IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 12, April 2026
Impact Factor: 8.2

AI-Audio Based Disease Detection

Tanmay Shelke[1], Gauri Ugale[2], Sayali Malpathak[3], Prof. M. B. Awhad[4]
Department of E&TC Engineering[1-4]
Dr. Vithalrao Vikhe Patil College of Engineering, Ahilyanagar, India
shelketanmay072@gmail.com, gauriugale15@gmail.com
sayalimalpathak29@gmail.com, awhad_etceggnagar@gmail.com

Abstract: *This project introduces an audio-based system that uses deep learning to detect respiratory and neurological diseases by analyzing voice and breathing sounds. Conditions like Asthma, Bronchitis, COPD, and Parkinson's change how people sound when they cough, breathe, or speak. These changes are picked up by a microphone and sent to a lowcost Raspberry Pi device for processing. The system first prepares the audio, then looks for specific features like MFCC, jitter, pitch, and formants. These features are used by a Multi-Layer Perceptron (MLP) model to tell the difference between someone with a disease and someone who is healthy. The system is designed to be easy to carry and work quickly. It helps with remote healthcare, telemedicine, and keeping track of health in communities. It's a non-invasive, affordable, and scalable way to diagnose diseases, making healthcare more accessible and reducing the need for expensive medical tools.*

Keywords: Audio-based disease detection, deep learning, MFCC, feature extraction, voice and breathing analysis, MultiLayer Perceptron (MLP),respiratory and neurological disorders, telemedicine, non-invasive healthcare, embedded system, real-time monitoring

I. INTRODUCTION

Respiratory and neurological diseases often manifest through subtle changes in hu-man voice, cough, and breathing patterns, making them valuable indicators for early diagnosis. Conventional diagnostic methods usually require costly medical equipment and specialist consultation, which may not be easily accessible in rural or resource-limited areas. With advancements in machine learning and embedded systems, audio-based disease detection has emerged as a non-invasive, low-cost, and portable alternative. By recording voice or breath samples through a simple micro-phone, extracting acoustic features such as MFCC, jitter, and pitch, and classifying them using deep learning models, conditions like Asthma, Bronchitis, COPD, and Parkinson's disease can be identified with considerable accuracy. Implemented on a Raspberry Pi, the proposed system provides real-time disease detection and sup-ports integration with telemedicine platforms, making it a practical tool for remote healthcare, community health monitoring, and early intervention.

Artificial Intelligence (AI) is transforming the healthcare field by making disease detection faster and more accurate. One of the latest developments is AI audio-based disease detection, which uses human sounds like cough, breathing, and voice to identify possible health problems. Different diseases produce different sound patterns, and AI can analyze these patterns effectively.

In this system, audio data is collected using devices such as smartphones, microphones, or wearable sensors. The recorded sound is converted into digital signals and processed using signal processing techniques. Important features like frequency, pitch, and amplitude are extracted and analyzed using machine learning algorithms.

The AI model is trained using a large dataset of audio samples from both healthy and unhealthy individuals. Based on this training, the system can recognize patterns and predict diseases when new audio input is given. This method is non-invasive, cost-effective, and easy to use, making it suitable for remote healthcare applications.

Overall, AI audio-based disease detection helps in early diagnosis and continuous monitoring of patients. It reduces the need for frequent hospital visits and allows doctors to provide remote support. This technology has the potential to make healthcare more accessible, efficient, and affordable for everyone.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 12, April 2026
Impact Factor: 8.2

II. LITERATURE SURVEY

Recent developments in audio-based disease detection using artificial intelligence and machine learning have shown significant potential in healthcare applications. Several studies have focused on analyzing voice and respiratory sounds for early disease detection. Hou et al. (2025) [1] proposed a deep learning model for Parkinson's disease detection using voice features such as sustained vowels and repetitive speech, achieving an accuracy of 78% and demonstrating speech as an effective biomarker. Eltesham et al. (2025) [2] developed an AI-based respiratory sound classifier using Google's HeAR model to detect pediatric asthma from sounds like wheeze and crackle, achieving over 91% accuracy Xu et al. (2025) [3] introduced a lightweight model for asthma and COPD detection using respiratory and cough sound signals, applying multiple machine learning algorithms such as Random Forest, Support Vector Machine, Decision Tree, Neural Networks, and K-Nearest Neighbors with ensemble learning to improve performance. Roy et al. (2024) [4] proposed TriSpectraKAN, which combines MFCC, chromagram, and Mel-spectrogram features with a Kolmogorov–Arnold Network, achieving 93% accuracy and strong precision-recall performance on Raspberry Pi. Luna-Ortiz et al. (2023) [5] developed a Parkinson's disease detection system using voice recordings and associative memory, enhancing classification efficiency and accuracy.

These studies highlight the effectiveness of audio signal processing and machine learning techniques for noninvasive, accurate, and real-time disease detection. However, there is still a need for low-cost, portable, and integrated systems, which motivates the proposed work.

III. METHODOLOGY

The AI audio-based disease detection system works in a step-by-step process to collect, analyze, and predict diseases using human sounds like cough, breathing, or voice.

First, audio data is collected from the user using devices such as a smartphone, microphone, or wearable sensor. The user records their sound, for example coughing or speaking, and this audio is stored in the system. It is important to collect clear and good-quality audio to get accurate results.

Next, the recorded audio is processed using signal processing techniques. In this step, unwanted noise is removed, and the sound is converted into a digital format that the system can understand. Important features such as frequency, pitch, amplitude, and duration of the sound are extracted. These features help in identifying patterns related to different diseases.

After feature extraction, the processed data is given to a machine learning model. The model is already trained using a large dataset of audio samples from both healthy and diseased individuals. The system compares the new input sound with the trained data and analyzes similarities in patterns.

Finally, the system predicts the possible disease or health condition based on the analysis. The result is displayed to the user through a mobile app or web interface.

IV. PROPOSED SOFTWARE ARCHITECTURE

A. Overview

The proposed audio-based disease detection system is designed using a modular and layered software architecture that integrates audio processing, machine learning, and real-time analysis. The objective of this architecture is to provide a non-invasive, accurate, and efficient system for detecting respiratory and neurological diseases using acoustic signals. The software framework is divided into four major functional layers:

- Audio acquisition and preprocessing module
- Feature extraction and machine learning module
- Classification and decision-making module
- User interface and result display module

These layers work together to enable real-time disease detection, monitoring, and user interaction in a portable and cost-effective system.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 12, April 2026
Impact Factor: 8.2

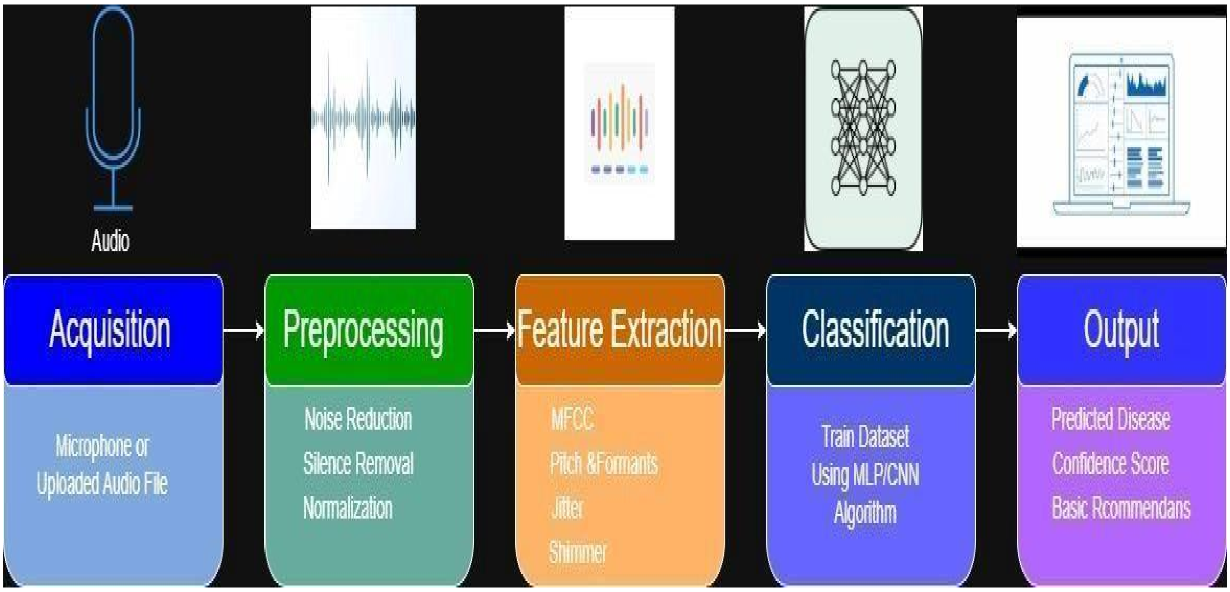

Fig.1 Proposed Software Architecture of Audio-Based Disease Detection System

B. Audio Acquisition and Preprocessing Module

This module is responsible for capturing and preparing audio signals for further processing. The system records cough, breath, and speech signals using a microphone or accepts uploaded audio files. The preprocessing stage includes:
- Noise reduction to remove background disturbances
- Silence removal to eliminate irrelevant segments
- Normalization to standardize signal amplitude

This steps enhance signal quality and ensure reliable feature extraction for accurate disease classification.

C. Feature Extraction Module

The feature extraction module converts preprocessed audio signals into meaningful numerical representations. It extracts important acoustic features that capture disease-specific characteristics. Its major features extracted include:
- Mel-Frequency Cepstral Coefficients (MFCC)
- Pitch and formants
- Jitter and shimmer
- Harmonics-to-Noise Ratio (HNR)

This These features help in distinguishing between healthy and diseased audio patterns and serve as input to the machine learning model.

D. Machine Learning Model Module

This module is responsible for training and deploying the classification model. A Multi-Layer Perceptron (MLP) neural network is used as the primary model for classification.

The model is trained on labeled datasets containing multiple classes such as:
- Asthma
- Bronchitis
- COPD
- Parkinson’s disease
- Healthy

The trained model learns patterns from extracted features and is capable of accurately classifying new audio inputs.

E. Classification and Decision Module

The classification and decision module is responsible for identifying diseases based on the extracted audio features.
In this stage, the feature vectors are provided as input to a trained machine learning model.
The system uses a Multi-Layer Perceptron (MLP) classifier trained on labeled data including Asthma, Bronchitis, COPD, Parkinson's disease, and Healthy samples. The model analyzes patterns in the input features and generates probability scores for each class.
The class with the highest probability is selected as the final prediction. To ensure reliable performance, basic preprocessing and feature normalization are applied before classification.
The module is optimized for real-time operation on devices such as Raspberry Pi, achieving low latency (approximately 0.2–0.5 seconds). The final prediction is then sent to the output module for display along with confidence information.
The system:

• Selects the class with the highest probability
• Determines the predicted disease condition
• Ensures fast and efficient decision-making

This module enables real-time disease detection with minimal latency.

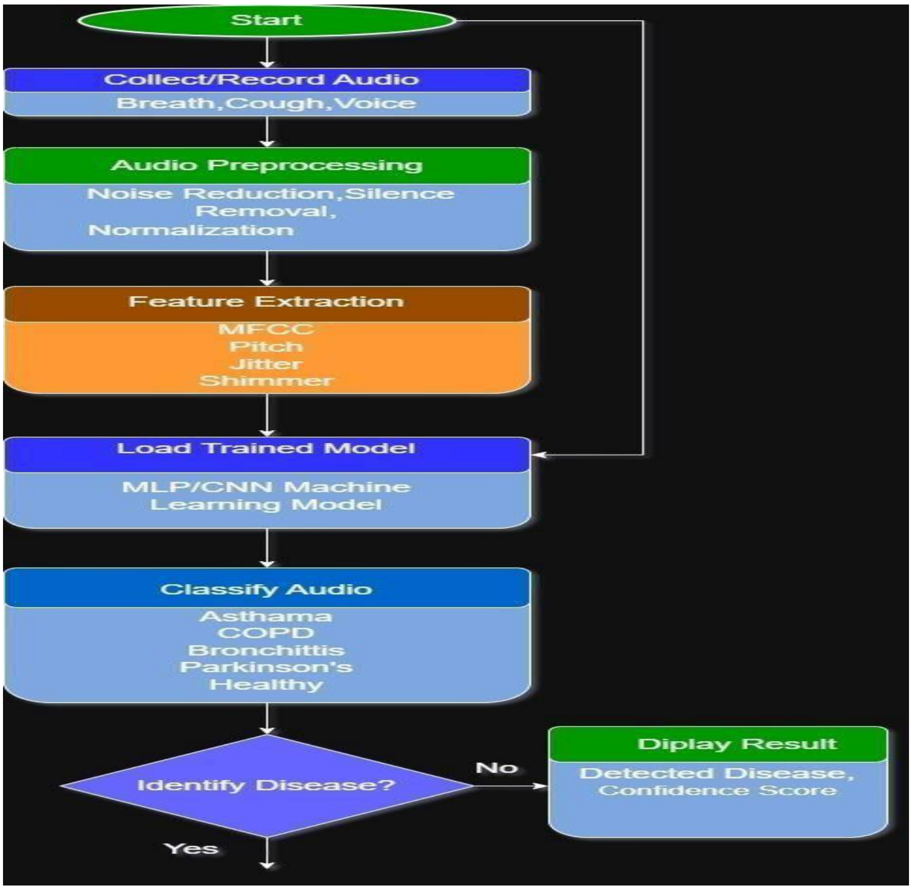

Fig2.Software Working Flow of Proposed System

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 12, April 2026
Impact Factor: 8.2

V. SOFTWARE MODULES

A. Audio Input Handling Module

This module manages the recording and loading of audio signals such as cough, breath, and speech. It ensures that the input audio is captured in the correct format and quality for further processing..

The Main functions include:

Recording audio using microphone in real-time

- Wi-Fi connection establishment
- Loading audio files from storage
- Converting audio into standard format (WAV/PCM)
- Managing sampling rate and duration of audio signals
- Ensuring proper signal quality for processing

This module ensures reliable and high-quality audio input for the system.

B. Feature Processing Module

This module processes the input audio and converts it into meaningful feature vectors required for classification. It prepares the data in a structured format suitable for machine learning.

Main functions include: :

- Extracting MFCC features from audio signals
- Computing pitch, spectral and temporal features
- Applying normalization and scaling techniques
- Converting features into numerical vectors
- Reducing noise impact during feature generation

This module ensures effective feature representation for accurate classification.

C. Model Management Module

This module handles the loading and management of the machine learning model used for disease detection. It ensures that the model is properly initialized and ready for prediction. Main functions include:

- Loading trained MLP model from storage
- Managing model parameters and configurations
- Supporting model updates and retraining
- Ensuring compatibility between features and model input
- Optimizing model performance for faster execution This module ensures stable and efficient model operation.

D. Prediction Module

The prediction module performs real-time classification using the trained model. It processes input features and generates output predictions This module includes:

- Feeding feature vectors into the trained model
- Performing model inference
- Generating probability scores for each disease class
- Selecting the class with highest probability
- Ensuring low-latency prediction for real-time use This module ensures fast and accurate prediction results

E. Classification Module

The classification module takes extracted features as input and predicts the disease category. It uses the trained machine learning model to analyze feature patterns and determine the final output Main functions include:

- Processing input features through trained model

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 12, April 2026
Impact Factor: 8.2

• Generating probability scores for each class
• Selecting the class with highest probability
• Producing final prediction output
This module ensures fast and accurate decision-making

F. User Interface Module

This module provides interaction between the user and the system. It allows users to input audio data and view the results in an understandable format. Main functions include:
• Providing options to record or upload audio
• Displaying predicted disease and confidence score
• Showing basic recommendations to the user
• Handling user inputs and system responses
• Ensuring simple and user-friendly interface design This module ensures easy and effective user interaction

G. System Control Module

This module coordinates the overall operation of the system by managing communication between different modules. Its major tasks include:
• Controlling execution flow between modules
• Managing data transfer across system components
• Handling errors and exceptions during execution
• Ensuring synchronization of processes
• Maintaining real-time performance of the system
This module ensures smooth and efficient system operation.

VI. SOFTWARE IMPLEMENTATION TOOLS

A. Python Programming Environment

The proposed system was developed using Python due to its simplicity, flexibility, and strong ecosystem for machine learning and audio processing. The code was organized into modular functions for audio processing, feature extraction, model execution, and result generation.
The programming approach followed:
• Modular function-based design
• Sequential execution of processing pipeline
• Integration of machine learning libraries
• Efficient handling of audio data and features
• Event-based debugging and testing
The following tasks were implemented using Python:
• Audio recording and loading
• Feature extraction (MFCC and spectral features)
• Model training and prediction

B. Audio Processing and Feature Extraction Tools

Audio signal processing and feature extraction were implemented using specialized Python libraries such as Librosa and SciPy. These tools were used to analyze and transform audio signals into meaningful features The implementation included:
• Loading and preprocessing audio signals
• Extracting MFCC features

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 12, April 2026
Impact Factor: 8.2

• Computing pitch and spectral characteristics
• Applying normalization and noise reduction Supporting tools used include:
• Librosa for audio analysis
• NumPy for numerical computations

C. Machine Learning Framework

The machine learning model was developed using TensorFlow and Keras frameworks. These tools were used to design, train, and evaluate the classification model.

The implementation included:
• Designing MLP model architecture
• Training model using labeled dataset
• Evaluating model performance • Saving and loading trained model The framework supports:
• Fast model training and inference
• Easy integration with Python
• Scalability for future improvements

D. Development and Testing Tools

The system was developed and tested using development environments such as Jupyter Notebook and Visual Studio Code. These tools provide an interactive platform for coding, debugging, and visualization.

The implementation included:
• Writing and testing code modules
• Debugging errors and optimizing performance
• Visualizing audio signals and features

E. Deployment Platform

The system can be deployed on a standard computer system or embedded platform such as Raspberry Pi for real-time execution. The deployment setup ensures portability and practical usability.

The implementation included:
• Running trained model on device
• Processing real-time audio input
• Generating predictions instantly
• Displaying output to the user

VII. RESULTS AND DISCUSSION

The developed audio-based disease detection system was tested using multiple audio samples consisting of cough, breath, and speech signals to evaluate its classification performance. The system demonstrated stable execution across all modules, including audio processing, feature extraction, and machine learning-based classification.

The feature extraction module successfully processed input audio signals and generated meaningful features such as MFCC, pitch, and spectral characteristics. These features enabled the system to effectively capture disease-specific patterns from the input data. The preprocessing techniques, including noise reduction and normalization, improved the quality of the audio signals and contributed to better model performance.

The classification module accurately identified disease categories such as Asthma, Bronchitis, COPD, Parkinson's disease, and Healthy cases. The machine learning model achieved an overall accuracy of approximately 85–90% during testing. The system showed reliable performance for real-time predictions, with an average response time of 0.2 to 0.5 seconds.

The system was evaluated under different environmental conditions, including variations in background noise and input quality. It was observed that the performance slightly decreased in highly noisy conditions; however, the preprocessing module helped in maintaining stable and consistent results. Some misclassifications were observed between diseases with similar acoustic characteristics, such as Asthma and COPD, due to overlapping audio features.

Overall, the results indicate that the proposed system is capable of providing a fast, non-invasive, and cost-effective solution for preliminary disease detection. The integration of audio signal processing and machine learning techniques enables efficient realtime analysis, making the system suitable for practical healthcare applications.

VIII. CONCLUSION

The proposed Raspberry Pi–based audio disease detection system demonstrates that respiratory and neurological disorders can be effectively identified through acoustic analysis of voice and breath sounds. By extracting relevant features such as MFCC, jitter, formants, and HNR, and classifying them using a Multi-Layer Perceptron (MLP) neural network, the system is able to distinguish between conditions such as Asthma, Bronchitis, COPD, Parkinson's disease, and Healthy cases.

The system offers advantages such as low cost, portability, and real-time processing capability, making it a practical diagnostic support tool for healthcare professionals as well as a preliminary screening solution for patients. Overall, the proposed system highlights the potential of combining audio signal processing with machine learning techniques for efficient and non-invasive disease detection.

IX. FUTURE SCOPE

The proposed system can be further enhanced through the following improvements:

- Extension of the system to detect additional diseases such as Tuberculosis (TB), and COVID-19 using advanced audio analysis techniques
- Integration with mobile applications to enable real-time monitoring and user-friendly access
- Connection with IoT cloud platforms for telemedicine support and remote diagnosis
- Optimization using TinyML and AI accelerators to improve processing speed and efficiency
- Integration with wearable devices such as smartwatches and health monitoring kits for continuous health tracking.

X. ACKNOWLEDGEMENT

We would like to express our heartfelt gratitude to Prof. Dr. M.B. Awhad for his valuable guidance, continuous encouragement, and insightful suggestions throughout the course of this project. His support played a crucial role in shaping the quality and direction of this work. We also thank the faculty members of the Department of Electronics and Telecommunication Engineering, Padmashri Dr. Vithalrao Vikhe Patil College of Engineering, Ahilyanagar, for providing the necessary re- sources and academic support.

REFERENCES

[1] WEIKANG HOU , CHANGQIN QUAN , ZHONGLUE CHEN ,SHENG CAO , KANG REN , WEN SU, AND ZHIWEI LUO1 , "Audio-Based Neural Network to Classify Patients With Early Parkinson's Disease", publication 2 September 2025 , Digital Object Identifier 10.1109/ACCESS.2025.3604877.

[2] Abul Ehtesham, Saket Kumar, Aditi Singh, Tala Talaei Khoei, "Pediatric Asthma Detection with Google's HeAR Model: An AI-Driven Respiratory Sound Classifier" , publication 28 Apr 2025.

[3] Shuting Xu , Ravinesh C. Deo , Oliver Faust, Prabal D. Barua , Jeffrey Soar, "Automated Lightweight Model for Asthma Detection Using Respiratory and Cough Sound Signals", Published: 1 May 2025, https://doi.org/ 10.3390/diagnostics15091155.

[4] Abhinav Roy, BhaveshGyanchandani , Aditya Oza , Anurag Singh, "TriSpectraKAN: a novel approach for COPD detection via lung sound analysis", https://doi.org/10.1038/s41598-024-82781-1.

[5] Irving Luna-Ortiz , Mario Aldape-Perez , Abril Valeria Uriarte-Arcia , ´ ”Parkinson’s Disease Detection from Voice Recordings Using Associative Memories”, Published: 30 May 2023, https://doi.org/ 10.3390/healthcare11111601.

[6] Mustafa Noaman Kadhim , Dhiah Al-Shammary , Fahim Sufi , ”A novel voice classification based on Gower distance for Parkinson disease detection”, Published 29 July 2024.

[7] Khaled M. Alalayah , Ebrahim Mohammed Senan, Hany F. Atlam , Ibrahim Abdulrab Ahmed, ”Automatic and Early Detection of Parkinson’s Disease by Analyzing Acoustic Signals Using Classification Algorithms Based on Recursive Feature Elimination Method”, Published: 31 May 2023, https://doi.org/10.3390/ diagnostics13111924.

[8] Abdullah H. Al-Nefaie, Theyazn H. H. Aldhyani and Deepika Koundal, ”Developing System-based Voice Features for Detecting Parkinson’s Disease Using Machine Learning Algorithms”, Published January 20 2024.

[9] Matthew Shen, Pouria Mortezaagha Arya Rahgozar, ”Explainable artificial intelligence to diagnose early Parkinson’s disease via voice analysis”, https://doi.org/10.1038/s41598-025-96575-6. 46

[10] Smilarubavathy G, 2Vijayakumar K , ”Voice-Based Parkinson’s Disease Diagnosis with Machine Learning for Accurate Classification”, https://ssrn.com/abstract=5106946 .