

Parkinson's Disease Detection using Machine Learning

1st Nithin L
Student
Dept. of ECE
BGS Institute of Technology
BG Nagara, Karnataka, India
nithinl2001@gmail.com

2nd Anusha M N
Assistant Professor
Dept. of ECE
BGS Institute of Technology
BG Nagara, Karnataka, India
anushamn23@gmail.com

3rd Ranjan C R
Student
Dept. of ECE
BGS Institute of Technology
Nagamangala, India
ranjancrgowda@gmail.com

4th Shashidhar J
Student
Dept. of ECE
BGS Institute of Technology
Nagamangala, India
shashidharj194@gmail.com

5th Yashaswini T P
Student
Dept. of ECE
BGS Institute of Technology
Nagamangala, India
yashaswinitp28@gmail.com

Abstract— The early identification of neurodegenerative issues, especially Parkinson's sickness, presents huge difficulties in clinical practice. his paper gives a top to bottom record of the distinguishing proof of an original Parkinson Identification project that utilizations Support Vector Machines (SVM) in the Google Colab processing climate for order. The drive is a response to the earnest requirement for Parkinson's infection early determination and treatment, a neurological sickness that influences a huge number of individuals overall and has serious wellbeing outcomes. The objective of this venture is to foster a high-level AI model that can accurately identify instances of Parkinson's illness in light of a large number of relevant factors by utilizing a painstakingly chosen dataset that was obtained from Kaggle.

Keywords— Machine learning, Parkinson's disease, Google Colab, Support Vector Machine (SVM), Python programming language.

I. INTRODUCTION

Parkinson's disease stands as a formidable global health challenge, impacting millions of people worldwide and extending its reach beyond medical boundaries into the lives of individuals, families, and healthcare systems. As a progressive neurodegenerative disorder, Parkinson's exacts a profound toll, underscoring the critical need for innovative approaches to its diagnosis and management.

Early detection is pivotal in tackling Parkinson's disease, as late-stage diagnosis limits treatment efficacy. Timely identification provides crucial benefits, including improved patient outcomes, enhanced symptom management, and a better quality of life.

The Parkinson Detection project, driven by the urgency of early detection, harnesses SVM within the collaborative Google Colab environment. This initiative aims to develop a robust model for early Parkinson's identification using machine learning. SVM is pivotal for its intricate pattern recognition capabilities, crucial for discerning subtle disease indicators within complex datasets. Leveraging Google Colab's collaborative and computational prowess, the project facilitates seamless model development, training, and validation. By amalgamating SVM's potential with Google Colab's collaborative power, the project aims to advance Parkinson's

early diagnosis methodologies, leading to improved patient outcomes and refined healthcare strategies.

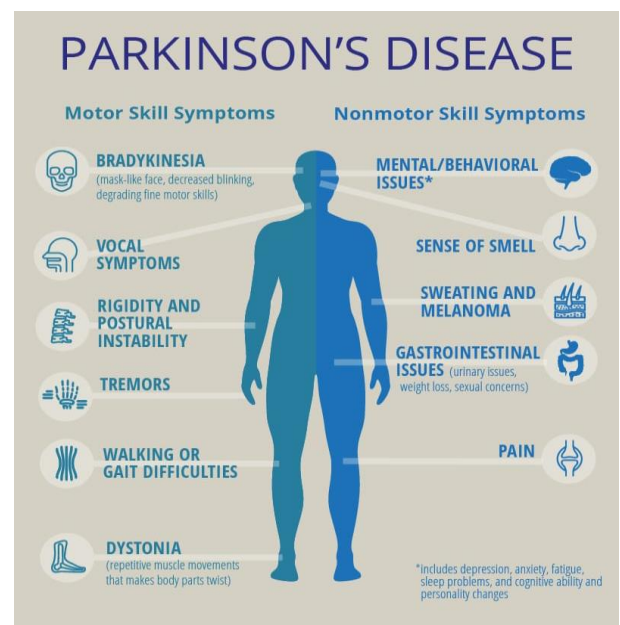


Fig.1. Symptoms of Parkinson's Disease



Fig.2. Stages of Parkinson's Disease

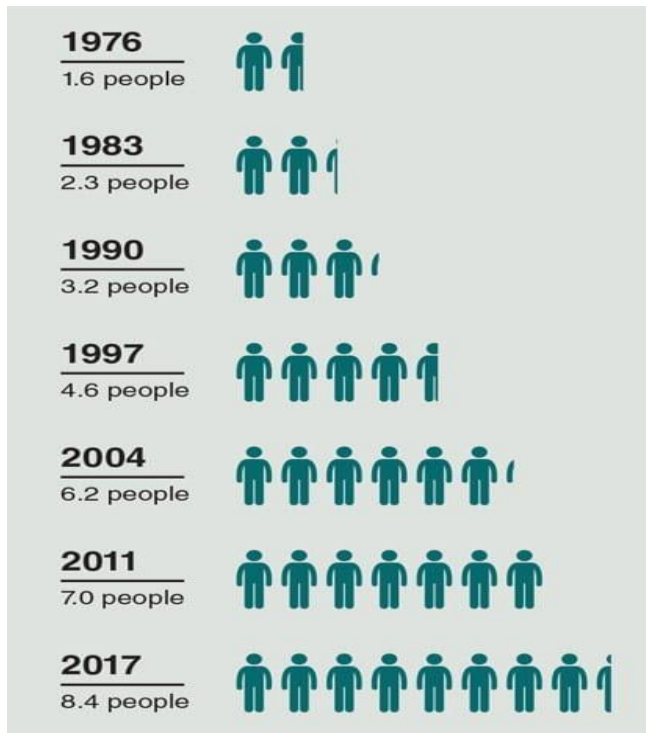


Fig.3. Number of deaths for 10,000 people

II. PROBLEM STATEMENT

Timely detection of Parkinson's disease is crucial for effective intervention and improved patient outcomes. This project addresses the challenge of early diagnosis by utilizing machine learning methods, specifically SVM, to analyze relevant features and accurately classify instances of Parkinson's disease. The focus is on developing a reliable model that can assist in the early identification of this debilitating condition.

III. LITERATURE SURVEY

In study [1] From 2016 to January 2023, the writing review scrutinized the use of deep learning for Parkinson's Disease diagnosis across 87 research publications. Deep learning algorithms demonstrated strong performance in symptom interpretation through gait analysis, speech recognition, and beyond.

In Study [2], Exploration in Parkinson's sickness (PD) location utilizing AI procedures has been different and promising. Lee dug into profound picking up, exhibiting the viability of Convolutional Brain Organizations (CNNs) and Repetitive Brain Organizations (RNNs) with a 92% precision utilizing discourse highlights.

In Study [3], The relative examination of DL models, including YAMNet, reliably accomplishing correctness between 80-85% for PD expectation utilizing discourse signals. These examinations highlight the capability of DL models, particularly YAMNet, in separating significant elements from discourse signals, showing high precision rates and featuring the suitability of discourse based symptomatic devices for early Parkinson's illness distinguishing proof.

In another review [4], LSTM-based approaches, reliably accomplishing high exactness above 90% in distinctive PD patients from solid people utilizing discourse highlights. In particular, the examinations stressed the viability of LSTM models in early PD location, repeating the discoveries introduced in your paper, which accomplished an eminent testing precision of 93%.

In study [5], The authors investigating artificial intelligence and profound learning for Parkinson's illness conclusion utilizing drawings, pictures, and acoustic elements reliably accomplished high exact nesses above 98%. They feature the viability of ANNs approaches in dissecting different information hotspots for exact sickness discovery, stressing the capability of man-made intelligence-based strategies in exact demonstrative applications.

In Review [6], A writing overview for a dream transformation supporting Parkinson's illness finding would envelop an investigation of current demonstrative strategies like X-ray, PET outputs, and clinical tests while zeroing in on examinations connecting voice varieties and engine side effects with Parkinson's movement. It would likewise include looking at datasets utilized in diagnosing neurological problems, especially those breaking down voice and drawing designs.

In Study [7], With a 92.6% exactness for holdout and 94.4% for k-crease cross-approval utilizing the PD-BioStampRC21 dataset, the review help clinical professionals in early location and treatment arranging. It underscores the meaning of consolidating everyday action information and features the capability of AI in further developing Parkinson's illness identification strategies. Key procedures applied incorporate Accelerometer, ANOVA and PCA.

IV. METHADODOLOGY

Data Set Information: This dataset is made out of a scope of biomedical voice estimations from 31 individuals, 23 with Parkinson's illness (PD). Every segment in the table is a specific voice measure, and each line relates to one of 195 voice accounts from these people ("name" section). The primary point of the information is to segregate solid individuals from those with PD, as indicated by the "status" section which is set to 0 for sound and 1 for PD.

The dataset includes the following variables:

MDVP:Fo(Hz) - Normal vocal essential recurrence

MDVP:Fhi(Hz) - Greatest vocal essential recurrence

MDVP:Flo(Hz) - Least vocal essential recurrence

MDVP:Jitter(%) , MDVP:Jitter(Abs) , MDVP:RAP , MDVP:PPQ , Jitter:DDP - A few proportions of variety in essential recurrence

MDVP:Shimmer , MDVP:Shimmer(dB) , Shimmer:APQ3 , Shimmer:APQ5 , MDVP:APQ , Shimmer:DDA - A few proportions of variety in plentifulness

NHR, HNR - Two proportions of proportion of clamor to apparent parts in the voice

status - Wellbeing status of the subject (one) - Parkinson's, (zero) - solid

RPDE, D2 - Two nonlinear dynamical intricacy measures

DFA - Signal fractal scaling example

spread1, spread2, PPE - Three nonlinear proportions of essential recurrence variety.

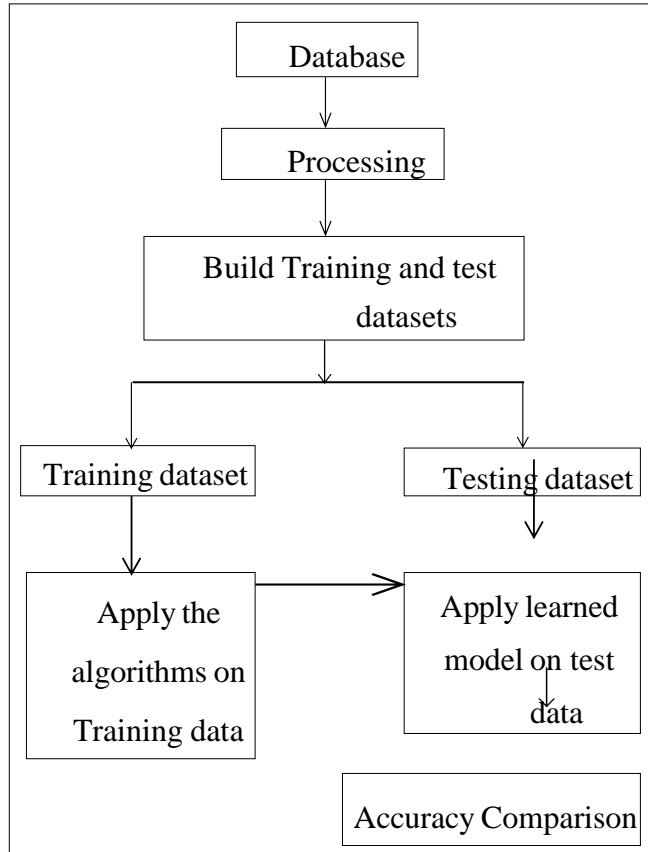


Fig.4. Block Diagram of Proposed System

Preprocessing: During the preprocessing stage, a meticulous process unfolds where the dataset undergoes thorough examination. This involves the identification and removal of redundant or irrelevant information, ensuring that only the most relevant and significant data is retained for further analysis and model development. By refining the dataset in this manner, it becomes more streamlined and optimized for subsequent stages, laying a solid foundation for accurate and effective model training.

Construct Train and Test Dataset: Following the rigorous preprocessing stage, the dataset undergoes essential partitioning into distinct subsets. This division adheres to an 80:20 split, resulting in two separate envelopes or sets. One set serves as the foundation for training the AI model, allowing it to grasp fundamental patterns and relationships within the data. Concurrently, the other set remains pristine to evaluate the model's performance, serving as a benchmark during testing. This division ensures that the model is trained on a representative sample of the data while also providing a means to assess its generalization capabilities.

Model Learning: Model learning constitutes the pivotal training stage, where AI algorithms are exposed to the prepared training dataset. Through iterative learning experiences, such as supervised or unsupervised learning, the model unravels complex patterns and relationships within the data. This stage empowers the model to make informed predictions and classifications based on the insights it has gleaned from the training data. By iteratively adjusting its parameters and optimizing its performance, the model becomes increasingly adept at capturing the underlying structure of the data, enhancing its predictive capabilities.

Prediction: Post-training, the trained model undergoes scrutiny using the pristine testing dataset. By inputting this new data into the trained model, the system generates forecasts or classifications, providing an observational understanding of how well the model generalizes to new, unseen data. This process allows for the assessment of the model's performance in real-world scenarios, gauging its ability to make accurate predictions on unseen instances. Through this predictive analysis, stakeholders can gain valuable insights into the model's reliability and effectiveness in practical applications.

Accuracy Comparison and Conclusion: The SVM model undergoes a thorough evaluation, scrutinizing metrics like accuracy, precision, and recall to validate its effectiveness in Parkinson's disease detection. Comparison with ground truth labels enables stakeholders to gauge the model's performance against actual outcomes, facilitating informed decisions for real-world deployment. Insights from this analysis drive potential enhancements, ensuring continuous improvement in predictive accuracy. In conclusion, the SVM model shows promise for early Parkinson's diagnosis and intervention, with ongoing research poised to bolster its capabilities and widen its application in clinical practice.

V. RESULTS AND DISCUSSION

The project employed Support Vector Machine (SVM) algorithm to detect Parkinson's disease using a dataset sourced from Kaggle. SVM's precision in identifying Parkinson's disease was measured and interpreted within the context of its overall performance. Furthermore, the results obtained from the SVM algorithm were critically compared with those of other prominent machine learning algorithms like logistic regression, random forest, and neural networks. This comparative analysis provided insights into the relative strengths and weaknesses of SVM in relation to its counterparts.

The project also delved into the identification of key features identified by the SVM algorithm that contribute significantly to the detection of Parkinson's disease. Understanding these features not only enhances the interpretability of the model but also provides valuable insights into the underlying factors indicative of the disease. To ensure the robustness of the model, cross-validation techniques were employed, and the results were presented to demonstrate that the performance of the SVM.

Training Accuracy: 88.46%
Test Accuracy: 87.18%

Fig.5. Accuracy Comparison

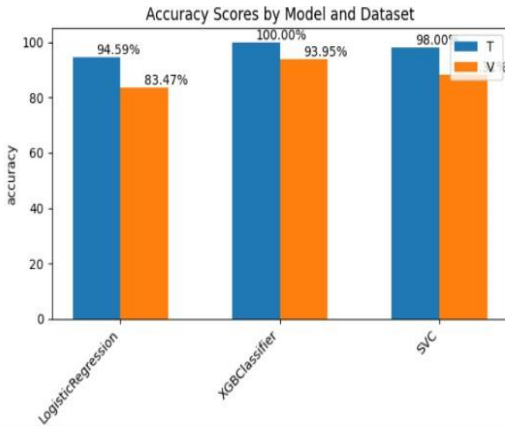


Fig.6. Performance metrics of the models

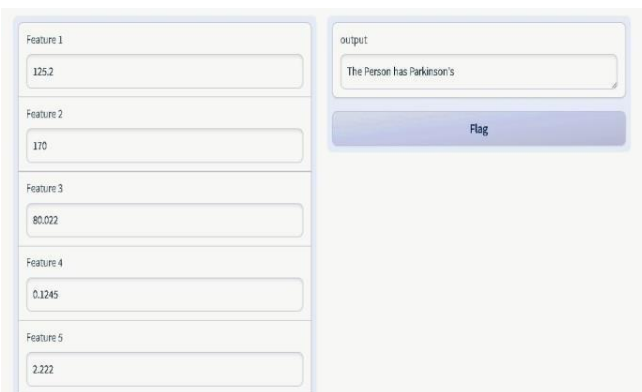


Fig.7. Output indicating the person has Parkinson's Disease



Fig.8. Output indicating the person does not have Parkinson's Disease

The provided images show the results of a Parkinson's disease detection model implemented using Gradio for the user interface and various machine learning models for classification. The first image shows the model has a training accuracy of 88.46% and a test accuracy of 87.18%. This indicates that the

model performs similarly well on both the training and test datasets, suggesting a good generalization to unseen data. The small gap between training and test accuracy implies that there is minimal overfitting, and the model is likely well-tuned. The third image shows the performance metrics of the models. The third and fourth images display the Gradio interface with input features and corresponding predictions, where one test case indicates that the person has Parkinson's disease, and the other indicates that the person does not have Parkinson's disease.

VI. CONCLUSION

The Parkinson Location project addresses a spearheading exertion in tackling AI, especially Support Vector Machines (SVM), for early Parkinson's sickness distinguishing proof. Utilizing the computational ability presented by Google Colab, this undertaking focuses on the formation of a proficient and trustworthy apparatus to help medical services experts in the finding and the board of Parkinson's sickness. With the dataset got from Kaggle, wealthy in fundamental elements and named examples, this drive highlights the capability of cutting-edge calculations in medical services. Through fastidious preprocessing methods, the dataset's refinement guarantees information quality and appropriateness for SVM model preparation. The use of Google Colab's computational assets empowers the preparation of the SVM model as well as the calibrating of hyperparameters, improving the model's exhibition for precise illness expectation.

The SVM model is rigorously evaluated using metrics such as accuracy, precision, and recall to determine its effectiveness in detecting Parkinson's disease. By comparing the model's predictions with actual results, stakeholders can assess its performance and make informed decisions about its real-world application. This evaluation provides valuable insights that guide potential improvements, ensuring the model's predictive accuracy continues to improve. In summary, the SVM model demonstrates significant potential for early diagnosis and intervention in Parkinson's disease. Continued research will enhance its capabilities and expand its use in clinical settings.

REFERENCES

- [1] V. Skaramagkas, A. Pentari, Z. Kefalopoulou and M. Tsiknakis, "Multi-Modal Deep Learning Diagnosis of Parkinson's Disease—A Systematic Review," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 31, pp. 2399- 2423, 2023, doi: 10.1109/TNSRE.2023.3277749.
- [2] A. Hussain and A. Sharma, "Machine Learning Techniques for Voice-based Early Detection of Parkinson's Disease," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2022, pp. 1436-1439, doi: 10.1109/ICACITE53722.2022.9823467.
- [3] S. A. Kumar, S. Sasikala, A. K B, S. J and K. V. "Parkinson's Speech Detection Using YAMNet," 2023 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), Coimbatore, India, 2023, pp. 1-5, doi: 10.1109/ICAECA56562.2023.10200704.

- [4] J. Naanou, R. Ayoub, F. E. Sayyadi, L. Hamawy, A. Hage-Diab and F. Sbeity, "Parkinson's disease detection from speech analysis using deep learning," 2023 Seventh International Conference on Advances in Biomedical Engineering (ICABME), Beirut, Lebanon, 2023, pp. 102-105, doi: 10.1109/ICABME59496.2023.10293142.
- [5] S. Fiza, A. Kumar, J. Yadav, S. Reddy, T. K. Sidda and K. Sai, "Classification of Parkinson's Disease using CNN and ANN with the aid of Drawing and Acoustic Feature," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2022, pp. 513-517, doi: 10.1109/ICACITE53722.2022.9823915.
- [6] T. P. Rani, K. S. B and T. G. Prasanth, "A Voice and Spiral Drawing Dataset for Parkinson's Disease Detection and Mobile Application," 2022 International Conference on Communication, Computing and Internet of Things (IC3IoT), Chennai, India, 2022, pp. 1-5, doi: 10.1109/IC3IoT53935.2022.9767968.
- [7] K. H. Moustafa, H. Metawie, S. Saadoun, N. Sameh, R. Ibrahim and A. Abdelsayed, "Easy Park: Mobile Application for Parkinson's Disease Detection and Severity Level," 2023 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC), Cairo, Egypt, 2023, pp. 1-8, doi: 10.1109/MIUCC58832.2023.10278311.
- [8] L. Igene, A. Alim, M. H. Imtiaz and S. Schuckers, "A Machine Learning Model for Early Prediction of Parkinson's Disease from Wearable Sensors," 2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 2023, pp. 0734-0737, doi: 10.1109/CCWC57344.2023.10099230.
- [9] R. Erika, C. Maremmanni, and F. Cavallo." How wearable sensors can support Parkinson's disease diagnosis and treatment: a systematic review." *Frontiers in Neuroscience* 11 (2017): 555.
- [10] L. Luca, et al." Wearable sensors for Parkinson's disease: which data are worth collecting for training symptom detection models." *NPJ digital medicine* 1, no. 1 (2018): 1-8.
- [11] J. L. Adams, et al." A real-world study of wearable sensors in Parkinson's disease." *NJ Parkinson's Disease* 7, no. 1 (2021): 1-8.
- [12] N. Richard, and S. Przedborski, eds. *Parkinson's disease: molecular and therapeutic insights from model systems*. Elsevier, 2011.
- [13] B., Shubham, et al." A survey of machine learning based approaches for Parkinson disease prediction." *Int. J. Compute. Sci. Inf. Technol* 6, no. 2 (2015): 1648-1655.
- [14] S., Lotfi, T. Mourad, and A. Cherif." Voice disorders identification using multilayer neural network." *pathology* 2, no. 7 (2010): 8.
- [15] S. Mohammad, D.T. Far, and E. Tahami." Speech analysis for diagnosing Parkinson's disease using genetic algorithm and support vector machine." *Journal of Biomedical Science and Engineering* 2014 (2014).
- [16] F.J.M. Shamrat, M. Asaduzzaman, A.S. Rahman, R.T.H. Tusher, Z. Tasnim" A comparative analysis of Parkinson disease prediction using machine learning approaches." *International Journal of Scientific and Technology Research* 8(11) (2019) 2576-2580.
- [17] Z. K. Senturk," Early diagnosis of Parkinson's disease using machine learning algorithms." *Medical hypotheses* 138 (2020): 109603.
- [18] B., Minja, et al" Artificial intelligence for assisting diagnostics and assessment of Parkinson's disease—A review." *Clinical neurology and neurosurgery* 184 (2019): 105442.