AI-Based Voice Analysis System for Predicting Parkinson’s Motor Symptom Severity

1. Project Overview

Parkinson’s Disease (PD) is a chronic neurodegenerative disorder that primarily affects motor function causing tremors, speech changes, and loss of coordination. Monitoring disease progression typically requires regular clinical assessments using the *Unified Parkinson’s Disease Rating Scale (UPDRS)*, a time-consuming and subjective process that depends on specialized expertise. Early detection and continuous monitoring of PD symptoms are critical for improving patient outcomes, yet traditional monitoring methods require **in-person clinical assessments**, which are **costly, subjective, and infrequent**.

There is a growing need for **non-invasive, accessible, and automated systems** that can monitor the progression and severity of PD symptoms remotely, particularly using speech, which is one of the earliest indicators of Parkinson’s.

Our project introduces an intelligent, voice-based system that automatically predicts a patient’s **motor UPDRS score** using **voice characteristics and patient metadata** (age, sex, and test time). The system leverages **machine learning and ensemble modeling** to deliver an accurate, non-invasive, and accessible tool for remote Parkinson’s symptom monitoring.

2. Literature Review

2.1 Voice-Based Parkinson’s Disease Monitoring

Tsanas et al. [1] introduced one of the most influential studies in this domain by demonstrating that motor_UPDRS scores can be accurately predicted using biomedical voice measurements. Their work validated speech as a reliable biomarker for telemonitoring Parkinson’s disease and provided the widely used Parkinson’s Telemonitoring Voice Dataset. This study laid the foundation for regression-based severity estimation rather than simple disease detection.

2.2 Dysphonia and Voice Feature Analysis

Little et al. [2] investigated the suitability of dysphonia measurements such as Jitter, Shimmer, Harmonics-to-Noise Ratio (HNR), and Noise-to-Harmonics Ratio (NHR) for Parkinson’s telemonitoring. Their findings confirmed strong correlations between these voice features and Parkinson’s motor symptoms.

While this study established the clinical relevance of voice biomarkers, it did not incorporate demographic features, temporal attributes, or severity interpretation. Our system extends this approach by integrating voice features with patient metadata (age, sex, test_time) and translating model outputs into clinically interpretable severity levels.

2.3 Parkinson’s Disease Detection Using Machine Learning

Sakar et al. [3] focused on Parkinson’s disease detection using multiple speech recordings and traditional ML classifiers such as Support Vector Machines (SVMs). Their work demonstrated that ML models can successfully discriminate between healthy individuals and PD patients.

However, this binary classification approach does not provide information about disease severity or progression, which limits its usefulness for long-term monitoring. Our work shifts the focus from detection to **continuous motor severity estimation**, which is more clinically valuable for patient management.

2.4 Deep Learning Approaches

Gunduz [4] explored deep learning models for Parkinson's disease classification using vocal features. While deep learning achieved competitive results, these models required large datasets, higher computational resources, and offered limited interpretability.

Given the structured nature of biomedical voice features and the moderate dataset size, our project adopts classical ensemble ML models, which provide better interpretability, efficiency and robustness for real-time deployment in healthcare settings.

2.5 Research Gaps and Our Contribution

Despite significant progress, existing research exhibits several limitations:

- Emphasis on binary disease detection rather than quantitative severity estimation
- Lack of **clinically interpretable regression outputs** for the end users that align directly with standard clinical scoring systems such as motor_UPDRS.
- Limited use of **ensemble learning techniques** to improve robustness and generalization on structured biomedical voice data.

Our project addresses these gaps by:

- Predicting **continuous Motor UPDRS scores** rather than performing simple classification.
- We employ an **optimized ensemble machine learning architecture** combining Random Forest, XGBoost, and LightGBM to improve accuracy and stability.
- Laying the foundation for **future longitudinal progression tracking**

References

[1] A. Tsanas, M. A. Little, P. E. McSharry, and L. O. Ramig, "Accurate telemonitoring of Parkinson's disease progression by non-invasive speech tests," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1015–1022, Apr. 2009.

[2] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, and L. O. Ramig, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1015–1022, Apr. 2009.

[3] B. E. Sakar et al., "Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 4, pp. 828–834, Jul. 2013.

[4] H. Gunduz, "Deep learning-based Parkinson's disease classification using vocal features," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 3, pp. 1151–1158, May 2019.

2. Problem Statement

Current Challenges:

Parkinson's disease diagnosis and progression tracking rely on **in-person evaluations**, which are:

- Subjective and prone to inter-rater variability.
- Inaccessible for patients in rural or resource-limited regions.
- Expensive and inconvenient for continuous monitoring.

Who is affected?

- Parkinson's disease patients (Specially in rural or resource-limited regions)
- Neurologists and clinicians

- Caregivers and healthcare institutions

The Unmet Need in Healthcare:

There is a strong need for a **remote, objective, non-invasive, and affordable system** to assess Parkinson's motor symptom severity without frequent hospital visits. Since voice impairments are among the earliest and most consistent symptoms, **voice analysis** offers an ideal, non-invasive biomarker for disease monitoring.

3. Dataset Description

Dataset Name: Parkinson's Telemonitoring Voice Dataset
Source: UCI Machine Learning Repository
Size: 5,875 voice recordings from 42 patients with early-stage Parkinson's disease.

Dataset Link: https://archive.ics.uci.edu/ml/datasets/Parkinsons+Telemonitoring

Attributes:

- 16 biomedical voice measurements(e.g., *Jitter*, *Shimmer*, *NHR*, *HNR*, *RPDE* etc.)
- Demographic attributes: *age*, *gender*
- Time-based attribute: *test_time*
- Target variables: *motor_UPDRS*

Why this dataset is appropriate:

- Clinically validated and widely cited
- Longitudinal recordings enabling severity modeling
- Rich biomedical voice features (Jitter, Shimmer, RPDE, DFA, PPE, etc.)
- Includes demographic and temporal attributes
- No missing values; high data quality

4. Proposed Solution

Our solution is a **Machine Learning–driven web application** that leverages voice-based biomedical signals to estimate Parkinson's motor severity levels.

System Workflow:

1. **Voice Feature Input -** Users (patients or clinicians) upload the voice recording with the demographic attributes through the web interface.
2. **Model Prediction -** The trained Ensemble ML model (Random Forest, XGBoost and LightGBM) predicts the motor_UPDRS score.
3. **Severity Interpretation -** The score is categorized into severity levels: *Mild, Moderate, Advanced, Severe*.
4. **Feedback to the user -** A clear, color-coded and interpretable summary is displayed in the frontend interface.

5. Methodology

5.1 Data Preprocessing Pipeline

The data preprocessing pipeline consisted of four sequential stages to ensure data quality and model readiness:

Stage 1: Data Cleaning and Validation

- Verified data integrity: no missing values detected across all features
- Removed non-predictive columns: subject#, total_UPDRS (keeping only motor_UPDRS as target)
- Validated feature distributions and identified outliers using descriptive statistics

Stage 2: Feature Selection
Two complementary feature selection methods were employed:

- **Statistical Method:** SelectKBest with f_regression scoring to identify top 15 features based on univariate statistical tests
- **Model-based Method:** Random Forest feature importance (100 estimators) to capture non-linear feature interactions
- Combined insights from both methods guided feature engineering decisions

Stage 3: Train-Test Split

- **Split Ratio:** 80% training (4,700 samples) / 20% testing (1,175 samples)
- **Random State:** 42 (for reproducibility)
- **Strategy:** Random sampling without stratification (continuous target variable)

Stage 4: Feature Scaling

- **Scaler:** RobustScaler (chosen over StandardScaler)
- **Implementation:** Scaler fitted exclusively on training data to prevent data leakage, then applied to test set

5.2 Model Architecture and Training Process

Our approach employed an ensemble methodology combining multiple regression algorithms to maximize prediction accuracy and robustness.

1. **Baseline Model Training (**Trained all 8 models with default hyperparameters**)**

- **Linear Models with Regularization**
  - Ridge Regression, Lasso Regression, ElasticNet
- **Tree-Based Ensemble Models**
  - Random Forest, Gradient Boosting, XGBoost and LightGBM
- **Support Vector Machine**
  - SVR: Support Vector Regression with RBF kernel

2. Hyperparameter Optimization Architecture:

Top 3 performing base models underwent hyperparameter optimization using GridSearchCV:

- **Random Forest:**

  n_estimators: [100, 200, 300], max_depth: [10, 20, None], min_samples_split: [2, 5, 10], min_samples_leaf: [1, 2, 4]

- **XGBoost:**

  n_estimators: [100, 200, 300], max_depth: [3, 6, 9], learning_rate: [0.01, 0.1, 0.2], subsample: [0.8, 0.9, 1.0]

- **LightGBM:**

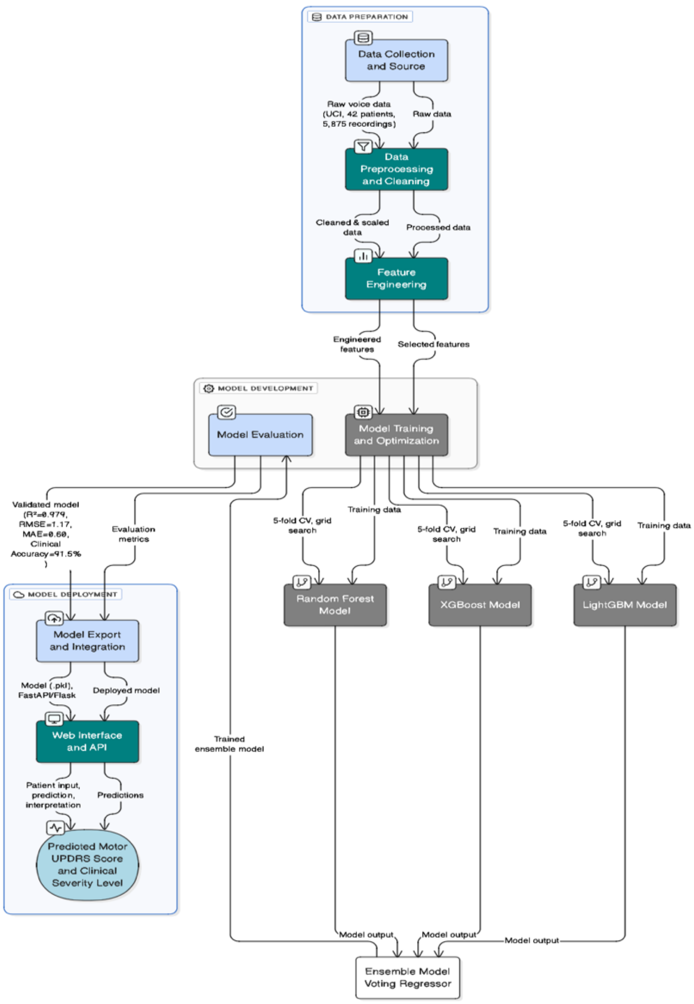
  n_estimators: [100, 200, 300], max_depth: [5, 10, 15], learning_rate: [0.01, 0.1, 0.2], num_leaves: [31, 50, 100]

5.3. Final Ensemble Architecture

- Combined top 3 optimized models using VotingRegressor
- **Aggregation Strategy:** Simple averaging (equal weights)

VotingRegressor (Averaging Ensemble)	→	Optimized Model 1 (Best Individual Model) + Optimized Model 2 (Second Best) + Optimized Model 3 (Third Best)

- Generated final predictions on test set



5.4 Validation Strategy

Cross-Validation (Primary Validation)

- **Method:** 5-Fold Cross-Validation
- **Metrics Computed:** Mean RMSE across 5 folds, Standard deviation of RMSE (stability measure)

Hold-Out Test Set (Final Evaluation)

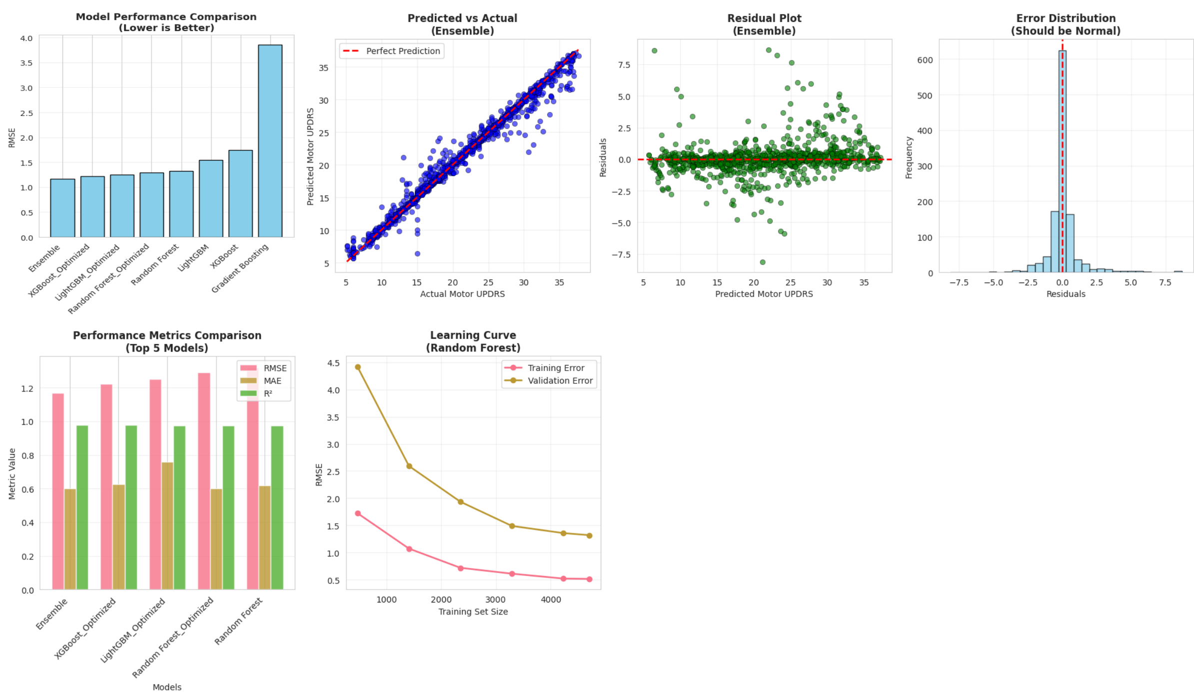
- **Metrics Computed:** RMSE, MAE and R² Score

6. Results and Performance

Model	Test RMSE	Test MAE	Test R^2	Performance
Ensemble (RF + XGB + LGBM)	1.169	0.600	0.979	Best overall
XGBoost (Optimized)	1.220	0.625	0.977	Excellent
LightGBM (Optimized)	1.251	0.759	0.976	Excellent
Random Forest (Optimized)	1.288	0.602	0.974	Excellent

Error Analysis

- **R^2 = 0.979**, explaining **97.9% of variance** in actual scores.
- **Average prediction error:** ±1.17 UPDRS points (clinically acceptable).
- Ensemble model achieved **91.5% accuracy within ±2 UPDRS points**.
- Residuals normally distributed, no systematic bias



Model Limitations

- Dataset limited to early-stage Parkinson's patients
- Voice recording environment variability
- Not designed for Parkinson's diagnosis in healthy individuals

7. Real-world Application

7.1 Deployment Scenario

- **Voice Feature Extraction:** Patient voice recordings are processed using **Parselmouth**, a Praat-based library, to extract clinically validated biomedical voice features such as Jitter, Shimmer, HNR, RPDE, DFA, and PPE.
- **Backend Service:** A **FastAPI-based backend** handles feature preprocessing, model inference, and response generation, enabling efficient and structured communication between the frontend and the ML model.
- **Frontend Interface:** A user-friendly **React-based web interface** allows users to input voice recordings and demographic data and view predicted motor_UPDRS scores with severity interpretation.
- **Containerization:** The entire system, including frontend and backend services, is **containerized using Docker**, ensuring consistency, portability, and ease of deployment.
- **Cloud Hosting:** The Dockerized application is deployed on **Microsoft Azure**, providing reliable hosting, scalability, and public accessibility.
- **Live System Availability:** The deployed system(MVP) is fully operational and accessible for demonstration through a public web link.
- **Live MVP Link:** https://parkinson-ai-eugccxchaseaascn.centralindia-01.azurewebsites.net

7.2 Potential Users

- Parkinson's disease patients (Specially in rural or resource-limited regions)
- Hospitals and Telemedicine platforms
- Neurologists and clinicians

7.3 Healthcare Workflow Integration

- **Patient Interaction:** Patients record and upload short voice samples during routine check-ups or from home, reducing the need for frequent in-person or hospital visits.
- **Clinical Decision Support:** The system provides clinicians with an objective motor_UPDRS estimate to support symptom assessment and treatment planning.
- **Follow-up & Monitoring:** Clinicians can use predicted severity scores alongside traditional evaluations to monitor symptom changes and guide follow-up decisions.

7.4 Risks & Limitations

- Not designed for Parkinson's diagnosis in healthy individuals.
- Requires standardized voice recording instructions.
- Not a diagnostic tool decision support only.
- Requires clinical validation before medical adoption.

8. Marketing & Impact Strategy

- **Target Adopters:** Parkinson's disease patients, Neurologists, hospitals, clinics, telemedicine platforms and healthcare research institutions seeking objective Parkinson's symptom assessment tools.
- **Practical Benefits:** Enables non-invasive, fast and consistent motor symptom evaluation, reducing clinician workload and improving patient follow-up.
- **Cost Effectiveness:** Low operational cost as it requires only short voice recordings, minimizing the need for specialized hardware or frequent hospital visits.
- **Accessibility & Reach:** Web-based design allows wide accessibility, supporting remote and underserved populations and enabling large-scale deployment.

9. Future Improvements

This project was developed as a **Minimum Viable Product (MVP)** within the limited timeframe of the **BioFusion Hackathon**. The current system focuses on predicting the **Motor Unified Parkinson's Disease Rating Scale (Motor UPDRS)** score using **voice recordings**, together with demographic and temporal features. This enables a **non-invasive, data-driven assessment** of Parkinson's motor symptom severity at a given point in time.

However, Parkinson's disease affects multiple motor domains, and voice-based analysis alone captures only a subset of motor impairments. Therefore, future work aims to extend this system into a **clinically robust, AI-based multimodal Parkinson's detection and progression monitoring platform**. Planned enhancements include the integration of **handwriting and drawing analysis**, **gait and posture analysis using cameras or wearable IMU sensors**, and **fine motor hand movement tests** to better reflect the comprehensive motor assessment performed in clinical practice.

In addition, future enhancements will extend the voice analysis component to support **longitudinal monitoring**. Patients will be able to **record and upload voice samples daily or weekly**, allowing the system to continuously analyze changes in motor symptom severity over time. Each new recording will be compared against the patient's historical data to detect and visualize **symptom progression trends**, such as whether symptoms are worsening, remaining stable, or improving.

These future extensions will transform the current model from a **single-time prediction tool** into a **continuous, AI-driven multimodal Parkinson's detection and progression monitoring system**, supporting informed clinical decision-making and personalized disease management.

Live MVP Link: https://parkinson-ai-eugccxchaseaascn.centralindia-01.azurewebsites.net