# Prediction of Freezing of Gait in Parkinson's Disease Patients Using Machine Learning

Sruthy Ajith
Department of Computer Science and Engineering
College of Engineering Trivandrum
Kerala, India
sruthyajith2210@gmail.com

***Abstract*— Parkinson's disease is a complex neurological disorder that severely impacts a person's motor control, leading to various movement difficulties. Among the most challenging symptoms being experienced is the freezing of gait, where individuals suddenly and temporarily struggle to start or maintain movement, causing them to freeze in place. Detecting and predicting these FOG episodes early is crucial for prompt intervention and improved patient results. This project introduces a machine learning method for predicting FOG episodes using 3D accelerometer data gathered from the lower backs of Parkinson's patients in both lab and home environments. The research uses two datasets: tDCSFOG, with 833 unique participants, and DeFOG, with 137 unique participants. The goal is to find the most accurate model for predicting FOG and to estimate the probability of occurrence of four specific events: Start Hesitation, Turn, Walking, and Normal movement, and to identify which event is most and least likely to occur among the subjects. Five machine learning models were used : LightGBM, Random Forest, Decision Tree, Gradient Boosting Classifier, and CatBoost. These models were assessed according to their accuracy, recall, precision, and f1-score. LightGBM performed the best, with an accuracy rate of 98.11%, a f1-score of 98.10%, a precision rate of 98.12%, and a recall rate of 98.11%. Therefore, LightGBM was chosen for the final probability prediction.The results show that 'Turn' is the most frequently occurring event, while 'Walking' is the least frequently occurring event among the subjects. The aim of this research is to improve the prediction and understanding the instances of FOG in patients, thereby contributing to better management and care for those affected by this challenging symptom.**

**Keywords— Freezing of Gait, Parkinson's Disease, Machine Learning, Start Hesitation, Turn, Walking, Light Gradient Boosting Machine.**

## I. INTRODUCTION

Parkinson's disease is a progressive neurological degenerative disease that significantly affects the quality of life for those who have it. It is characterized by a variety of motor symptoms. Among these symptoms, FOG stands out as a particularly challenging and debilitating aspect. It refers to sudden and temporary episodes where people with PD struggle to begin or maintain a steady gait, leading to a transient halt in movement. Recognizing the urgent need for advanced predictive tools in managing Parkinson's disease, this study focuses on leveraging machine learning methods for precise FOG prediction. Through the development of a sophisticated model, this research aims to analyze data collected from wearable sensors, providing insights into the temporal patterns and characteristics associated with FOG episodes. The use of algorithms can enhance our understanding of FOG in Parkinson's disease patients, ultimately paving the way for more timely interventions and personalized treatment strategies. The study employs cutting-edge technology, utilizing data from wearable 3D lower back sensors to train an algorithm. The model focuses on predicting FOG episodes by analyzing time-series data recorded during specific protocols and patient characteristics. With a particular emphasis on categorizing and differentiating types of freezing episodes, such as Start Hesitation, Turn, and Walking, the research seeks to offer a nuanced understanding of the challenges posed by each category. Utilizing diverse datasets obtained from laboratory environments and the residences of subjects undergoing FOG-provoking protocols, this study seeks to advance the creation of a dependable predictive tool for FOG. The potential impact of this research extends to enhancingmanagement and care strategies in PD. Within the field of healthcare, along with technology convergence, this research project stands at the forefront, aiming to bridge gaps in understanding and treatment options for Parkinson's patients. The anticipation of FOG in PD arises as a pivotal area of study, holding the promise to revolutionize treatment and support for affected ones.

Previous research has primarily focused on detecting Parkinson’s disease using machine learning techniques without addressing specific symptoms. In contrast, my work aims to predict the occurrence of a particular symptom, freezing of gait , in Parkinson’s disease patients. This involves a comprehensive comparison of different machine learning models to identify the best-performing model for predicting FOG. The selected model predicts the probability of the occurrence of four specific events in FOG: start-hesitation, turning, walking, and normal movement. By predicting FOG and its associated events, my work aims to enhance patient management and intervention strategies tailored to the unique movement challenges faced by Parkinson’s disease patients.

## II. LITERATURE SURVEY

In this study [1], researchers address the growing prevalence of Parkinson’s disease by proposing a method that integrates machine learning with wearable sensors to identify optimal exercises for effective PD detection. The study outlines 15 common diagnostic tasks used in clinical practice and employs ML techniques on data collected via a compact wearable sensor node. This approach identifies three exercises with the highest discriminative power, achieving an ROC AUC of 0.9 for each task. The methodology tackles key challenges, including PD stage determination, classification between healthy and PD subjects, and differentiation between PD and other neurological disorders. Exercises are categorized into gross motor, clinical evaluation, fine motor, and tremor at rest classes. Utilizing SensorTile1 IMU modules for wireless data collection, the study involves data preprocessing, dimensionality reduction, and feature extraction, employing algorithms such as SVM, RF, LR, NB, kNN, Boosted Trees, and SE. The study recommends two exercises as efficient for PD diagnosis and suggests a model based on these exercises. Despite limitations related to wireless communication and patient discomfort, this system has potential as a second-opinion tool to aid accurate diagnosis and assist doctors in recommending suitable diagnostic exercises, although it shows a lower accuracy of 65% in some cases.

This research [2] explores the application of ML with smartphone-based sensors to detect PD through a 20-step walking test. The study involves 103 participants, with data from 58 subjects used in the analysis. It compares three feature selection methods—mRMR, SFS, and SBS—alongside nine classification algorithms. For individual step accuracy, SFS combined with NB achieves the highest accuracy at 75.3%. In overall subject classification, the kNN algorithm shows the best result at 84.5%. The study highlights the potential of smartphone-based systems for monitoring PD symptoms, offering a cost-effective and accessible solution. However, challenges include transitioning from controlled clinic data to real-world conditions, potential misclassifications due to medication effects, age-related health issues, and the need for a robust walking detection algorithm suitable for uncontrolled environments. Further research is needed to adapt and validate these methods in home settings, aiming to provide objective insights into PD symptoms for personalized patient care.

This paper [3] proposes a novel system integrating wearable sensors and AI to address the challenges of diagnosing PD, particularly in early-stage patients. Data from 56 individuals, including those with PD tremor and other tremor-related disorders, was collected using SensorTile wearable sensors. The study focuses on tremor and bradykinesia analysis, incorporating exercises based on the MDS-UPDRS scale. ML methods such as RF and SVM were used for feature extraction and classification. The study reveals challenges in visually distinguishing different tremor types and employs statistical analysis to highlight similarities in amplitude and frequency distributions. Despite an accuracy rate of 85%, the inclusion of bradykinesia features significantly improves PD diagnosis accuracy, achieving an f1-score of 0.88 for exercise 3. The findings suggest potential applications in telemedicine, providing a valuable second-opinion tool for clinicians. Future research should address limitations such as sample size and enhance the diagnostic accuracy of the proposed system.

This research [4] presents a method for PD diagnosis using the SensHand V1 motion sensor to capture upper limb data during specific tasks from three groups: healthy subjects, individuals with IH, and PD patients. The study involves 30 participants per group and analyzes temporospatial and frequency data, yielding 48 parameters per side. Feature selection and classification are conducted using RF, NB, and SVM across three datasets, resulting in a high accuracy of 97% in distinguishing healthy subjects from PD patients. The study demonstrates the system's potential for supporting objective PD diagnosis but identifies challenges in distinguishing IH from healthy subjects and PD patients. Future research should address sample size limitations, consider PD severity, and establish normative data for clinical validation. Integrating kinematic evaluation with smell assessment tests could create a non-invasive, low-cost procedure for analyzing individuals at risk of PD development.

This study [5] introduces a hierarchical architecture for automated PD symptom assessment, focusing on early and heterogeneous stages of the condition. The wPAD wearable, featuring inertial sensors, captures motor features during different tasks. The architecture comprises three layers: HMMs identify abnormality classes specific to motor tasks; ML classifiers, including FFNN, SVM, and RF, quantify single-symptom severity; and regression models integrate severity estimations across symptoms to generate a MSS score. The evaluation involves 15 PD patients and 15 healthy subjects, demonstrating accurate symptom-specific movement recognition and high classification accuracy for tremor and bradykinesia. The generated MSS scores show significant correlations with clinical UPDRS scores. The study highlights the need for further validation with larger cohorts to enhance generalizability.

This paper [6] develops an auxiliary diagnostic system for PD using limb data from 100 participants. The system uses wrist sensors to gather movement data, a genetic algorithm random forest classifier for PD classification, and a user-friendly GUI for neurologists. The GA optimizes RF parameters, achieving a high classification accuracy of 94.4% through leave-one-out cross-validation. Different movement tasks are explored, revealing that the alternating hand movement task alone yields an accuracy of 92.9%. The study emphasizes the importance of having sensors on both wrists for a comprehensive assessment. Despite promising results, limitations include the small scale of training data and the challenge of distinguishing PD from similar diseases. Future research should integrate hardware for real-time monitoring and expand data for advanced ML algorithms, contributing to the development of an effective and user-friendly PD assessment system with clinical applications.

This study [7] tackles the challenge of detecting freezing of gait in PD using a RNN and a single waist-worn tri-axial accelerometer in real-life environments. The dataset includes acceleration signals from 21 PD patients in both ON and OFF states during scripted home activities. Four data representations are evaluated: hand-crafted features, MFCCs, spectral data based on FFT, and contextual windows. The study employs LOSO cross-validation to train and evaluate a generic FOG detection model. Comparisons with supervised methods like RF, AdaBoost, SVM, DNN with CNN, and CNN-LSTM show the proposed model's superior performance, achieving an AUC of 0.939 with 87.1% sensitivity and specificity, and reducing EER from 17.5% to 12.7%. Despite limitations like a small dataset and imbalanced data distribution, the results suggest a promising direction for enhancing FOG detection systems. Further research is needed to address these constraints and improve generalizability.

This paper [8] develops a predictive model for FOG in PD using acceleration signals from 12 patients during FOG-provoking walking tasks. ML techniques, specifically AdaBoost, are used to build FOG models. The models are trained with standard and impaired sets using uniform and individualized labeling. Model D shows better performance, with an accuracy of 82.7% in subject-specific evaluation and 77.9% in patient-independent testing. The research uses a stage-based classification method to reduce the time to 0.93 seconds. Limitations include a small dataset and lower sensitivity in patient-independent models, suggesting the need for larger datasets and further refinement to enhance generalizability.

This study [9] proposes a system for detecting FOG in PD using motion signals from devices and ML techniques. The study involves 24 Parkinson’s patients and collects acceleration signals from sensors placed on seven body parts. Data preprocessing includes handling missing values and filtering, followed by stepwise feature extraction to separate acceleration signals as body, gravity, and FOG data. Extracted features are used to train classifiers, with LGBM emerging as the most effective, achieving an accuracy of 88.62%, a sensitivity of 87.69%, and an AUC of 88.49%. Despite promising results, further research is needed to validate the method on a larger dataset, explore other data types, and compare its performance against existing approaches.

This paper [10] introduces a system for early PD detection, emphasizing the importance of timely diagnosis. The study employs ML techniques including object recognition, NLP, speech identification, and machine vision, with a LightGBM classifier for PD detection. The dataset comprises 195 records with 22 columns from Kaggle. The methodology involves data preprocessing, noise handling, and normalization. The LightGBM classifier achieves a high accuracy of 97.56%, outperforming other boosting algorithms. However, the study lacks exploration of limitations, dataset representativeness, and ethical considerations in healthcare ML. Future work should focus on incorporating gait analysis or other motor symptoms to enhance the system's clinical impact.

## III. DATASET USED

The proposed study utilizes data gathered from a three-dimensional accelerometer placed on the lumbar region of participants. This data comes from two primary datasets: tDCSFOG and DeFOG. These datasets include series collected both in controlled laboratory environments and at participants' residences, where participants followed a protocol designed to provoke FOG episodes. Both datasets include:

1. Time: An integer representing the timestep, recorded at a frequency of 128 Hz for the tDCSFOG dataset and 100 Hz for the DeFOG dataset.
2. AccV, AccML, and AccAP: Acceleration readings along the vertical, medial-lateral, and anterior-posterior axes, respectively.
3. Walking, Start Hesitation, Turn: Flags indicating the presence of each respective event type.

Both datasets come with their respective metadata files, named tdcsfog_metadata.csv and defog_metadata.csv. These files uniquely identify each data series by participant, session, assessment, and medication status. Additionally, the events.csv file includes metadata for every episode in all datasets, specifying event occurrence, timing (columns for start and end times in seconds), type, and movement characteristics.

## IV. MACHINE LEARNING MODELS USED

### *A. Random Forest*

Random forest is a popular ensemble learning method used for classification and regression tasks. It builds multiple decision trees during training and merges their results to improve accuracy and control overfitting. The method involves bootstrap sampling, where multiple subsets of the training data are created by sampling with replacement. For each subset, a decision tree is independently trained, and at each node, a random subset of features is considered for splitting to introduce further randomness. In the aggregation step, the final prediction for classification is determined by the mode of the classes predicted by individual trees, and for regression, by the average of the predictions. They are valued for their robustness, ability to handle large datasets, and capability to manage both numerical and categorical data effectively.

### *B. Decision Tree*

A decision tree is a supervised learning algorithm used for both classification and regression tasks. It models decisions and their possible consequences as a tree-like structure of nodes, where each internal node represents a "test" on an attribute, each branch represents the outcome of the test, and each leaf node represents a class label (in classification) or a continuous value (in regression). The process starts at the root node and involves splitting the dataset into subsets based on the attribute that results in the highest information gain or lowest impurity. At each node, the algorithm evaluates the attributes to determine the optimal split, often using metrics like gini impurity or information gain for classification, and variance reduction for regression. To avoid overfitting, the tree may be pruned by removing branches that have little importance. They are favored for their simplicity, interpretability, and ability to handle both numerical and categorical data.

### *C. Gradient Boosting Classifier*

A gradient-boosting classifier is a powerful ensemble learning method used for classification tasks. It builds a series of decision trees, where each tree corrects the errors of the previous one, to create a strong predictive model. The process begins with an initial model, often a simple one like predicting the mean of the target values. In each iteration, a new tree is trained on the residual errors (the difference between the actual and predicted values) of the current model. The model adjusts to minimize the loss function by learning from these residuals. The predictions of the new tree are then combined with those of the previous trees to form an improved overall prediction. They are highly effective for their ability to improve prediction accuracy, handle various types of data, and provide robust performance in a wide range of applications.

### *D. CatBoost*

Categorical Boosting is a state-of-the-art gradient boosting algorithm particularly effective for handling categorical features in both classification and regression tasks. Developed by Yandex, CatBoost aims to improve prediction accuracy and reduce overfitting while providing fast training. It converts categorical features into numerical representations using an efficient encoding method, preserving the natural order of the data. The algorithm employs ordered boosting to prevent overfitting by using an unbiased subset of the data in each iteration. It builds an ensemble of decision trees trained on the residuals of previous trees to minimize the loss function and uses symmetric (oblivious) trees, where each level splits on the same feature, for faster training and prediction. It is highly valued for its efficient handling of categorical data, robustness, and superior performance across various applications.

### E. Light Gradient Boosting Machine

LightGBM , a highly efficient gradient-boosting framework developed by Microsoft, prioritizing speed and performance. Particularly adept with large datasets and high-dimensional data, It offers rapid training and low memory usage. Its working mechanism involves histogram-based decision tree learning, where continuous feature values are binned into discrete intervals, enhancing computational speed and reducing memory requirements. It adopts a leaf-wise growth strategy, unlike traditional level-wise tree growth, by selecting the leaf with the maximum split gain to expand, resulting in deeper trees and potentially improved accuracy. Employing gradient boosting, it sequentially constructs an ensemble of decision trees, each aimed at rectifying the errors of its predecessors by minimizing a specified loss function. It is valued for its efficiency, scalability, and aptitude for managing extensive and complex data, and is widely favored across diverse practical applications demanding rapid and precise predictive modeling.

## V. METHODOLOGY

The workflow presented in Fig. 1 outlines the sequence of operations within the system, covering several essential stages. First, dataset collection involves gathering data recorded under various conditions. During data preprocessing, tasks include addressing missing data, converting categorical data into numerical form, and standardizing acceleration units. Feature engineering entails crafting novel features or altering existing ones, while feature selection identifies the most pertinent features for the training process. In model selection, it involves choosing the best-suited model. In model training, it undergoes training with the pre-processed data. Model evaluation involves assessing the performance of the testing data. Finally, the model predicts the likelihood of all four events occurring and identifies the most and least frequent events among the subjects.

### A. Dataset Collection

The proposed work uses data collected from the 3D accelerometer on the lower back of subjects. The data series encompass two datasets collected under distinct circumstances: the tDCSFOG (tdcsfog) dataset, which comprises data series collected in a laboratory setting as subjects completed a FOG-provoking protocol, and the DeFOG (defog) dataset, consisting of data series collected in the subject's home under similar FOG-provoking conditions.

### B. Dataset Preprocessing

Data preprocessing involves preparing and cleaning the dataset for subsequent analysis. This step addresses various aspects, starting with handling missing data by employing imputation methods like mean or median to replace null values. The treatment of categorical features follows, where these variables are transformed into a numerical format using methods like one-hot encoding. Data normalization is then applied to scale and standardize numerical features, ensuring uniformity and mitigating the impact of different scales on model performance.

### C. Feature Engineering

It is a pivotal process in enhancing the forecasting abilities of a model, particularly in the context of detecting events using acceleration data. This process involves introducing new features derived from the existing ones to capture essential patterns and information. In this scenario, statistical calculations are applied to time series data using a sliding window with a predetermined size of 1024. The newly created features include metrics such as the sum of acceleration, cumulative sum, and rolling sum. These engineered features provide a more comprehensive representation of the underlying acceleration data, enabling the model to better discern and understand relevant event patterns, contributing to improved predictive accuracy and model performance.

### D. Feature Selection

Feature selection is the process of identifying and selecting the most relevant features from a dataset to use in a machine learning model. It is a crucial step that enhances model performance by reducing dimensionality, removing irrelevant data, and improving the learning process. This process utilizes the Mutual Information (MI) technique to identify the most informative features. A higher MI score indicates that a feature provides more information about the target variable. The features are evaluated based on their scores. To prioritize the most relevant features, the MI scores are sorted in descending order. A threshold value is set to select the top features, ensuring that only those with significant information gain relative to the target variable are chosen. This method ensures that the selected features contribute meaningfully to the model's predictive power.

### E. Model Selection

In this project, we selected five ML models to establish a resilient predictive framework for the disease, each contributing unique strengths. This multi-class classification problem involves LightGBM for its high efficiency, speed, and capacity to handle large-scale and imbalanced datasets, which is crucial for medical data. Random Forest is included for its robustness and versatility in handling both numerical and categorical data. The decision tree offers simplicity, which is essential for clear decision-making paths in medical diagnostics. The gradient-boosting classifier is chosen for its high predictive accuracy and ability to capture complex patterns, which are crucial for modeling the multifaceted symptoms of the disease. Lastly, CatBoost excels at handling categorical features and preventing overfitting, ensuring efficient and accurate predictions even with smaller datasets. Together, these models form a comprehensive toolkit for developing robust predictive models for this critical application, offering reliable and accurate solutions for medical professionals diagnosing and treating disease.

### F. *Model Training*

Model training is the stage where the data is fed into a machine learning algorithm to help it learn and identify optimal values for all involved attributes. Subsequently, the dataset is partitioned into two sets, typically allocating 80% for training and 20% for testing. During this phase, all five models undergo hyperparameter tuning to enhance their performance. This tuning entails adjusting parameters that dictate the training process, set before the learning begins and not learned from the data. To achieve this, we employ the GridSearchCV method, which systematically explores various parameter combinations for each model. It conducts an exhaustive search over a specified parameter grid, assessing model performance through cross-validation. This process involves splitting the data into partitions, educating the model using selected partitions, and then evaluating its performance on the remaining subsets to ensure the selected parameters offer the best performance on unseen data.

### G. *Model Evaluation*

Model evaluation is a critical stage in assessing a model's performance and its capacity to provide appropriate predictions. Evaluation criteria, such as recall, accuracy, f1-score, and precision, are integral to this process. Accuracy as in (1) reflects the ratio of correct predictions to the total number of input samples. Precision as in (2) quantifies the proportion of accurate positive predictions among all positive predictions made. Recall as in (3) determines the ratio of actual positives correctly identified by the model. The F1-score as in (4), a blend of precision and recall, aims to effectively balance these two metrics for each class. A confusion matrix serves as a tabular representation of a model's performance. It contrasts the predicted labels with the actual labels and showcases the count of true positives, true negatives, false positives, and false negatives. Additionally, a classification report provides a comprehensive overview of the primary metrics derived from the confusion matrix. It conducts a detailed assessment of the model's performance by computing metrics for each class, furnishing valuable information regarding the model's effectiveness across different classes.

$$\text{Accuracy} = \frac{TP+TN}{\text{TP + TN + FP + FN}} \qquad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \qquad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \qquad (3)$$

$$\text{F1-Score} = 2.\ \frac{Precision\ .\ Recall}{Precision + Recall} \qquad (4)$$

## VI. RESULTS

### A. *Performance Evaluation*

The findings presented in Table 1 showcase LightGBM's remarkable accuracy rate of 98.11%, surpassing all other models in performance. The selected evaluation metrics confirm LightGBM's superiority in FOG prediction without a doubt.

### B. *Prediction*

The model predicts the probability of all four events—start hesitation, turn, walking, and normal—for each subject in the test set. Additionally, it identifies the most and least frequent events among the majority of subjects, with 'Turn' being the most common and 'Walking' being the least common. The probability distribution is illustrated in Fig. 2. The classification report is illustrated in Fig. 3 and the confusion matrix in Fig. 4.

### C. *Figures and Tables*

Table 1 : Comparison Table

| Model | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| LightGBM | 98.11 | 98.12 | 98.11 | 98.10 |
| CatBoost | 96.81 | 96.82 | 96.81 | 96.79 |
| Gradient Boosting | 94.45 | 94.42 | 94.45 | 94.38 |
| Decision Tree | 93.25 | 93.42 | 93.25 | 93.23 |
| Random Forest | 91.15 | 91.37 | 91.15 | 91.05 |

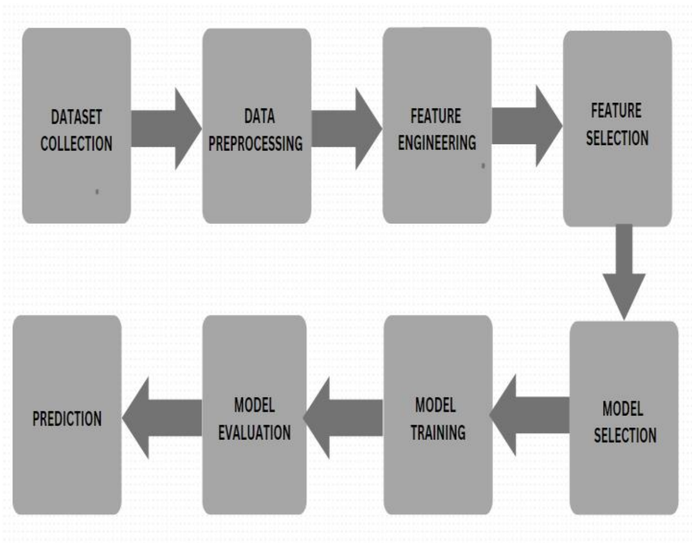


Fig. 1. Framework of Implemented System

| | ID | StartHesitation | Turn | Walking | Normal | PredictedClass |
|---|---|---|---|---|---|---|
| 0 | 1 | 0.006320 | 0.000015 | 0.003695 | 0.989969 | Normal |
| 1 | 2 | 0.008571 | 0.000025 | 0.991258 | 0.000147 | Walking |
| 2 | 3 | 0.014589 | 0.000019 | 0.985180 | 0.000212 | Walking |
| 3 | 4 | 0.077371 | 0.000034 | 0.922269 | 0.000326 | Walking |
| 4 | 5 | 0.273970 | 0.000110 | 0.006994 | 0.718925 | Normal |
| ... | ... | ... | ... | ... | ... | ... |
| 614475 | 614476 | 0.940030 | 0.000043 | 0.059676 | 0.000251 | StartHesitation |
| 614476 | 614477 | 0.000010 | 0.999980 | 0.000010 | 0.000000 | Turn |
| 614477 | 614478 | 0.001489 | 0.000106 | 0.998362 | 0.000043 | Walking |
| 614478 | 614479 | 0.121772 | 0.000084 | 0.877999 | 0.000145 | Walking |
| 614479 | 614480 | 0.012211 | 0.000010 | 0.004209 | 0.983569 | Normal |

614480 rows × 6 columns

Fig. 2. Probability Prediction

Classification Report:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Normal | 0.98 | 0.95 | 0.96 | 153620 |
| StartHesitation | 1.00 | 1.00 | 1.00 | 153620 |
| Turn | 0.96 | 0.98 | 0.97 | 153620 |
| Walking | 0.99 | 1.00 | 1.00 | 153620 |
| accuracy | | | 0.98 | 614480 |
| macro avg | 0.98 | 0.98 | 0.98 | 614480 |
| weighted avg | 0.98 | 0.98 | 0.98 | 614480 |

Fig. 3. Classification Report

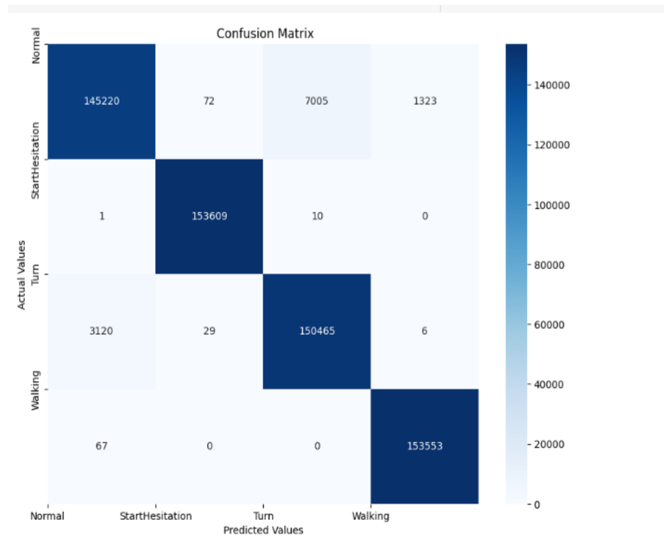


Fig. 4. Confusion Matrix

## VII. CONCLUSION

In summary, this project has effectively developed a machine learning method for forecasting FOG incidents utilizing 3D accelerometer data obtained from the lower back of individuals with Parkinson's disease in both laboratory and home settings. By leveraging two datasets, tDCSFOG and DeFOG, and assessing five machine learning models, LightGBM emerged as the top performer, achieving an accuracy rate of 98.11%, precision rate of 98.12%, recall rate of 98.11%, and f1-score of 98.10%. The model forecasts the likelihood of four events—start-hesitation, turn, walking, and normal—with Turn being the most probable and Walking the least likely to occur. These findings carry substantial implications for enhancing the management and care of patients experiencing FOG episodes by facilitating proactive interventions that enhance mobility, security, and the overall standard of living.

The future scope of this project is vast and exciting, with numerous opportunities for growth and expansion, including continuously refining the predictive model through new data sources and advancements in machine learning, integrating models into clinical decision support systems to aid healthcare providers, exploring other motor symptoms and complications of PD, developing personalized predictive models that account for individual differences, investigating the utilization of wearable devices and real-time monitoring systems to enable continuous tracking of FOG episodes, and providing timely interventions.

## REFERENCES

[1] Aleksandr Talitckii , Anna Anikina , Ekaterina Kovalenko , Aleksei Shcherbak , Oscar Mayora , Olga Zimniakova , Ekaterina Bril , Maxim Semenov and Dmitry V. Dylov "Defining optimal exercises for efficient detection of Parkinson's disease using machine learning and wearable sensors." in IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-10, 2021, Art no. 2512010, doi: 10.1109/TIM.2021.3097857.

[2] Juutinen, Milla, et al. "Parkinson’s disease detection from 20-step walking tests using inertial sensors of a smartphone: Machine learning approach based on an observational case-control study." PloS one 15.7 (2020): e0236258.

[3] Ekaterina Kovalenko, Anna Anikina, Olga Zimniakova, Maksim Semenov, Ekaterina Bril, Aleksei Shcherbak, Dmitry V. Dylov et al. "Avoiding misdiagnosis of Parkinson’s disease with the use of wearable sensors and artificial intelligence." IEEE Sensors Journal 21.3 (2020): 3738-3747.

[4] Filippo Cavalloa, Alessandra Moschettia, Dario Espositoc, Carlo Maremmanib, Erika Rovini "Upper limb motor pre-clinical assessment in Parkinson's disease using machine learning." Parkinsonism & related disorders 63 (2019): 111-116.

[5] C. Wang, L. Peng, Z. -G. Hou, Y. Li, Y. Tan and H. Hao, "A Hierarchical Architecture for Multisymptom Assessment of Early Parkinson’s Disease via Wearable Sensors," in IEEE Transactions on Cognitive and Developmental Systems, vol. 14, no.4,pp.1553-1563,Dec.2022,doi:10.1109/TCDS.2021.3123157.

[6] M. Chen, Z. Sun, F. Su, Y. Chen, D. Bu and Y. Lyu, "An Auxiliary Diagnostic System for Parkinson’s Disease Based on Wearable Sensors and Genetic Algorithm Optimized Random Forest," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 2254-2263, 2022, doi: 10.1109/TNSRE.2022.3197807.

[7] Sigcha L, Costa N, Pavón I, Costa S, Arezes P, López JM, De Arcas G. Deep Learning Approaches for Detecting Freezing of Gait in Parkinson's Disease Patients through On-Body Acceleration Sensors. Sensors (Basel). 2020 Mar 29;20(7):1895. doi: 10.3390/s20071895. PMID: 32235373; PMCID: PMC7181252.

[8] Zhang, W. Yan, Y. Yao, J. B. Ahmed, Y. Tan and D. Gu, ”Prediction of Freezing of Gait in Patients With Parkinson’s Disease by Identifying Impaired Gait Patterns,” in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 28, no. 3, pp. 591-600, March 2020, doi: 10.1109/TNSRE.2020.2969649.

[9] Xihua Liu, Shengdi Chen, Kang Ren, Jin Zhao "A Freezing of Gait Detection System Based on Stepwise Feature Extraction Method," 2020 Chinese Control And Decision Conference (CCDC), Hefei, China, 2020, pp. 2060-2065, doi: 10.1109/CCDC49329.2020.9164331.

[10] G. V. D. Kumar, V. Deepa, N. Vineela and G. Emmanuel, "Detection of Parkinson’s disease using LightGBM Classifier," 2022 6th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2022, pp. 1292-1297, doi: 10.1109/ICCMC53470.2022.9753909.