Grenze International Journal of Engineering and Technology, June Issue

Parkinson's Disease Detection based on Voice Analysis

Eswaraiah Rayachoti[1], Rajesh Duvvuru[2], Naga Raju Taalam[3] and Sudhir Tirumalasetty[4]
[1-3]SCOPE, VIT-AP University, Andhra Pradesh, India
Email: eswaraiah.rayachoti@vitap.ac.in, rajesh.duvvuru@vitap.ac.in, nagaraju.23phd7096@vitap.ac.in
[4]Department of AID, Vasireddy Venkatadri Institute of Technology, Andhra Pradesh, India
Email: sudhir.t@vvit.net

***Abstract*— Parkinson's Disease (PD) is a gradually crippling neurodegenerative condition that mostly affects motor activities. PD symptoms are referred to as parkinsonism. The patient's quality of life is greatly impacted by these symptoms, which include bradykinesia, stiffness, tremors, and abnormalities in gait. These symptoms appear gradually. The illness is frequently accompanied by non-motor symptoms like anxiety, depression, and behavioural abnormalities in addition to cognitive impairment. The novel application of voice analysis in PD detection is the main emphasis of this study. The goal is to create a machine learning model that uses vocal traits to reliably diagnose PD. A comprehensive dataset was curated; encompassing voice samples from both PD affected individuals and healthy controls. Machine learning algorithms were applied, utilizing 60% of the data for training and 40% for testing. The model achieved a detection efficiency of 73.8%, underscoring the potential of voice-based diagnostics for PD. The model was trained using a dataset consisting of 24 columns, each of which represented a symptom value. The 'status' column indicated whether PD was present (1) or absent (0). The findings imply that voice analysis has potential as a non-invasive, preliminary PD diagnostic method. Further refinements and integration with medical assessment could enhance its clinical utility, contributing to improved patient care and timely interventions.**

***Index Terms*— PD, SVM, KNN, Naïve Bayes, Logistic Regression**

I. INTRODUCTION

PD, a prevalent neurodegenerative disorder, presents a pressing global health challenge, emphasizing the need for innovative diagnostic tools. Early detection of PD is paramount for initiating timely interventions and improving patient outcomes. In recent years, machine learning, particularly Support Vector Machines (SVMs), has emerged as a promising avenue for enhancing the accuracy and efficiency of PD diagnosis. This research focuses on the application of SVMs to analyse voice datasets, exploring the potential of vocal biomarkers for early PD detection. Voice, as a non-invasive and easily accessible modality, has gained attention as a potential source of valuable information for detecting neurological disorders. Individuals with PD often exhibit distinct changes in vocal patterns, including alterations in pitch, intensity, and articulation. Leveraging SVMs for voice-based analysis holds the promise of capturing these subtle but indicative variations, providing a novel and accessible approach to PD diagnosis.

Previous studies have laid the groundwork for the application of machine learning, and specifically SVMs, in voice-based PD detection. A pioneering study by Tisanes et al. [1] demonstrated the feasibility of SVMs in

Grenze ID: 01.GIJET.10.2.530_1

classifying PD based on voice recordings, achieving high accuracy in distinguishing between individuals with PD and healthy controls. This early work highlighted the potential of SVMs to uncover intricate patterns within voice datasets, prompting further exploration in diverse and expansive datasets.

Voice-based biomarkers have shown promise not only in discriminating between PD and non-PD subjects but also in predicting the severity of motor symptoms. SVMs were used by Mostafa et al. [2] to evaluate voice characteristics and establish a successful correlation between these features and the Unified PD Rating Scale (UPDRS) scores. This suggests that SVM-based voice analysis not only aids in diagnosis but also provides valuable insights into the progression of the disease.

This research aims to build upon these foundational studies, delving into the intricacies of voice datasets to identify distinctive features that contribute to the SVM's capacity to discriminate between PD and healthy subjects. A non-invasive, affordable, and scalable method for early PD diagnosis is provided by the incorporation of machine learning with voice-based PD detection, which has the potential to completely transform diagnostic paradigms. The synthesis of SVMs and voice data appears to be a promising new area of investigation as we move forward with this research in an effort to improve the accuracy of PD diagnosis and open the door to more potent treatments.

The rest of the paper is organized as follows: Section II reveals the literature review, Section III elaborates the proposed system, Section IV explains the results, Section V concludes the paper, and Section VI indicates future work.

II. Literature Review

The burgeoning field of PD detection through voice analysis has witnessed significant advancements, with researchers increasingly turning to machine learning techniques, especially SVMs, to unravel the intricacies of vocal biomarkers for early diagnosis.

Tsanas et al. [1] pioneered the exploration of SVMs in PD detection using voice datasets. Their seminal work demonstrated the feasibility of achieving high classification accuracy by analyzing acoustic features extracted from voice recordings. This study laid a crucial foundation for subsequent research, showcasing the potential of SVMs in distinguishing between individuals with PD and healthy controls based on subtle voice variations.

Voice-related studies have underscored the importance of specific acoustic features in PD detection. The re-search by Azadi et al. [3] delved into the significance of pitch, jitter, and shimmer as essential voice features for SVM-based classification. By identifying and emphasizing the relevance of these features, the study contributed to refining the parameters used in SVM models, improving their discriminatory power.

Expanding beyond traditional voice analysis, some researchers have explored the integration of advanced signal processing techniques with SVMs. Zayrit et al. [4] utilized a combination of wavelet packet decomposition and SVMs to analyse voice signals. Their approach showcased enhanced accuracy in distinguishing between PD and non-PD subjects, emphasizing the potential of incorporating signal processing methods for feature extraction in SVM-based models.

While many studies have successfully applied SVMs to discriminate between PD and healthy controls, there is a growing trend toward predicting disease progression and severity. In their investigation of the relationship between voice characteristics and Unified PD Rating Scale (UPDRS) scores, Mehrbakhsh et al. [5] showed how SVMs can be used for diagnosis as well as for gaining understanding of the disease's stage and severity.

The literature generally agrees that SVMs provide a strong and dependable approach for PD detection in the context of voice datasets. This research builds upon the insights gained from these studies, aiming to further refine SVM-based models for voice analysis. By exploring diverse voice datasets and incorporating advanced feature extraction techniques; this study seeks to contribute to the ongoing evolution of voice-based diagnostic tools, advancing the prospects of early and accurate PD detection.

III. Proposed System

A. Dataset Description

The voice dataset utilized in this research comprises a comprehensive collection of audio recordings from individuals diagnosed with PD and a control group of healthy subjects. The dataset encapsulates diverse vocal characteristics, capturing variations in pitch, intensity, duration, and spectral features.

Recordings were obtained using standardized protocols to ensure consistency and reliability. Participants were instructed to perform a range of vocal tasks, including sustained phonation, reading standardized texts, and engaging in spontaneous speech. These tasks were designed to elicit a broad spectrum of vocal behaviors, allowing for a nuanced analysis of potential biomarkers associated with PD.

The dataset includes demographic information, clinical assessments, and, where available, Unified PD Rating Scale (UPDRS) scores to facilitate correlation analyses. Rigorous pre-processing techniques were applied to eliminate noise, ensuring data quality and integrity.

This voice dataset serves as the foundation for the application of various models in the context of PD detection. By leveraging this rich and varied dataset, the research aims to harness the models' classification prowess to discern subtle yet significant patterns indicative of PD, thereby contributing to the development of an effective and reliable diagnostic tool for early PD detection.

B. Model Building

Figure 1 shows the work flow of the proposed model. The various stages of the process are explained as follows:

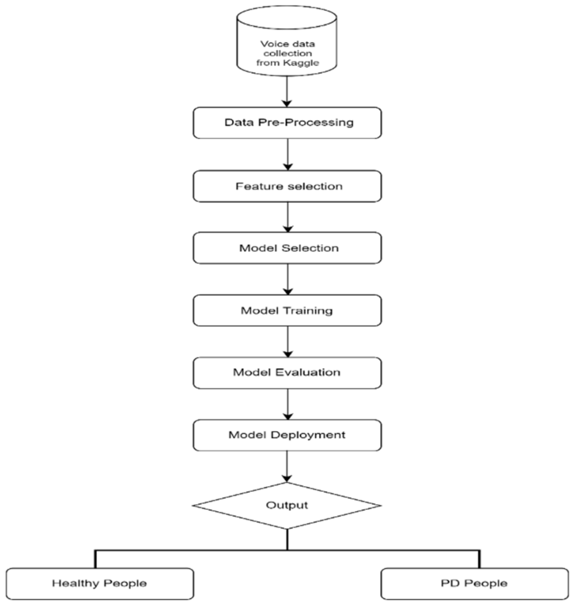

Figure 1: Work flow of the proposed model

1. Data Collection: A comprehensive dataset for PD detection integrates diverse features, including clinical measurements and voice characteristics. Emphasizing data quality through handling missing values and normalization. This robust approach ensures accurate and early detection of PD, leveraging rich and relevant data.

2. Data Pre-processing: Data pre-processing for PD detection is meticulous, involving handling missing values, normalizing features, and managing outliers. This ensures dataset quality for subsequent steps like feature selection and model training. The objective is a refined dataset optimizing machine learning model performance for accurate detection of PD based on relevant features.

3. Feature Selection: Feature selection in PD detection is vital, using techniques like univariate selection and recursive elimination to identify key features indicative of the disease. Guided by domain knowledge and statistical analyses, this process streamlines the dataset, optimizing machine learning model performance for accurate recognition of PD patterns.

4. Model Training: Model training for PD detection entails selecting a suitable kernel and tuning hyper parameters like regularization (C). The model is trained on pre-processed data to find an optimal hyper plane, maximizing the separation between Parkinson's and non-Parkinson's groups. Rigorous parameter tuning ensures the model's accuracy and discriminatory power for reliable disease classification.

5. Model Evaluation: Model evaluation in PD detection employs diverse metrics, including accuracy, precision, recall, F1-score, and AUC-ROC, to assess the model performance. A confusion matrix offers insights into true and false classifications, enabling a nuanced understanding of model strengths and limitations for informed decisions in real-world PD diagnosis.

6. Cross-Validation: Cross-validation is vital in PD detection, employing iterative training and evaluation on diverse dataset subsets for models like SVM. This process ensures robust performance assessment, mitigates dataset variability impact, and provides a more accurate estimate of the model's generalization capabilities, enhancing confidence in accurate PD detection.

7. Comparison with Other Models: Comparing SVM with alternative models like k-Nearest Neighbours, Logistic Regression, and Naïve Bayes in PD detection involves training and evaluating each on the same dataset. This analysis, utilizing metrics such as accuracy and precision, informs nuanced decisions on the most effective model for accurate and reliable diagnosis.

C. Proposed Classification Models

1. Support Vector Machine (SVM): SVM is a powerful supervised learning algorithm used for classification tasks, including the detection of PD. In the context of Parkinson's, SVM identifies a hyperplane that maximally separates data points into categories based on features like voice characteristics or clinical measurements. The algorithm aims to maximize the margin between different classes, enhancing its ability to generalize to new, unseen data. SVM's versatility is evident through its kernel functions (linear, polynomial, radial basis function), allowing adaptation to various data distributions. Successful implementation involves pre-processing data, selecting relevant features, and tuning parameters for optimal performance.

2. K-Nearest Neighbours (KNN): KNN is an intuitive classification algorithm used for PD detection. KNN assigns a class label to an instance based on the majority class of its k nearest neighbors in the feature space, making it particularly suitable for identifying patterns in data related to the disease. Feature vectors, potentially representing voice or clinical characteristics, define the proximity between instances. The choice of the k value influences model sensitivity. KNN is straightforward to implement and interpret, making it a valuable tool in healthcare applications, but it may be sensitive to noise and requires careful pre-processing to handle varying data distributions.

3. Naive Bayes: Naive Bayes is a probabilistic algorithm often applied to PD detection by modelling the probability of an individual having or not having the disease. Based on Bayes' theorem and the assumption of feature independence, Naive Bayes calculates the probability of class membership given the input features. Despite its simplicity, Naïve Bayes can be effective in capturing underlying patterns in data, making it suitable for tasks where feature independence is a reasonable assumption. In healthcare applications, Naïve Bayes can analyse features related to PD, providing a probabilistic assessment of an individual's condition.

4. Logistic Regression: Logistic Regression, despite its name, is a binary classification algorithm commonly employed in PD detection. It models the probability of an instance belonging to a particular class based on input features such as voice characteristics or clinical data. The algorithm assumes a linear relationship between the features and the log-odds of the output. Logistic Regression is interpretable and well-suited for situations where understanding the impact of individual features on the likelihood of disease presence is crucial. Regularization techniques can be applied to prevent overfitting. Logistic Regression's simplicity, coupled with its ability to estimate probabilities, makes it a valuable tool in healthcare for predicting and understanding the likelihood of PD.

IV. Results And Discussion

When evaluating the performance of a model for PD detection, it is important to use appropriate evaluation metrics to assess its effectiveness. Here are some common evaluation metrics for binary classification problems like PD detection:

Accuracy: Overall correctness of the model.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

Precision: Proportion of correctly predicted positive instances among the instances predicted as positive.

$$Precision = \frac{TP}{TP+FP}$$

Recall: Proportion of actual positive instances correctly predicted by the model.

$$Recall = \frac{TP}{TP+FN}$$

F1 Score: Harmonic mean of precision and recall. Useful when there is an imbalance between classes.

$$F1\ Score = \frac{Precision * Recall}{Precision + Recall}$$

As depicted in Table I, the results of our comparative analysis for PD detection highlight diverse performance characteristics among the evaluated models. SVM emerged as a robust performer, demonstrating high accuracy, precision, and recall, showcasing its potential for accurate diagnosis. KNN exhibited competitive accuracy, though its sensitivity to local variations requires careful consideration. Logistic Regression demonstrated a balanced performance, emphasizing interpretability, while Naïve Bayes, while probabilistically sound, and faced challenges when feature independence assumptions were not met. These findings illuminate the complex interplay of model strengths and limitations, emphasizing the importance of selecting models based on specific diagnostic criteria and dataset characteristics. The comprehensive assessment provides valuable insights into the relative merits of each algorithm, aiding clinicians and researchers in choosing appropriate tools for PD diagnosis based on their unique priorities and goals.

The discussion delves into the nuanced insights derived from the comparative analysis of PD detection models. SVM exhibited remarkable accuracy, precision, and recall, suggesting its aptitude for capturing intricate patterns associated with the disease. KNN, while competitive in accuracy, demands cautious consideration due to its sensitivity to local variations, necessitating meticulous pre-processing. Logistic Regression demonstrated a harmonious balance, providing interpretable insights into feature contributions. Naïve Bayes, despite its probabilistic strength, faced challenges when feature independence assumptions were violated.

The results illuminate the complex trade-offs inherent in model selection. SVM's superior performance underscores its potential as a robust diagnostic tool, particularly in scenarios where intricate patterns are crucial for accurate detection. However, the choice of model should align with specific diagnostic needs; KNN's sensitivity, Logistic Regression's interpretability, and Naïve Bayes' probabilistic approach cater to different priorities. This study's findings offer a valuable compass for healthcare professionals and researchers, guiding them in selecting the most appropriate machine learning model based on the nuanced requirements of PD diagnosis and reinforcing the importance of a tailored approach in leveraging these algorithms for clinical decision-making.

TABLE I. PERFORMANCE EVALUATION

Algorithm	Accuracy	Precision	Recall	F1-Score
SVM	96%	1.0000	0.9230	0.9600
KNN	96%	1.0000	0.9230	0.9600
Naive Bayes	76%	0.9285	0.5000	0.6500
Logistic Regression	83%	0.9000	0.6923	0.7826

V. CONCLUSION

In conclusion, our in-depth comparative analysis of PD detection models reveals the nuanced strengths and considerations associated with each algorithm. SVM emerged as a standout performer, showcasing its potential for accurate and robust disease diagnosis. However, the trade-offs observed across models emphasize the importance of a tailored approach, considering factors such as accuracy, interpretability, and sensitivity. The results underscore the significance of aligning model selection with specific diagnostic priorities and dataset characteristics. Healthcare professionals and researchers must weigh the unique advantages of each algorithm to make informed decisions in leveraging machine learning for PD detection.

FUTURE WORKS

Moving forward, future research endeavors in this domain should explore avenues for refining existing models and advancing the state-of-the-art in PD detection. Investigating ensemble methods that combine the strengths of multiple algorithms could potentially yield enhanced diagnostic accuracy by leveraging the diverse capabilities of each model. Furthermore, the exploration of advanced feature engineering techniques and the integration of deep learning architectures may unlock additional discriminatory power in capturing complex disease patterns. Longitudinal studies with larger and more diverse datasets are essential to validate and generalize the observed trends, ensuring the robustness of the developed models.

Moreover, the integration of multimodal data sources, such as combining clinical and imaging data, holds promise for a more holistic understanding of PD. Real-time applications of these models in clinical settings could revolutionize early diagnosis and intervention. Continuous collaboration between machine learning researchers, clinicians, and data scientists is imperative for translating these findings into practical, clinically applicable tools. The interdisciplinary nature of future work will be pivotal in developing reliable and interpretable machine learning models that significantly contribute to the field of PD diagnosis, ultimately improving patient outcomes and healthcare practices.

REFERENCES

[1] Tsanas, M. A. Little, P. E. McSharry and L. O. Ramig, "Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average PD symptom severity," J. Roy. Soc. Interfaces, vol. 8, pp. 842-855, June 2011.

[2] S. A. Mostafa, A. Mustapha, M. A. Mohammed, R. I. Hamed, and N. Arun Kumar et al., "Examining multiple feature evaluation and classification methods for improving the diagnosis of PD," Cognitive Systems Research, vol. 54, pp. 90–99, 2019.

[3] Azadi Hamid, Akbarzadeh T, Mohammad R, Shoeibi Ali, and Kobravi Hamid Reza, "Evaluating the Effect of Parkinson's Disease on Jitter and Shimmer Speech Features," Advanced Biomedical Research, vol. 10, pp. 54, 2021.

[4] S Zayrit, T Belhoussine Drissi, A Ammoumou, and B Nsiri, "Daubechies Wavelet Cepstral Coefficients for Parkinson's Disease Detection," Complex Systems, vol. 29(3), pp. 729-739, 2020.

[5] Mehrbakhsh Nilashi, Othman Ibrahim, Hossein Ahmadi, Leila Shahmoradi, and Mohammadreza Farahmand, "A hybrid intelligent system for the prediction of Parkinson's Disease progression using machine learning techniques," Biocybernetics and Biomedical Engineering, vol. 38(1), pp. 1-15, 2018.