DEEP LEARNING APPROACH TO PARKINSON'S DISEASE DIAGNOSIS USING VOCAL FEATURE

Project Reference No.: *48S_BE_5158*

College : Bangalore Institute of Technology, Bengaluru
Branch : Department of Computer Science and Engineering
Guide(s) : Prof. Suma L
Student(s) : Ms. Shreya Shresth
Mr. Syed Mohammed Naqi Raza
Mr. Tushar Shetty
Mr. Sankar Rai

Keywords:

Parkinson's Disease, MFCC, Audio Classification, Machine Learning, CNN

Introduction:

Parkinson's Disease (PD) is a chronic, progressive neurological disorder that primarily affects motor function, leading to tremors, stiffness, bradykinesia, and speech impairments. It also affects cognitive and emotional well-being, significantly reducing quality of life over time.

Early diagnosis is critical, as timely intervention can slow disease progression and improve patient outcomes. However, traditional diagnostic methods rely heavily on clinical expertise and subjective evaluations, often resulting in delayed or inaccurate detection.

This project introduces a non-invasive, AI-driven approach for early PD diagnosis using voice analysis. Since vocal impairments are among the earliest symptoms of PD, speech signals offer a valuable diagnostic biomarker.

The system extracts **Mel-Frequency Cepstral Coefficients (MFCCs)** from recorded audio and uses machine learning and deep learning models to classify individuals as either healthy or potentially affected by PD.

It supports both real-time voice input and pre-recorded files through a user-friendly interface. The integration of multiple models enhances diagnostic accuracy while providing a **Parkinson's likelihood score** to aid interpretation.

This approach offers a scalable, low-cost solution that can assist clinicians, support telemedicine, and make early screening accessible even in remote or underserved regions.

Objectives:

1. **Extract Vocal Features**: Use techniques like MFCC to capture key speech characteristics related to Parkinson's Disease.
2. **Implement ML/DL Models**: Develop and compare Random Forest, SVM, KNN, and CNN classifiers for accurate detection.
3. **Enable Real-Time Detection**: Support live audio input for instant prediction of Parkinson's Disease.
4. **Evaluate Model Performance**: Use metrics like accuracy, precision, recall, and confusion matrix to assess model effectiveness.
5. **Support Early, Non-Invasive Diagnosis**: Provide a cost-effective tool to assist clinicians in early-stage Parkinson's diagnosis using voice analysis.

Methodology:

This project employs a structured and modular pipeline for the early detection and classification of Parkinson's Disease (PD) using vocal biomarkers extracted from speech recordings. The methodology integrates both classical machine learning (ML) and deep learning (DL) techniques to ensure high diagnostic accuracy and real-time applicability. The system is organized into four core subsystems: **S1 – Random Forest Classifier (RFC)**, **S2 – Support Vector Machine (SVM)**, **S3 – K-Nearest Neighbors (KNN)**, and **S4 – Convolutional Neural Network (CNN)**. The process begins with the acquisition of voice data either from the Kaggle Parkinson's Voice Dataset or through live microphone input using the sounddevice library. All audio samples are standardized using the librosa library to ensure consistency in format and sampling rate. From each sample, 22 **Mel-Frequency**

Cepstral Coefficients (MFCCs) are extracted to capture the most relevant vocal features. These extracted features are structured using pandas and numpy for ease of processing. The dataset is then split into training and testing sets. Subsystems S1, S2, and S3 involve the implementation of the RFC, SVM, and KNN models respectively using scikit-learn. Each model is trained on the MFCC features and evaluated using performance metrics such as accuracy, F1-score, and confusion matrices. In **S4**, a CNN is developed using TensorFlow/Keras to learn complex patterns in the reshaped MFCC inputs. The CNN consists of convolutional layers, max-pooling, dense layers, and a sigmoid output layer for binary classification. Real-time audio recordings are processed and passed through the CNN model for instant prediction. The final application is presented through a **Tkinter-based GUI**, enabling users to either record or upload audio, receive predictions from all models, and view a Parkinson's likelihood score—making the system user-friendly, scalable, and suitable for clinical and telemedicine environments.

Results & Conclusion:

After training and testing the machine learning models, the CNN model stood out with the highest accuracy in classifying PD from HC. Evaluation metrics such as accuracy and F1-score were used to compare model performance. The models were tested on unseen audio samples and produced consistent, reliable outputs. The GUI successfully allows users to interact with the backend models, select or record audio, and receive predictions instantly.

The Parkinson's likelihood ratio provided alongside the classification enhances the trust and interpretability of the system. The project has demonstrated the effectiveness of combining digital signal processing with AI for healthcare diagnosis and has potential utility in telemedicine and screening environments.

Project Outcome & Industry Relevance:

This project successfully delivered a technically sound and application-ready system for early diagnosis of Parkinson's Disease using vocal features. The team designed a complete pipeline involving audio preprocessing, feature extraction using MFCCs, and classification through multiple machine learning and deep learning models, including CNN.

A key outcome was the development of a GUI-based interface that enables real-time and file-based predictions, enhancing usability for non-technical users. The CNN model demonstrated superior accuracy and robustness in classification tasks, highlighting the effectiveness of deep learning for voice-based medical diagnostics.The project is highly relevant to industries focused on telemedicine, mobile health applications, and AI-assisted diagnostics. Its non-invasive nature, affordability, and portability make it suitable for clinical use and deployment in rural or resource-constrained environments.

Furthermore, the system's modular architecture allows easy extension to support additional neurodegenerative disorders. It exemplifies how artificial intelligence can be leveraged to transform conventional healthcare practices, offering accessible and scalable diagnostic tools for early disease detection.

Working Model vs. Simulation/Study:

The project resulted in a **working model** capable of real-time Parkinson's Disease detection using voice input. A GUI built with Tkinter supports both live audio recording and file upload for prediction. The backend processes audio, extracts MFCC features, and classifies them using trained models like CNN, SVM, KNN, and Random Forest.

Real-time predictions were achieved, but background noise during live testing affected accuracy. To address this, controlled simulations using clean, pre-recorded data were conducted to evaluate model performance.

Hence, the project involved both a **functional prototype** and **simulation-based analysis**, validating the system across real-world and ideal conditions.

Project Outcomes and Learnings:

Through this project, the team gained hands-on experience in applying machine learning and deep learning techniques to real-world healthcare problems. The development process involved data preprocessing, feature engineering, model building, and integration of a responsive user interface, reinforcing technical skills across the AI pipeline.

Key learnings included understanding the nuances of audio signal processing, the challenges of handling real-time noisy data, and the importance of model evaluation for reliable predictions. The team also developed a deeper appreciation for interdisciplinary problem-solving, combining knowledge from computer science and biomedical domains. Iterative testing and debugging further strengthened collaborative skills, logical thinking, and system design proficiency.

Future Scope:

The developed system lays a strong foundation for real-world deployment and can be significantly enhanced through the following future directions:

1. **Advanced Model Architectures**: Incorporating more sophisticated deep learning architectures such as **Long Short-Term Memory (LSTM)** networks, **Transformers**, or **Attention-based models** can improve temporal feature recognition and overall diagnostic accuracy.
2. **Dataset Expansion and Diversity**: To improve generalizability, the system can be trained on **larger and more diverse datasets**, including speech samples from **multiple languages, accents, and demographics**.
3. **Cloud and Mobile Integration**: Deployment as a **mobile or cloud-based application** would allow real-time voice-based PD screening in remote or underserved areas, making the tool more accessible and scalable.
4. **Clinical Collaboration and Validation**: Partnering with healthcare professionals and hospitals will enable **clinical trials and validation**,

refining the system for practical medical use and ensuring compliance with health data standards.

5. **Multimodal Disease Detection**: The framework can be adapted to detect other **neurodegenerative diseases** such as **Amyotrophic Lateral Sclerosis (ALS)** or **Alzheimer's Disease**, using speech patterns as early indicators.