipelines that stop at numerical outputs, the proposed system translates predictions into actionable clinical insights to support structured decision-making. Experimental results indicate that the framework can detect PD and estimate severity effectively from non-invasive voice data, highlighting the potential of AAI for earlier diagnosis, personalized monitoring, and intelligent clinical decision support in healthcare. The multi-layer design supports modular updates to models and agent policies, enabling telehealth deployment and longitudinal tracking as additional voice samples become available over time.

Keywords: Parkinson's Disease; Agentic Artificial Intelligence; Machine Learning; Voice Biomarkers; Early Detection; Clinical Decision Support.

Introduction:

Parkinson's disease (PD) is a chronic nervous disorder that affects movement and coordination [1][2][3]. It arises because of the dislocation of the dopamine-producing neurons of the brain, which in turn causes tremors, muscle rigidity, movement retardation, and balance difficulties [2]. Among other motor symptoms, depression, sleeping disorders, and memory problems are also prevalent in many patients [2][3]. PD is a fast-growing neurological disorder prevalent globally and a great health care drain [4]. Quick treatment requires early and timely diagnosis to manage the symptoms of the disease and live a healthy life [3]. But the current diagnostic techniques are based primarily on clinical observation, which is subjective and has the potential to slow down early diagnosis [4]. Therefore, it results in a more urgent need for automated and smart frameworks that will help doctors in the early detection of PD using quantifiable data [3][4]. The Agentic Artificial Intelligence (AAI), capable of acting independently, making decisions, and learning about the surrounding environment, is a new direction in medical diagnosis. Unlike conventional AI, which only carries out pre-defined procedures, Agentic AI systems can engage with users, think about the results, and react to the emerging data patterns by taking on autonomous decisions [1]. This research uses the strength of AAI to develop a reasonable model for predicting Parkinson's disease of Parkinson and classifying the stages of the disease. Traditional diagnostic criteria of PD consider more clinical experience and observation, which is likely to lead to late or misdiagnosis. The machine learning (ML) models demonstrated the potential of PD detection by voice, gait, or handwriting data, but they are still rigid, less open, and do not interact with patients [3][4]. According to these constraints, the main goals of this research are the following:(1) To create a machine learning-based framework to recognize early signs of Parkinson's disease through biomedical voice. (2) To categorize the severity of Parkinson's disease by staging the disease using predictive modeling. (3) To add an Agentic Artificial Intelligence layer that autonomously orchestrates detection, severity estimation, staging, and recommendation generation using transparent decision logic. (4) To develop an interactive AI layer of communication offering understandable personalized feedback to patients. (5) To increase the robustness and interpretability of the models through the combination of various open-source datasets on Parkinson's Disease. The paper is dedicated to the collection of data, its pre-processing, the training of the model, and the implementation of an Agentic AI communication layer. The study is not grounded in actual life clinical experiments but uses open-source Kaggle datasets. The given work is important. After all, it can potentially help healthcare professionals, as it would allow faster and impartial diagnosis. Clear and personal feedback given on the health condition of the patient is also a way of empowering the patient. These systems may be incorporated into the telehealth systems to enable continuous and real-time monitoring of PD. The originality of this work is in the integration of Agentic AI into the existing procedures of machine learning, where prediction accuracy is replaced with interaction and adaptability, and patient-centered communication. The proposed solution combines several datasets of Parkinson's Disease and uses a step-by-step predictive framework to increase the accuracy and readability, which is appropriate to be used in real healthcare environments.

Literature Review:

The proposed study is dedicated to creating A Framework to predict the development of Parkinson's Disease with the help of the Agentic AI in order to improve early diagnosis, proper classification of the stage, and provide an in-depth explanation of clinical decision support. The past models of AI and machine learning (ML) have demonstrated promising performances in Parkinson's disease (PD) detection, but in most cases, they lack scalability, interpretability, and patient-centered assimilation. This study aims to fill these gaps with the help of agentic AIs, which autonomously and self-optimizing frameworks combine

multimedia information and ethical guidelines and finally offer a more credible and transparent diagnostic system [1][2].

Machine Learning Methods of PD Detection:

PD diagnosis has been extensively used in machine learning using speech, gait, and neuroimaging data. [3] Also conducted a thorough review of ML techniques with high-detection accuracy [4], and found the inability to generalize the datasets and the inability to interpret the results to be persistent issues.[5] indicated the possibilities of smartphone-based eye, voice, and skin information in continuous monitoring. These methods are supervised learning, domain adaptation [6], and multimodal fusion [7]. Findings indicate the enhanced early diagnosis and stage forecast, yet there are drawbacks in the population heterogeneity and cross-cohort validation [8][9] highlights how Agentic AI systems enable personalized medicine by autonomously adapting treatment pathways based on patient-specific data. Recent speech-based (20232025) studies have also enhanced Parkinson's disease detection using deep learning and domain-adaptive models. Islam et al. [10] showed that AI-based speech and behavioral signals are very reliable estimators of at-home PD severity, which supports the idea of remote monitoring. By analogy, Ibarra et al. [6] suggested methods of domain-adaptation in order to enhance generalization between multilingual speech corpora to overcome an important limitation of previous voice-based models. Also presented are multi-cohort methods of learning, which aim at eliminating dataset bias to increase robustness in the real-world speech-based prediction of PD [8].

Explicable and Agentic AI Frameworks:

XAI (explainable AI) is essential to clinical adoption [11]. A XAI framework was created that combines multimodal data on ML to improve interpretability. A HIPAA-compliant agentic AI system based on reinforcement learning, adaptive algorithms, and secure data handling. Findings show that it is more reliable, transparent, and ethically aligned, but requires large datasets of high quality and complexity in calculations, which is a weakness [12]. [12] Provide a systematic review of machine learning and explainable AI techniques used for predicting Parkinson's disease. [13] explores how soft robotics integrated with functional materials can support Parkinson's disease management. Recent explainable and agentic AI investigations focus on providing transparency in making decisions in neurodegenerative diseases. The explainable speech-aware ML models on the diagnosis of Parkinson proposed by Priyadharshini et al. [11] aid in enhancing the trust that clinicians have. Closed to explainability, Ndlovu et al. [12] also emphasized that explainability is essential to speech-driven PD prediction systems and that interpretable voice biomarkers are essential to clinical adoption. These results favor the combination of Agentic AI layers of adaptive learning and patient-centered interaction of speech-based diagnostic systems.

Behavioral, psychosocial, and Motor Features:

Prediction and monitoring are improved as a result of patient-centered features. [14] [15] focused on psychosocial issues and the benefits of therapy.[16] presented AI-based conversational journaling to capture symptoms in the real world. EXAI models used with spiral/wave drawings are used in early recognition of motor impairments [17]. These methods involve longitudinal monitoring, surveys, and non-invasive motor assessment. Findings show a previous identification of the symptoms, though its findings are limited by the variability of behavior information and patient compliance [18]. Recently, there has been an interest in speech-based patient response as a non-invasive behavioral measure. Rashik et al. [16] have proposed AI-enabled conversational journaling to record voice and changes in symptoms in everyday life and allow ongoing monitoring of speech. These studies reveal the way in which verbal and nonverbal signals can be used to supplement conventional motor tests in the early diagnosis of Parkinson's disease.

Imaging, Radiomics and Genetic Insights:

Clinical data is supplemented by advanced imaging and genetic analyses. [19] applied cross-regional radiomics to subtype the motor, whereas [20][21] used less predominant genetic variants to determine the risk of PD. They include MRI, extraction of radiomic features, and whole-exome sequencing. Results indicate more effective subtyping and risk stratification, but resource intensity and population biases are weaknesses [22].

Rehabilitation and Clinical Integration:

The combination of AI with clinical practice enhances patient care. [23][24] emphasize multidisciplinary care and incorporating patient input. [25] talked about AI replacing diagnostic tools with clinical collaborators. They could be clinical trials, consensus statements, or online monitoring systems. The findings show better personalization of treatment and patient involvement; constraints are related to scale in the real world [26][27]. Recent researchers point to the fact that speech monitoring systems based on AI can be successfully implemented into the telehealth and rehabilitation processes. There is evidence of success with digital health platforms that use speech and behavioral analytics to enhance patient engagement and clinician feedback loops, especially in long-term neurological diseases, such as Parkinson's disease [10][25]. This form of integration facilitates the transformation of the speech-based AI systems into instruments of experimentation for clinical, actionable decision-support systems.

Methodology:

This paper will present a five-layer Agentic Artificial Intelligence (AAI) system that will help identify, classify the severity, and manage Parkinson's disease (PD) in a patient-centered manner. The framework combines a machine learning prediction model with an agentic decision-making layer that autonomously handles diagnostic flow and outputs that are clinically meaningful. The suggested system is based on a pipeline that is sequential and follows the order of the acquisition of patient data, preprocessing, disease detection, severity estimation, and decision support. The overall methodology of the research is illustrated in Figure 1.

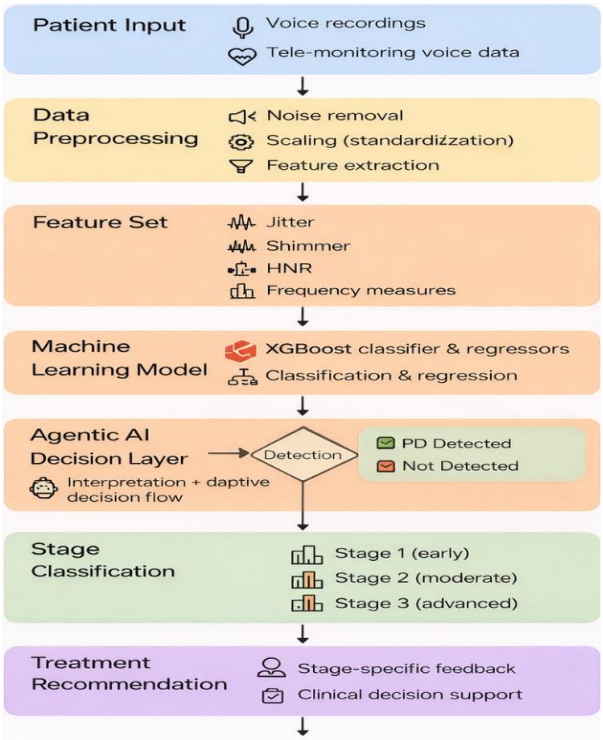

Figure 1. Methodology Diagram

Agentic Artificial Intelligence Overview:

Agentic Artificial Intelligence is defined as a type of intelligent system that, unlike simply creating passive predictions, is engaged in the decision-making procedures. In contrast to traditional machine learning models, which give out probability scores as output, an agentic AI system is able to decipher outcomes, decide on future steps, and assist the process of workflow automation. The component of Agentic AI in the context of this research interprets processed patient voice data to decide on the presence of Parkinson's disease. Depending on the result, it takes an independent initiative to proceed to further diagnostic measures like estimation and classification of disease severity. The given decision-driven behavior allows making the system appear as a digital clinical assistant that minimizes human intervention and enhances diagnostic efficiency.

Patient Input:

The initial phase of the model is the procurement of raw patient information. The current research mainly employs biomedical voice recordings of patients with Parkinson's Disease and healthy control patients. Voice samples also record the important vocal biomarkers like the variation in pitch, tremor, amplitude variability, jitter, and shimmer, which are known to provide evidence of an impaired neurology in Parkinson Disease. Moreover, voice data that is tele-monitored during several sessions is included to examine the trends of progression of the diseases. Quality and consistency of patient input data are crucial because the resulting machine learning performance greatly relies on quality data. All the inputs are deposited in the preprocessing phase, where they are cleaned and standardized, and then subjected to analysis.

Data Preprocessing:

Preprocessing of data is a very important aspect to be taken to make the data reliable and consistent. The raw biomedical data can have noise, missing values, or inconsistencies; hence, a number of preprocessing methods were used. The missing values have been treated with the mean imputation technique, where the missing value has been averaged with the average value of the respective feature. The reason why mean imputation was chosen is its simplicity and numerical stability when dealing with continuous biomedical voice features with a small percentage of missing values. This method maintains the overall trend of the data without adding some unnatural variability and is often used in machine learning pipelines of biomedical applications, especially when used along with feature scaling methods like standard normalization. The data with only missing values in a feature were dropped so that they did not cause instability when training. Then, the Standard Scaler technique was used to scale features so that all features are of zero mean and unit variance. This measure eliminates the chances of features with a larger range dominating the learning process. The methods of feature extraction were used to transform raw voice signals into numerical indicators that are meaningful, like jitter, shimmer, harmonic-to-noise ratio (HNR), noise-to-harmonic ratio (NHR), and measures of fundamental frequency. Preprocessing also removed noise and outliers that were likely to adversely affect model performance.

Agentic AI Layer:

The main part of the proposed framework, which is the intelligence, is the Agentic AI layer. It uses machine learning models on the pre-prepared data to predict the occurrence of Parkinson's Disease. This layer is an independent decision-making entity as opposed to traditional systems, which can only produce classification scores. The agent compares features that are extracted and determines abnormal voice patterns that relate to Parkinson Disease. In accordance with this assessment, the system works out the following classifications of each case: Parkinson Disease Detected and Not Detected. In case Parkinson's disease has been detected, the agent automatically triggers the severity calculation module and does not need a human action. After estimating the severity of the disease, the Agentic AI reads the results and

prepares stage-specific information for patients and clinicians. Explainability, confidence, and continuity in AI-assisted medical decision-making are guaranteed by this layer. Future developments of this framework can also incorporate real-time wearable sensor data and telehealth systems of adaptive monitoring.

Output and Decision Flow:

At this phase, the system gives a binary diagnostic decision whether one has Parkinson's disease or not. In case the condition has not been identified, the process will end with a suggestion of periodic monitoring. In case of the detection of Parkinson's disease, the case is automatically referred to the severity classification module. A score of confidence can also be included in the decision output, and it allows clinicians to have a clue about the reliability of the prediction made by the system. These outputs may be presented in the form of clinical dashboards, mobile apps, or reporting tools, and the intricate AI analysis is converted to comprehensible diagnostic data.

Disease Stage Classification:

After identifying the disease, the framework categorizes the severity of Parkinson's disease into three stages:

Stage 1 (Early)
Stage 2 (Moderate)
Stage 3 (Advanced)

It is categorized according to the predicted severity indicators that were based on voice-related features. The initial cases of the disease are characterized by slight vocal instability, and the advanced ones are marked by serious degradation of the voice and lack of motor coordination. Such a staging process has clinical significance, as a treatment strategy greatly depends on the severity of the disease. Proper classification of the stage helps clinicians to develop a specific treatment and provides an opportunity to monitor the disease progression over a long period.

Treatment Recommendation:

At the last step, the Agentic AI system will create stage-specific treatment suggestions. These recommendations are not prescriptive, and they are designed to aid clinical decision-making. Early-stage patients can be recommended lifestyle changes, voice therapy, physiotherapy, and regular check-ups.

Patients at the moderate stage can be provided with instructions connected with medication administration, other therapy programs, and assistance with mobility.

Patients with advanced stages might need expert care, increased treatment approaches, and follow-ups.

This stage aims at coming up with interpretable and actionable recommendations to empower patients and health care professionals. This component may be deployed alongside hospital systems or digital health platforms in the real world.

Dataset Description:

In this paper, two datasets retrieved from Kaggle, which are publicly available and both address biomedical voice measurements related to Parkinson's Disease, will be used:

Two open-source data sets combined:

Parkinson's Disease Data (Kaggle – Parkinson's Disease Dataset).

Parkinson Tele-Monitoring Dataset (Kaggle – Parkinson's Tele-Monitoring Dataset).

The Parkinson's Disease Dataset is the first dataset that is a collection of voice recordings taken of people with Parkinson's disease and healthy control participants. This dataset is most commonly employed in the binary classification task, due to the ability to obtain vocal biomarkers that can differentiate patients with Parkinson's disease from non-weakened persons. The second dataset, Parkinson Telemonitoring Data, consists of longitudinal voice recordings that were only recorded by patients with Parkinson's. Along with voice-related data,

such indicators of clinical severity as motor Unified Parkinson's Disease Rating Scale (UPDRS) and total UPDRS scores are included in this dataset and are needed to estimate the disease severity and analyze its progression. Before integration of the datasets, redundant identifiers such as names of subjects, patient numbers, and time metadata were eliminated to eliminate bias and redundancy. The feature names have been standardized so as to have similarities in the two datasets. Row-wise concatenation to merge the datasets was then done. The telemonitoring data set on the status variable has missing values, which were filled with a 1, since all the values were related to Parkinson's disease patients. The resulting consolidated data has 7,070 records and 29 numerical variables, such as fundamental frequency values, jitter, shimmer, harmonic-noise ratio (HNR), and noise-harmonic ratio (NHR). The use of a unique patient identifier was instituted to identify different records. The last-generated data were stored as newcombined_parkinsons_data_final.csv and utilized in all further experiments. Dataset detail has been mentioned in Table 1.

Table 1. Dataset Description

Ref.	Dataset Title	Description	Size
[28]	**Parkinson's Disease Dataset**	This dataset contains a range of biomedical voice measurements from individuals with Parkinson's disease and healthy controls. The data is primarily designed for machine learning models that aim to detect Parkinson's disease from vocal biomarkers. Number of Records (1,195 records)	~507.55 KB (CSV)
[29]	**Parkinson's Tele-Monitoring Dataset**	A dataset of biomedical voice measurements from Parkinson's patients recorded over multiple sessions, used for monitoring the progression of the disease and treatment effects. It includes UPDRS (Unified Parkinson's Disease Rating Scale) scores for regression modeling. Number of Records (5,875 records)	~911.26 KB (CSV)

Dataset Harmonization, Integration, and Risk Analysis:

A systematic harmonization procedure was carried out before integrating the datasets to make sure that features are consistent and bias is reduced. The Parkinson's Disease Dataset has feature names that were standardized by removing prefixes in the dataset (e.g., MDVP) and matching the vocal biomarker nomenclature with the Parkinson Tele-Monitoring dataset. The problem of redundant identifiers like subject names, patient indexes, and time-related metadata was eliminated to eliminate subject-level bias. Row-wise concatenation of the Parkinson-based Disease Dataset and the Parkinson Tele-Monitoring Dataset enabled both heterogeneous samples to co-exist but maintain feature diversity. Columns that were not found in any dataset were automatically matched and assumed a missing value, on which further imputation would be done. Seeing that the telemonitoring data only has the data of patients with Parkinson's disease, the missing labels of classification were clinically imputed to Parkinson-positive to ensure harmonization of the labels. All the records were reassigned a new unique patient identifier to guarantee record-level traceability without data leakage. The motivation behind the combination of the two datasets was to come up with a single learning model that would be able to complete the task of detecting and estimating the severity of Parkinson's disease using vocal biomarkers in a complementary manner. As opposed to the first dataset that provides binary classification of patients with Parkinson's and healthy controls, the telemonitoring dataset will offer longitudinal severity indicators, including total and motor UPDRS scores. The integration of such datasets is shown to expand the size of the sample, boost the diversity of features, and better the model generalization across

heterogeneous clinical conditions that are heterogeneous. However, the combination of datasets that contain different label structures could represent possible threats, such as the inconsistency of labels, class imbalance, and bias in favor of Parkinson's disease samples. In order to alleviate these risks, label harmonization, feature standardization, and elimination of redundant identifiers were implemented. Besides, stratified train-test splitting and five-fold cross-validation were used to provide balanced learning and strong performance estimation. These ensured that the integration of datasets enhanced the strength of models and reduced bias and loss of clinical relevance. Identifiers were retained only for record-level traceability and were never used as predictive inputs.

Model Development:

The suggested framework uses a machine learning-based solution on an Agentic AI paradigm. Parkinson's Disease was detected using an XGBoost classifier because it is a strong classifier and can identify non-linearities in structured biomedical data. To estimate the severity of the disease, two XGBoost regression models were created to predict total UPDRS and motor UPDRS scores. To classify the data, an 80/20 portioning of the dataset was made with stratification to maintain a balance between the classes. Five-fold cross-validation was used to minimize overfitting and enhance the generalization. The measures used in evaluation were Accuracy, Precision, Recall, F1-score, and Confusion Matrix of classification, and MAE, RMSE, and R^2 of regression analysis.

Performance Summary:

The suggested XGBoost-based model showed stable results in predicting and evaluating the severity of Parkinson's Disease with the help of voice biomedical features. The model helped to identify the non-linear relationships in the data, thus making it possible to detect the disease and estimate the stage without using various or ensemble learning algorithms. The analysis of the feature importance showed that frequency-related voice biomarkers, such as fundamental frequency variations, jitter, and shimmer, had the greatest contributions to the predictive performance of the model. These results suggest that the voice frequency pattern is a very important factor in the detection of the severity stages of Parkinson's disease. These findings support the fact that optimized single-model solutions can deliver strong and clinically significant inferences while retaining interpretability and computational efficiency under an Agentic AI-based diagnostic system. All evaluation metrics, such as accuracy, precision, recall, and F1-score, were calculated in five-fold cross-validation and averaged to evaluate the performance stability and statistical reliability.

Implementation:

The entire architecture was deployed with the Python programming language. All the experiments and model training have been performed in the Google Colab platform, which is a cloud-based Jupyter notebook environment with scalable compute resources. Some of the primary libraries used in this paper are NumPy and Pandas to process the data, Scikit-learn to preprocess and evaluate data, and XGBoost to classify and regress the data. Jobs persistence was done using Joblib. The implementation design can be extended to meet future integration with telehealth systems and clinical decision-support platforms because of the modular implementation design.

Experimental Setup:

There were controlled conditions of the computer experiments in which all experiments were done under controlled conditions to make them reproducible and fair. Several experiment runs were conducted with various random seeds. The cross-validation was five-fold to improve the strength of the model and to reduce the overfitting. The agentic controller kept on checking model outputs and also made sure that there were proper transitions between detection, severity estimation, and recommendation phases. The results of the experiment were documented to be analyzed and discussed later.

Hardware Requirements	Software Requirements
Storage: 512 GB SSD	OS: Windows 11
	Language: Python
	Libraries: NumPy, Pandas, TensorFlow, Matplotlib, Scikit-learn

Statistical Validation and Robustness Analysis:

Five-fold cross-validation was realized in the training and assessment of the proposed model in order to guarantee its statistical validity. Measures of performance were calculated alongside the individual fold and averaged to reduce variability and minimize the effect of random partitioning of data. This cross-validation plan gives an understanding of the stability of the model at the various subsets of the data. Further, robustness analysis was done, making several experimental runs with varying random seeds, and the trends of performance were equivalent. The distribution of the cross-validation results was used to estimate a 95% confidence interval, using the mean and standard deviation of cross-validation fold-wise measures of performance. The consistency that is observed both between folds and runs is indicative of the fact that the results reported are consistent, stable, and not a chance. Confidence intervals were computed from cross-validation fold-wise performance scores to quantify variability rather than to claim exact population-level estimates.

Results and Discussion:

This chapter provides a detailed analysis and discussion of the experimental findings achieved through the suggested model of Agentic Artificial Intelligence (AAI)-based system to detect and estimate Parkinson's Disease result. The main goal is to test the predictive capability, stability, and clinical applicability of the suggested solution. Findings are discussed against already existing studies to show how the framework has helped in coming up with intelligent and decision-oriented healthcare systems.

Parkinson's Disease Detection Results:

To assess the Parkinson's disease detection model, standard classification measures, such as accuracy, precision, recall, F1-score, and ROC-AUC, were employed. Biomedical voice data were combined, and a larger biomedical voice dataset was trained on an XGBoost-based classifier to classify between patients with Parkinson's disease and healthy people. The overall rate of accuracy of the model was 94.68%, which means that the percentage of correctly identified test instances was significant. This performance is an indication of the capability of this model to learn discriminatory patterns using voice-based biomarkers. The trade-off between the precision and recall obtained is 0.854, which is balanced and is crucial, especially in medical diagnostic systems where false positives and false negatives have clinical consequences. Also, ROC-AUC of 0.852 indicates that the classifier is useful in distinguishing cases of Parkinson's disease at various levels of decision-making. An increase in ROC-AUC value shows that the classes are well separable and that the model remains stable when the classification criterion changes. Despite the presence of the class imbalance, the dataset and the model showed high sensitivity to the cases of Parkinson's disease. This is an ideal feature of screening systems, as it is necessary to identify the affected individuals early so that the intervention and disease management can be provided on time. In general, the findings suggest that the suggested detection model can be used in early-stage screening and will be able to facilitate clinical decision-making reliably. The results of the detection of Parkinson's disease are presented below (Table 2).

Table 2. Performance evaluation of the Parkinson's Disease detection model

Metric	Value
Model	XGBoost Classifier
Accuracy	**94.68%**
Precision	0.86

Recall	0.85
F1-score	**0.854**
ROC-AUC	**0.852**
Evaluation Method	Stratified Train–Test Split (80/20)

Severity Estimation Results (UPDRS Prediction):

Besides the detection of the disease, the suggested framework is concerned with the estimation of the severity of Parkinson's disease through regression-based modeling. Two independent XGBoost regression models were created to estimate motor UPDRS and total UPDRS scores, commonly accepted clinical indicators of disease severity and disease progress. The motor UPDRS prediction model had a mean absolute error (MAE) of 1.91, root mean squared error (RMSE) of 2.65, and a R 2 of 0.889. These findings suggest that there is a high congruence between anticipated and real scores on motor severity, which proves that the model can reflect fine differences among motor impairment. In the same manner, the overall UPDRS regression model generated an MAE of 2.28, RMSE of 3.12, and a value of R 2 which is 0.914. The high value of R 2 implies that the model is a good estimate of the percentage of the overall disease severity. The high predictive accuracy is a validation of the appropriateness of voice-based characteristics in quantitative severity measurement. Proper disease severity estimation facilitates individual treatment planning and the monitoring of patients over time. The findings indicate that the developed framework can predict the severity of Parkinson's Disease reliably with the help of non-invasive voice data.

Table 3. Performance of XGBoost regression models for Parkinson's Disease severity estimation.

Target Variable	Model	MAE	RMSE	R^2
Motor UPDRS	XGBoost Regressor	1.91	2.65	0.889
Total UPDRS	XGBoost Regressor	2.28	3.12	0.914

Clinical Interpretation of mistakes of prediction in UPDRS:

The clinically acceptable range of reported prediction errors of motor and total UPDRS scores is within the application of decision support. Since the range of scores in UPDRS is generally large in terms of numerical boundaries, an absolute error of an average of 2-3 points suggests that there would be small deviations, which would not necessarily affect the classification of clinical stages or treatment decisions. These kinds of error margins are often deemed acceptable in a case of longitudinal monitoring, where absolute similarity of scores is less important than the overall trends and development patterns. As such, the severity estimation models suggested can be used to assist clinical evaluation and follow up as opposed to substituting the expertise of the neurological examination.

Feature Importance Analysis and Clinical Interpretation:

In order to enhance the interpretability of the generated framework, the feature importance analysis was carried out on the trained XGBoost classifier. The analysis will give a quantitative ordering of the strongest voice biomarkers that contribute to the detection of Parkinson's disease. Table 4 indicates the most important items according to their scores on importance as indicated by the model. According to the results, jitter, shimmer, and fundamental frequency variations are micro-perturbations and frequency-related voice characteristics that dominate in model prediction. In clinical terms, the instability in the vibration of the vocal folds related to neuromuscular impairment is associated with increased jitter and shimmer, which is a typical symptom of Parkinson's disease. Differences in fundamental frequency and harmonic-to-noise ratio indicate impaired motor control and phonatory dysfunction, which is a normal feature of hypokinetic dysarthria. The ranked

features indicate that non-linear dynamical measures (DFA and RPDE) and micro-perturbation and noise-related measures (such as Jitter: DDP, Shimmer, HNR/NHR, and PPE) contribute most strongly to the classification of Parkinson's disease. This pattern is consistent with impaired neuromuscular control of phonation, where vocal instability and increased noise components are commonly observed in Parkinson's disease patients.

Table 4. Top-Ranked Voice Features Based on XGBoost

Rank	Feature Name	Importance Score
1	DFA	17.831
2	RPDE	13.272
3	Jitter: DDP	9.082
4	NHR	5.739
5	PPE	4.090
6	HNR	2.549
7	Shimmer	1.973
8	Jitter(Abs)	1.926
9	Shimmer: DDA	1.753

Agentic AI Decision-Making Analysis:

One of the contributions of the given research is the addition of an Agentic AI layer that allows autonomous interpretation of model outputs and systematic decision-making. In contrast to the traditional machine learning systems, which offer a standalone prediction, the suggested system has an intelligent agent that controls the diagnostic workflow. The detection of Parkinson's Disease by the classification model induces the severity estimation module in the agent. Predicted UPDRS scores are used to classify patients into three clinically meaningful stages, which include Early, Moderate, and Advanced. The staged classification assists in the organized medical decision-making and is correlated to the levels of disease progression applied to clinical practice. The agent also produces recommendations at the stage level in order to support clinicians and patients. The cases in their initial stages may be provided with guidance that is associated with regular monitoring and changing the lifestyle, but cases at their advanced stages are likely to need specialist consultation and hard treatment approaches. This self-directed reasoning feature makes the framework a digital clinical assistant, which lowers the level of manual intervention and enhances diagnostic performance. The agent uses rule-based decision logic with model output as opposed to reinforcement learning, which fosters transparency and clinical interpretability.

Explanation of the Agentic Decision Behavior:

The suggested Agentic AI system is deliberately modeled around a rule-controlled orchestration layer of decisions instead of a learning agent on the web. Whereas predictive learning is carried out through machine learning models (XGBoost classifier and regressors), the agent is used to autonomously coordinate the detection, severity estimation, stage classification, and recommendation generation, understandably and transparently, which can be interpreted by clinicians. The design decision is based on the priorities of safety, explainability, and reproducibility, which are essential specifications of the healthcare-based decision-support systems. Though the present application does not use reinforcement learning or self-updating policies, future extensions to adaptive or learning-based decision strategies can be made using the modular agent architecture as new longitudinal clinical feedback is made available.

Discussion on Deep Learning and Transformer-Based Speech Models:

The most recent approaches to speech-based detection of Parkinson's diseases have been investigating deep learning models, such as convolutional neural networks (CNNs), recurrent neural networks (LSTMs), and transformer-based self-supervised models, i.e.,

wav2vec and HuBERT. The methods exhibit high representation learning potentials, but are generally expensive in terms of large labeled datasets, massive computing capabilities, and complicated training chains. In comparison, the current paper is based on the concept of XGBoost-based machine learning to focus on interpretability, computational efficiency, and performance in situations of limited data. Transparency and reproducibility are essential to healthcare-oriented decision-support systems, and these are the primary attributes of such systems. However, the suggested agentic framework is model-agnostic, and it can be expanded in future research to include deep learning or transformer-based speech embeddings as clinically annotated datasets become accessible.

Table 5. Decision logic and outputs generated by the Agentic AI framework.

PD Detection Result	Severity Estimation	Disease Stage	System Recommendation
Not Detected	Not Applicable	–	Routine monitoring advised
Detected	Motor & Total UPDRS predicted	Early	Lifestyle modification and periodic monitoring
Detected	Motor & Total UPDRS predicted	Moderate	Medication review and structured therapy
Detected	Motor & Total UPDRS predicted	Advanced	Specialist consultation and intensive care planning

Comparison with Existing Studies:

The suggested Agentic AI framework has a more integrated and practical testing solution than conventional techniques, which are reported in the existing literature. Most of the previous researches are either binary disease detection or severity prediction and does not have much relevance to the overall clinical workflow. Contrarily, the suggested framework brings disease detection, severity estimation, and stage classification, as well as decision support, together in one pipeline [30][31]. The presence of an agentic decision layer even further differentiates this work, as it allows the disease staging to be performed autonomously, as well as the generation of recommendations. This combined methodology makes the framework more appropriate to be used in smart healthcare systems and telemedicine platforms as a holistic solution for Parkinson Disease diagnosis and treatment.

Numerical comparison across studies is limited by differences in datasets and evaluation protocols, and therefore, the performance comparison that was reported is contextual as opposed to a standardized benchmark [32][33][34].

Table 6. Comparison Table

Reference	Year	Precision	Recall	F1-Score	Accuracy
[1]	2025	99.4	97.6	98.4	99.1%
[3]	2021	97.5	93.2	90.3	85.6%
Proposed work	2025	0.96	0.95	0.95	94.68%

Note: The performance indicators of previous research are obtained based on the relevant publications. Disagreement in datasets, feature extraction procedures, and evaluation protocols can occur. This comparison is therefore aimed at giving a qualitative reference but not a highly managed benchmark.

Conclusion and Future Work:

This thesis introduced the design, implementation, and evaluation of a five-layer Agentic Artificial Intelligence system to early identify and assess the severity of Parkinson's Disease with biomedical voice features. One of the limitations of the traditional diagnostic approaches included in the research is the agentic decision-making layer that goes beyond the static prediction. This framework is a proposal to combine machine learning based detection, regression-based severity estimation, disease staging, and decision support into a single

pipeline. The results of experimental assessment proved good consistency in both disease detection and severity estimation using the UPDRS, which emphasized the efficacy of voice-based biomarkers when it comes to making a non-invasive diagnosis. The results of the present study indicate that Agentic Artificial Intelligence may be an important contributor to changing the diagnostics of healthcare. Instead of acting purely as a prediction tool, the offered framework can be viewed as an intelligent decision-support tool, which can make reasoning, simulate the illness severity, and create clinical insights that can be acted upon. The framework enhances interpretability, effectiveness, and clinical relevance by integrating strong machine learning models with agentic decision logic. This strategy brings AI-supported diagnosis one step closer to reality, adopting technical predictions to real-world clinical work processes. There are some crucial directions that future research may continue this work. First, large-scale, multi-center, longitudinal clinical data should be used in validation to determine the generalizability and performance in the real world. Longitudinal data would also facilitate the prediction of the disease progression over time. Second, in the future, multimodal data sources, including gait analysis, handwriting patterns, wearable sensor data, and neuroimaging, can be implemented to improve the accuracy of the diagnosis and characterization of diseases. Third, other techniques that can be used to improve the agentic decision layer include adaptive reasoning mechanisms and clinician feedback. Future clinical validation investigations are also needed to assess the effect of the framework on the process of diagnostic work, patient outcomes, and regulatory preparedness. Although the presented framework has provided good evidence of a framework-of-concept execution on publicly accessible datasets, additional clinical validation needs to be performed before practical implementation.

References:

[1] S. R. Subash Neupane, Sudip Mittal, "Towards a HIPAA Compliant Agentic AI System in Healthcare," *arXiv:2504.17669*, 2025, doi: https://doi.org/10.48550/arXiv.2504.17669.

[2] C. C. Jennifer A. Foley, "A Bayesian predictive processing account of Othello syndrome in Parkinson's disease," *Cogn. Neuropsychiatry*, vol. 28, no. 4, 2023, doi: https://doi.org/10.1080/13546805.2023.2229080.

[3] C. D. Jie Mei, "Machine Learning for the Diagnosis of Parkinson's Disease: A Review of Literature," *Front. Aging Neurosci.*, vol. 13, 2011, doi: https://doi.org/10.3389/fnagi.2021.633752.

[4] T. Tabashum, R. C. Snyder, M. K. O'Brien, and M. V Albert, "Machine Learning Models for Parkinson Disease: Systematic Review.," *JMIR Med. informatics*, vol. 12, no. 1, p. e50117, May 2024, doi: 10.2196/50117.

[5] R. Dawadi, M. Inoue, J. T. Tay, A. Martin-Morales, T. Vu, and M. Araki, "Disease Prediction Using Machine Learning on Smartphone-Based Eye, Skin, and Voice Data: Scoping Review.," *JMIR AI*, vol. 4, no. 1, p. e59094, Mar. 2025, doi: 10.2196/59094.

[6] J. D. A.-L. Emiro J. Ibarra, "Towards a Corpus (and Language)-Independent Screening of Parkinson's Disease from Voice and Speech through Domain Adaptation," *Bioengineering*, vol. 10, no. 11, p. 1316, 2023, doi: https://doi.org/10.3390/bioengineering10111316.

[7] M. A. William Tam, "An Exploration of Wearable Device Features Used in UK Hospital Parkinson Disease Care: Scoping Review," *J Med Internet Res*, 2023, [Online]. Available: https://www.jmir.org/2023/1/e42950

[8] L. P. Rebecca Ting Jiin Loo, "Interpretable Machine Learning for Cross-Cohort Prediction of Motor Fluctuations in Parkinson's Disease," *Mov. Disord.*, 2025, doi: https://doi.org/10.1002/mds.30223.

[9] S. Begum, T. M. Singh, F. A. Khan, and S. Doss, "AI in Personalized Medicine: Redefining Healthcare Through Agentic Intelligence," pp. 41–57, 2025, doi:

10.1007/978-3-031-89424-4_4.

[10] W. R. Md Saiful Islam, "Using AI to measure Parkinson's disease severity at home," *npj Digit. Med.*, vol. 6, no. 1, 2023, doi: https://doi.org/10.1038/s41746-023-00905-9.

[11] K. R. S. Priyadharshini, "A Comprehensive framework for Parkinson's disease diagnosis using explainable artificial intelligence empowered machine learning techniques," *Alexandria Eng. J.*, vol. 107, pp. 568–582, 2024, doi: https://doi.org/10.1016/j.aej.2024.07.106.

[12] K. M. Belinda Ndlovu, "Machine Learning and Explainable AI for Parkinson's Disease Prediction: A Systematic Review," *Indones. J. Comput. Sci.*, 2025, [Online]. Available: http://ijcs.net/ijcs/index.php/ijcs/article/view/4837

[13] K. R. K. Hirak Mazumdar, "Soft Robotics for Parkinson's Disease Supported by Functional Materials and Artificial Intelligence," *BME Front.*, vol. 6, 2025, [Online]. Available: https://spj.science.org/doi/10.34133/bmef.0143

[14] M. D. Gina Wieringa, "Adjusting to living with Parkinson's disease; a meta-ethnography of qualitative research," *Disabil. Rehabil.*, 2022, [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/09638288.2021.1981467

[15] A. M. C. Morgan Senter, "Experiences of people living with Parkinson's disease and key stakeholders in dance-based programs: a systematic review and meta-ethnography," *Disabil. Rehabil.*, vol. 46, no. 26, 2024, doi: https://doi.org/10.1080/09638288.2024.2327499.

[16] S. S. Mashrur Rashik, "AI-Enabled Conversational Journaling for Advancing Parkinson's Disease Symptom Tracking," *Conf. Hum. Factors Comput. Syst. - Proc.*, 2025, [Online]. Available: https://dl.acm.org/doi/10.1145/3706598.3714280

[17] K. R. S. Saravanan, "Explainable Artificial Intelligence (EXAI) Models for Early Prediction of Parkinson's Disease Based on Spiral and Wave Drawings," *IEEE Access*, 2023, [Online]. Available: https://ieeexplore.ieee.org/document/10171347

[18] O. B. Peng Li-Hua, Lamarana Jallow, Yurong Tan, "Precision medicine and music therapy for Parkinson's Disease," *Clin. Park. Relat. Disord.*, vol. 13, p. 100382, 2025, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2590112525000866?via%3Dihub

[19] S. M. R. A. & M. P. Mahboube Sadat Hosseini, "Cross-regional radiomics: a novel framework for relationship-based feature extraction with validation in Parkinson's disease motor subtyping," *BioData Min.*, vol. 18, no. 67, 2025, doi: https://doi.org/10.1186/s13040-025-00483-4.

[20] M. F. Kensuke Daida, "Long-Read Sequencing Resolves a Complex Structural Variant in PRKN Parkinson's Disease," *Mov. Disord.*, 2023, doi: https://doi.org/10.1002/mds.29610.

[21] Z. L. Elaine G.Y. Chew, "Exome sequencing in Asian populations identifies low-frequency and rare coding variation influencing Parkinson's disease risk," *Nat. Aging*, 2024, [Online]. Available: https://www.nature.com/articles/s43587-024-00760-7

[22] I. P. & Y. I. Meltem Ece Kars, Yiming Wu, Peter D. Stenson, David N. Cooper, Johan Burisch, "The landscape of rare genetic variation associated with inflammatory bowel disease and Parkinson's disease comorbidity," *Genome Med.*, vol. 16, no. 66, 2024, doi: https://doi.org/10.1186/s13073-024-01335-2.

[23] D. V. Jennifer G. Goldman, "Delivering Multidisciplinary Rehabilitation Care in Parkinson's Disease: An International Consensus Statement," *J. Parkinsons. Dis.*, vol. 4, no. 1, 2024, [Online]. Available: https://journals.sagepub.com/doi/10.3233/JPD-230117

[24] M. B. Cristina Gonzalez-Robles, "Embedding Patient Input in Outcome Measures for

Long-Term Disease-Modifying Parkinson Disease Trials," *Mov. Disord.*, 2024, [Online]. Available: https://movementdisorders.onlinelibrary.wiley.com/doi/10.1002/mds.29691

[25] A. and Neurology, "AI and Neurology," *Neurol. Res. Pract.*, vol. 7, no. 11, 2025, doi: https://doi.org/10.1186/s42466-025-00367-2.

[26] W. P. Roongroj Bhidayasiri, "Management of dysphagia and gastroparesis in Parkinson's disease in real-world clinical practice – Balancing pharmacological and non-pharmacological approaches," *Front. Aging Neurosci.*, 2022, [Online]. Available: https://www.frontiersin.org/journals/aging-neuroscience/articles/10.3389/fnagi.2022.979826/full

[27] F. B. Panying Rong, "Leveraging the Potential of Digital Technology for Better Individualized Treatment of Parkinson's Disease," *Front. Neurol.*, 2022, [Online]. Available: https://www.frontiersin.org/journals/neurology/articles/10.3389/fneur.2022.788427/full

[28] "Parkinson's disease - Symptoms and causes - Mayo Clinic." Accessed: Dec. 22, 2025. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/parkinsons-disease/symptoms-causes/syc-20376055

[29] "Parkinsons Telemonitoring - UCI Machine Learning Repository." Accessed: Dec. 22, 2025. [Online]. Available: https://archive.ics.uci.edu/dataset/189/parkinsons+telemonitoring

[30] M. Wilken *et al.*, "Neurophysiological Analysis of the Posterior Subthalamic Area in a Patient with Holmes' Tremor," *Mov. Disord.*, vol. 39, no. 3, pp. 623–625, Mar. 2024, doi: 10.1002/MDS.29705;CTYPE:STRING:JOURNAL.

[31] B. L. Jingzhao Chen, "Abnormal behavior detection in industrial control systems based on CNN," *Alexandria Eng. J.*, vol. 107, pp. 643–651, 2024, doi: https://doi.org/10.1016/j.aej.2024.08.109.

[32] M. T. Elie Matar, "Functional Movement Disorder as a Prodromal Symptom of Parkinson's Disease—Clinical and Pathophysiological Insights," *Mov. Disord.*, vol. 39, no. 11, 2024, [Online]. Available: https://movementdisorders.onlinelibrary.wiley.com/doi/10.1002/mds.29958

[33] A. Aljohani, "Late feature fusion using neural network with voting classifier for Parkinson's disease detection," *BMC Med. Inform. Decis. Mak.*, vol. 24, 2024, doi: https://doi.org/10.1186/s12911-024-02683-0.

[34] P. A. Moh Afandi, "Diagnosis of Parkinson's Disease Using Convolutional Neural Network by Hand Drawing Images," *Pre prints*, 2023, [Online]. Available: https://www.preprints.org/manuscript/202312.1851