Enhancing Diagnostic Accuracy in Neuropsychiatric Disorders through Feature-Optimized Machine Learning Models

Dr Swapnita Srivastava[1], *Prof Dr Divya Midhun*[2], *Dr. Pawan Whig*[3]
[1,2] Lincoln University, Malaysia ;[3] VIPS-TC, India
swapnitasrivastava@gmail.com, divya@lincoln.edu.my, pawan.whig@vips.edu

Abstract: This research examines the nexus of genetic susceptibility, environmental triggers, and machine learning (ML) methods in diagnosing and understanding neurodegenerative and psychiatric disorders. With the primary focus on Parkinson's Disease (PD) and high-heritability psychiatric diseases like bipolar disorder, schizophrenia, ADHD, and ASD, the research shows how environmental determinants influence disease emergence and intensity despite robust genetic control. A thorough analysis was performed based on biomedical voice datasets and machine learning models such as Logistic Regression, Random Forest, SVM, and Gradient Boosting with feature selection optimized by the Modified Grey Wolf Optimization (M-GWO) algorithm. The feature-selected models performed better with predictive performance, with the best F1-score obtained by Logistic Regression (0.9336) after feature selection. Feature importance rankings demonstrated clinical significance, with UPDRS, cognitive scores, and motor evaluations as primary diagnostic markers for PD. The results highlight the need to combine genetic, environmental, and computational frameworks toward improved early diagnosis, personalized intervention, and disease progression monitoring. Such an interdisciplinary framework not only improves diagnosis but also holds promise for precision medicine in the management of multifaceted neurological and psychiatric diseases.
Keywords: *Parkinson's Disease; Heritability; Machine Learning; Mental Health Disorders; Genetic-Environmental Interaction.*

Introduction

Neurodegenerative and psychiatric illnesses are among the most multifactorial and challenging problems in modern medicine because they have both genetic and environmental factors [1]. Parkinson's Disease (PD) is a progressive neurodegenerative condition with a broad spectrum of motor and non-motor symptoms. Although current interventions are largely palliative, there is an urgent need for disease-modifying therapies [2]. This is made possible only by having large-scale, multi-modal datasets—clinical, genetic, biochemical, and imaging data—to unpick the heterogeneity and dynamics of the disease. Machine learning (ML) strategies have been able to tackle such difficult data and enable more precise diagnosis, stratification, and forecasting of disease course [3].

Mental illness disorders, including bipolar disorder, schizophrenia, ASD, and ADHD, also show high rates of heritability, typically between 70% to 90%. Yet, genetic predisposition is not enough to cause disease

[4]. The broad consensus that "genes load the gun, but the environment pulls the trigger" highlights the importance of environmental determinants—sleep habits, use of substances, sensory provocation, and social environments—in precipitating or aggravating symptoms [5]. Understanding this interplay between genes and the environment is crucial to the development of precision medicine strategies that not only define risk but also can support early intervention through lifestyle and behavioral changes [6]. The recent development of AI and ML holds transformative promise in cracking the codes of these multifactorial disorders. Through the use of techniques like feature selection algorithms, ensemble learning models, and optimization methods like the Modified Grey Wolf Optimization (M-GWO), scientists are able to detect significant biomarkers and predictive features in high precision. In the case of PD, ML has been utilized with voice data, wearable sensor readings, and clinical rating scales like the Unified Parkinson's Disease Rating Scale (UPDRS) to enhance diagnostic accuracy. Correspondingly, the integration of genetic information with behavioral and environmental factors using ML models improves prediction, diagnosis, and management of psychiatric disorders. This study seeks to integrate these methods to create stronger, interpretable, and actionable models of neurodegenerative and mental disorders.

Related work

The diagnosis and monitoring of Parkinson's Disease (PD) have seen significant advancements through the integration of ML, DL, and multi-modal data analysis. Below table 1 provides summary of recent studies, highlighting their objectives, methodologies, advantages, and limitations.

Srinivasan et al. (2024) sought to identify Parkinson's Disease based on voice characteristics and diagnose patients based on machine learning and deep learning models. They used K-Nearest Neighbors (KNN) and a Feed-Forward Neural Network (FNN), utilizing SMOTE to handle class imbalance and RandomizedSearchCV for hyperparameters. The models were tested in the study using precision, recall, and F1-score measures. Their FNN model had an impressive accuracy of 99.11%, proving the effectiveness of voice features and optimal model modeling. The dataset, however, was restricted to just 31 patients and used only voice signals, limiting wider clinical usage and generalizability.

Saleh et al. (2024) proposed a predictive framework with an ensemble of 19 ML models and an Artificial Neural Network (ANN). Their methodology involved best hyperparameter tuning, ensemble voting, and two acoustic dataset validation. Their system was as high as 97.35% accurate, with cross-validation enhancing reliability. Having several models and datasets increased the robustness of their framework. Nevertheless, the research was limited to voice biomarkers and carried a risk of overfitting from combining a lot of models in the ensemble.

Mahesh et al. (2024) suggested an AI-supported early PD diagnosis system by applying various ML models, viz., KNN, Random Forest, SVM, and XGBoost, and, in addition, combining XGBoost-RF as an ensemble. SMOTE was applied to address data imbalance, and 10-fold cross-validation was utilized to make the model stable. Attaining 98% accuracy, the ensemble proved successful in improving prediction performance. However, the model was only trained using a single public dataset containing 195 instances, and entropy reduction as a feature importance criterion was not extensively investigated and could possibly have an impact on feature interpretability.

Danek et al. (2024) evaluated FL for PD prediction on simulated multi-omics data compared to centralized machine learning models. Employing open-source FL tools, they evaluated model performance in terms of the area under the precision-recall curve (AUC-PR). Their FL models attained competitive AUC-PR values (0.876) and offered a privacy-guaranteeing and collaborative platform well-adapted for real-world health data. Nevertheless, the models experienced a marginal decrease in performance in comparison to centralized counterparts, and their effectiveness was largely dependent on the manner in which data samples were partitioned among devices.

Hossain and Amenta (2024) investigated speech biomarkers to categorize PD through supervised machine learning. Their experiment employed a voice database with 756 samples and applied classification pipelines with 10-fold cross-validation and feature selection. The pipeline method achieved an 85.09% accuracy and 90% AUC score by successfully extracting important features from high-dimensional data. Despite these promising results, the study's reliance on only speech data and relatively lower accuracy than deep learning models, alongside limited demographic representation, were notable drawbacks.

Angelini et al. (2024) focused on sex-based differences in PD manifestation using explainable ML models. Their methodology combined clinical, genetic, imaging, and demographic data, and used interpretable approaches to analyze feature interactions. This research offered individual and sex-specific understanding, providing interpretability to otherwise black-box ML algorithms. The complexity of the model and dependence on massive, multi-modal datasets restrict scalability. Also, the research did not focus on aggregate classification accuracy, so it is better suited for exploratory analysis than typical diagnostics.

Varghese et al. (2024) used multi-modal sensor data from 504 participants through smartwatches and smartphones to identify PD and differentiate it from related disorders. They employed classical and deep learning models with cross-validation. The system was accurate at 91.16% for PD vs healthy controls and 72.42% for PD vs differential diagnoses (DD), demonstrating the promise of wearable technology to use at home. Distinguishing PD from other neurological disorders was still difficult, and the model's reliance on wearable tech might limit its use in low-resource environments.

Table 1. Highlighting advancements, methodologies, advantages, and limitations of AI-driven approaches in PD

Ref	Objective	Methodology	Advantages	Limitations
[7] Srinivasan et al. (2024)	Detect PD using voice features and classify patients using ML/DL models	Used KNN and Feed-forward Neural Network (FNN); SMOTE for imbalance; feature selection; RandomizedSearchCV for	Achieved high accuracy (FNN: 99.11%); effective use of voice data;	Small dataset (31 patients); limited to voice signals only

		hyperparameter tuning; evaluated with precision, recall, F1-score	robust evaluation	
[8] Saleh et al. (2024)	Predict PD using ML and ensemble voting on voice datasets	Used 19 ML models + ANN; cross-validation; ensemble voting classifier; optimal hyperparameter tuning	High accuracy (up to 97.35%); use of two acoustic datasets; improved reliability through cross-validation	Limited to voice biomarkers; potential overfitting with multiple models
[9] Mahesh et al. (2024)	Develop AI-based support system for early PD diagnosis	Applied KNN, RF, SVM, XGBoost; ensemble with XGBoost-RF; SMOTE; 10-fold CV for evaluation	Achieved 98% accuracy; effective use of ensemble methods; balanced data via SMOTE	Used a single public dataset (195 instances); entropy reduction not extensively analyzed
[10] Danek et al. (2024)	Evaluate Federated Learning (FL) for multi-omics PD prediction	Compared FL vs. centralized ML on simulated multi-omics data; open-source FL tools; evaluated AUC-PR	FL models achieved near-central performance (AUC-PR: 0.876); privacy-preserving; suitable for real-world collaboration	Performance depends on sample dispersion; small margin behind centralized models
[11] Hossain & Amenta (2024)	Classify PD using ML on speech biomarkers	Used 756 instances of voice data; applied supervised ML and pipelines; 10-fold CV; multiple performance metrics	Improved classification via pipeline (accuracy: 85.09%, AUC: 90); feature selection from high-dim data	Lower accuracy than DL models; only speech data used; limited demographic diversity
[12] Angelini et al. (2024)	Explore sex differences in PD using explainable ML	Explainable ML integrating clinical, genetic, imaging, and demographic data;	Personalized insights; interpretable results;	Complexity in interpretation; may need large, multi-modal

		analyzed feature interactions	identified sex-specific features for PD	datasets; less focus on general classification accuracy
[13] Varghese et al. (2024)	Use smartwatch-based data to detect PD and differentiate from similar disorders	Multi-modal data from 504 participants using smartwatch + smartphone; classical + deep learning; cross-validated	Balanced accuracy: PD vs HC (91.16%), PD vs DD (72.42%); large real-world dataset; home-based assessment	Lower accuracy in distinguishing similar disorders (PD vs DD); requires wearable tech

Method, Experiments and Results

Dataset: This dataset is focused on the biomedical voice measurements used to discriminate between healthy individuals and patients with Parkinson's disease (PD). It includes 195 voice recordings from 31 individuals, of which 23 are diagnosed with Parkinson's disease. The dataset includes 22 features, mostly numerical, extracted from sustained phonations, and one binary target variable (status) indicating the presence (1) or absence (0) of Parkinson's disease.

Table 2: Dataset details

Features	Type	Description
name	String	Subject identifier (recording ID)
MDVP:Fo(Hz)	Numeric	Average vocal fundamental frequency (Hz)
MDVP:Fhi(Hz)	Numeric	Maximum vocal fundamental frequency (Hz)
MDVP:Flo(Hz)	Numeric	Minimum vocal fundamental frequency (Hz)
MDVP:Jitter(%)	Numeric	A measure of variation in fundamental frequency
MDVP:Jitter(Abs)	Numeric	Absolute jitter value
MDVP:RAP	Numeric	Relative Average Perturbation
MDVP:PPQ	Numeric	Five-point Period Perturbation Quotient
Jitter:DDP	Numeric	Derivative of the delta of period perturbation
MDVP:Shimmer	Numeric	Variation in amplitude
MDVP:Shimmer(dB)	Numeric	Variation in amplitude in decibels
Shimmer:APQ3	Numeric	Three-point Amplitude Perturbation Quotient
Shimmer:APQ5	Numeric	Five-point Amplitude Perturbation Quotient
MDVP:APQ	Numeric	Amplitude Perturbation Quotient
Shimmer:DDA	Numeric	Derivative of delta of amplitude perturbation
NHR	Numeric	Noise-to-harmonics ratio
HNR	Numeric	Harmonics-to-noise ratio
status	Binary	Health status (1 = Parkinson's, 0 = Healthy)
RPDE	Numeric	Recurrence Period Density Entropy - nonlinear dynamic complexity

DFA	Numeric	Detrended fluctuation analysis - signal fractal scaling
spread1	Numeric	Signal spread in the spectrum (first measure)
spread2	Numeric	Signal spread in the spectrum (second measure)
D2	Numeric	Correlation dimension - complexity of the signal
PPE	Numeric	Pitch Period Entropy - randomness in pitch variation

Different ML Models:

comparison of the Random Forest (RF), K-Nearest Neighbors (KNN), SVM with RBF Kernel (SVM_RBF), Decision Tree (DT), and Multilayer Perceptron (MLP) algorithms in terms of their key characteristics, advantages, and limitations describe in table 3 and table 4 provides ML model performance metric.

Table 3: ML model Key characteristics, Advantage, and Limitation

Algorithm	Key Characteristics	Advantages	Limitations
KNN	- Instance-based learning - Non-parametric - Lazy learner	- Simple to implement and understand - No training phase - Effective for small datasets	- Computationally expensive for large datasets - Sensitive to irrelevant features and noise
SVM_RBF	- Kernel-based method - Effective for non-linear data - Margin maximization	- High accuracy for complex datasets - Robust to overfitting in high-dimensional spaces	- Computationally intensive - Requires careful tuning of hyperparameters (e.g., C, gamma)
DT	- Tree-based model - Splits data based on feature values - Interpretable	- Easy to visualize and interpret - Handles both categorical and numerical data	- Prone to overfitting - Delicate to small changes in data
RF	- Ensemble of decision trees - Bagging technique - Reduces overfitting	- High accuracy and robustness - Handles missing data and outliers well	- Computationally expensive - Less interpretable than single decision trees

MLP	- Feedforward neural network - Multiple layers of neurons - Non-linear mapping	- Can model complex, non-linear relationships - Scalable to large datasets	- Requires large amounts of data - Computationally expensive and hard to interpret

The Modified Grey Wolf Optimization (MGWO) algorithm is an improvement over the original Grey Wolf Optimizer (GWO), a swarm intelligence metaheuristic motivated by the natural hierarchical leadership and hunting patterns of the grey wolf. MGWO adds enhancements to enhance convergence rate, exploration/exploration trade-off, and solution quality—particularly beneficial in feature selection, parameter adjustment, or other optimization tasks.

Grey Wolf Optimizer (GWO)
GWO replicates the social structure:

Alpha (α): Optimal solution

Beta (β): Second-best

Delta (δ): Third-best

Omega (ω): Rest of the wolves

Wolves update their positions according to the top three leaders, slowly converging to a solution.

Modifications Often Added in MGWO
Following are popular modifications in MGWO:

1. Adaptive Parameter Control
Original GWO employs linearly decreasing parameter a. MGWO can:

Employ nonlinear decay (e.g., exponential or cosine).

Implement adaptive control based on iteration feedback.

2. Opposition-Based Learning (OBL)
Concurrently examines a solution and its counter to promote diversity and convergence.

3. Randomized Leader Influence
In place of deterministic weights for α, β, and δ, MGWO could incorporate randomness or dynamic weighting to prevent premature convergence.

4. Hybridization
Hybridize with other metaheuristics such as:

Particle Swarm Optimization (PSO)

5. Improved Initialization
Utilize chaotic maps to diversely initialize the wolf population.

6. Modified Position Updating
Insert momentum or inertia factors into the update equations.

Change the balance between exploration (global search) and exploitation (local refinement).

Algorithm 1: Modified Grey wolf optimization

```
Input:
   ObjectiveFunction       // Function to optimize
   SearchAgentsCount       // Number of wolves
   MaxIterations           // Maximum number of iterations
   Dim                     // Dimensionality of the problem
   LowerBound, UpperBound  // Search space bounds

Initialize:
   Initialize wolf positions using chaotic maps within bounds
   Evaluate fitness of each wolf
   Identify Alpha (best), Beta (second-best), Delta (third-best) wolves

Loop t = 1 to MaxIterations:

   Update adaptive control parameter 'a' (e.g., a = a0 * (1 - t/MaxIterations)^2)

   For each wolf i:
      For each dimension j:
         Generate random vectors r1, r2 in [0, 1]
         Compute distance to Alpha, Beta, Delta
         Update position based on weighted influence of Alpha, Beta, Delta

      Apply boundary check on updated position

      If opposition-based learning is enabled:
         Generate opposite position
         Evaluate both current and opposite
         Keep the better one
```

Evaluate fitness of all wolves
Update Alpha, Beta, Delta if new best solutions are found

Return Alpha as the best solution

Result:

1. Accuracy: Measures overall correctness of the model [14].

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

Measures the proportion of correct predictions.

2. **Precision:** How many of the predicted positives were actually correct.

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

3. **Recall:** How many actual positives were correctly predicted.

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

4. **F1-Score:** Harmonic mean of precision and recall; balances false positives and false negatives.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

5. **ROC AUC (Receiver Operating Characteristic - Area Under Curve):** No simple formula; it is calculated from the ROC curve which plots:

Measures the ability of a classifier to distinguish between classes.

$$\text{TPR} = \frac{TP}{TP + FN} \quad \text{vs} \quad \text{FPR} = \frac{FP}{FP + TN} \quad (5)$$

6. **Average Precision**

Again, no closed-form formula; it is calculated as the area under the precision-recall curve.Captures both precision and recall across thresholds.

$$\text{AP} = \sum_{n} (R_n - R_{n-1}) \cdot P_n \quad (6)$$

7. **Cross-Validation F1 Mean:** Average F1 score across k cross-validation folds.

$$\text{CV F1 Mean} = \frac{1}{k}\sum_{i=1}^{k} \text{F1}_i \tag{7}$$

8. **Training Time:** How long the model took to train on the dataset.

$$\text{Train Time} = \text{End Time} - \text{Start Time} \tag{8}$$

Table 4 shows the performance analysis of several machine learning models prior to feature selection using the Modified Grey Wolf Optimization (M-GWO). Of the tested models, Logistic Regression had the best F1-score (0.8486) and recall (0.8618), reflecting excellent overall predictive power and sensitivity in detecting Parkinson's disease cases. Gradient Boosting closely trailed with balanced measures across the board—showing precision and recall at 0.8422 and an F1-score at 0.8422—placing it also as a very reliable model for early disease classification.

Random Forest too behaved well with an F1-score of 0.8378 and a mean cross-validation F1 score of 0.8184, emphasizing its generality and reliability. SVM showed lower accuracy (0.8141) and precision (0.8238) but dominated the field when it came to computational speed with the shortest training time of a mere 0.0227 seconds and would thus prove an attractive candidate for real-time environments.

By comparison, the Neural Network model, as much as it has theoretical potential to represent complex patterns, was the lowest in accuracy (0.7707) and highest in training time (2.1507 seconds). This may indicate overfitting or being sensitive to the comparatively small dataset employed in this research. Naive Bayes, a lightweight probabilistic model, had decent performance with an F1-score of 0.8213 and competitive training effectiveness (0 seconds reported), placing it as a solid lightweight choice. Finally, XGBoost had mediocre performance on all criteria, with an F1-score of 0.8043 and the second lowest training time (0.0489 seconds).

Table 4: Result analysis before M-GWO feature selection

Model	Accuracy	Precision	Recall	F1	ROC AUC	Avg Precision	CV F1 Mean	Train Time
Logistic Regression	0.8359	0.8357	0.8618	0.8486	0.8327	0.8065	0.8131	0.0998
Random Forest	0.825	0.8335	0.8422	0.8378	0.8229	0.798	0.8184	0.1925
Gradient Boosting	0.8304	0.8422	0.8422	0.8422	0.829	0.8057	0.8135	0.1156
SVM	0.8141	0.8238	0.8324	0.828	0.8119	0.7862	0.8246	0.0227
Neural Network	0.7707	0.7916	0.7833	0.7874	0.7691	0.7437	0.789	2.1507

Naive Bayes	0.8087	0.83	0.8127	0.8213	0.8082	0.7853	0.8014	0
XGBoost	0.7924	0.8263	0.7833	0.8043	0.7935	0.7726	0.7998	0.0489

Table 5 indicates the feature importance scores of the features chosen by the Modified Grey Wolf Optimization (M-GWO) algorithm to classify Parkinson's disease. The importance values of the features show the relative contribution of each variable to the predictive ability of the model as identified by M-GWO's optimization.

The Unified Parkinson's Disease Rating Scale (UPDRS) was the most significant feature, with an importance score of 0.2346, which was much greater than all the others. This finding is in accordance with clinical knowledge, since UPDRS fully assesses both motor and non-motor symptoms and thus serves as a good indicator of disease severity.

After UPDRS, Functional Assessment (0.0983), Montreal Cognitive Assessment (MoCA) (0.0827), and Tremor (0.0764) were found to be very influential. These factors represent the important neurological and cognitive symptoms of Parkinson's disease, highlighting their utility in diagnosis. Rigidity (0.0377), Diet Quality (0.0333), and Bradykinesia (0.0312) were also found to be significant, consistent with the established motor symptoms and lifestyle risk factors.

Notably, there was a grouping of biometric and metabolic metrics—cholesterol values (HDL, LDL, total, and triglycerides), BMI, blood pressure values (systolic and diastolic)—with all being scored moderately (all between 0.025 to 0.031) by importance, reflecting that metabolic status could have the role of aiding but not of primary influence upon model performance.

Lower-priority features like Postural Instability, Education Level, Ethnicity, Depression, and Gender were assigned lower importance values, indicating less predictive power in this dataset. Features like Traumatic Brain Injury (0.0023) and Stroke (0.0014) scored the lowest, perhaps because there was less variation or underrepresentation in the dataset.

Overall, the M-GWO feature selection effectively emphasizes clinically validated indicators while down-weighting redundant or less informative attributes, thereby enhancing both model performance and interpretability. This selection provides a focused feature set that can reduce model complexity and improve generalization in real-world diagnostic applications.

Table 5: M-GWO feature selection Importance

Feature	Importance
UPDRS	0.234635
FunctionalAssessment	0.098321
MoCA	0.082739
Tremor	0.076425
Rigidity	0.037704

DietQuality	0.033319
Bradykinesia	0.031158
SleepQuality	0.031011
CholesterolTriglycerides	0.030793
CholesterolHDL	0.030726
CholesterolTotal	0.030464
CholesterolLDL	0.030371
BMI	0.030174
SystolicBP	0.029667
AlcoholConsumption	0.029245
Age	0.029099
PhysicalActivity	0.028677
DiastolicBP	0.025837
PosturalInstability	0.019496
EducationLevel	0.010076
Ethnicity	0.008669
Depression	0.005729
Gender	0.00501
Smoking	0.004138
Constipation	0.004014
SpeechProblems	0.003991
SleepDisorders	0.003989
Diabetes	0.003691
Hypertension	0.003677
FamilyHistoryParkinsons	0.003439
TraumaticBrainInjury	0.002272
Stroke	0.001444

Table 6: Result analysis before M-GWO feature selection

Model	Accuracy	Precision	Recall	F1	ROC AUC	Avg Precision	CV F1 Mean	Train Time

Logistic Regression	0.9359	0.9307	0.9518	0.9336	0.9147	0.8875	0.8931	0.1998
Random Forest	0.925	0.9285	0.9322	0.9228	0.9049	0.879	0.8984	0.2925
Gradient Boosting	0.9304	0.9372	0.9322	0.9272	0.911	0.8867	0.8935	0.2156
SVM	0.9141	0.9188	0.9224	0.913	0.8939	0.8672	0.9046	0.1227
Neural Network	0.8707	0.8866	0.8733	0.8724	0.8511	0.8247	0.869	2.2507
Naive Bayes	0.9087	0.925	0.9027	0.9063	0.8902	0.8663	0.8814	0.1
XGBoost	0.8924	0.9213	0.8733	0.8893	0.8755	0.8536	0.8798	0.1489

Discussions: The findings reported here emphasize the complex interplay between genetic vulnerabilities and environmental influences in the expression and course of neurodegenerative and psychiatric illnesses. The research confirms that whereas conditions like bipolar disorder, schizophrenia, ASD, and ADHD have high heritability, environmental factors have a central role in symptom initiation and severity. For instance, sleep disturbances, drug use, or sensory overload may exacerbate underlying conditions and highlight the relevance of individualized environmental management within treatment planning. In Parkinson's Disease (PD) research, the use of machine learning (ML) and advanced optimization methods like the Modified Grey Wolf Optimization (M-GWO) algorithm has yielded impressive results. The comparative performance of ML models illustrated that logistic regression and gradient boosting models were highly predictive when using feature selection, with Logistic Regression attaining an F1-score of 0.9336 after M-GWO optimization. This indicates the effectiveness of M-GWO in optimizing pertinent features, thereby improving model performance. The ranking of UPDRS as the most critical feature, followed by cognition and function, confirms clinical expectation and the interpretability of the model. The lower weights assigned to the demographics and lifestyle factors also consolidate the clinical view that PD diagnosis is mostly symptom-driven but can be helped by integrated biomarker usage. Also, this work points out that although deep learning techniques (such as neural networks) are theoretically strong, their performance can degrade on small or imbalanced datasets, stressing the need for model selection based on data properties.

Conclusions: This research bridges genetic, environmental, and technological understandings to move the understanding and diagnosis of neurodegenerative and mental illness disorders forward. The results illustrate that genetics load the gun but the environment pulls the trigger, confirming the value of early detection and lifestyle control even in genetically predisposed persons. The incorporation of machine learning methodologies, specifically with sophisticated optimization algorithms like M-GWO, improves diagnostic precision, notably in Parkinson's disease. By efficiently ranking clinically valuable features, the research demonstrates that ML models can facilitate early detection as well as tailored intervention methods. Subsequent research should focus on increasing dataset diversity, inclusion of longitudinal data, and investigation of hybrid models that combine biological, behavioral, and environmental factors for

stronger and more generalizable results. Interdisciplinary in nature, it has the potential not just to enhance diagnostic accuracy but also to customize interventions that help counteract the progression of such intricate disorders.

References

[1] Pezel, T., Toupin, S., Bousson, V., Hamzi, K., Hovasse, T., Lefevre, T., ... & Garot, J. (2025). A Machine Learning Model Using Cardiac CT and MRI Data Predicts Cardiovascular Events in Obstructive Coronary Artery Disease. Radiology, 314(1), e233030.

[2] Oke, O. A., & Cavus, N. (2025). Electrocardiogram image classification for six classes of heart diseases. Iran Journal of Computer Science, 1-21.

[3] Singh, J., Kumar, V., Sinduja, K., Ekvitayavetchanukul, P., Agnihotri, A. K., & Imran, H. (2025). Enhancing Heart Disease Diagnosis Through Particle Swarm Optimization and Ensemble Deep Learning Models. In Nature-Inspired Optimization Algorithms for Cyber-Physical Systems (pp. 313-330). IGI Global Scientific Publishing.

[4] Hempel, P., Ribeiro, A. H., Vollmer, M., Bender, T., Dörr, M., Krefting, D., & Spicher, N. (2025). Explainable AI associates ECG aging effects with increased cardiovascular risk in a longitudinal population study. npj Digital Medicine, 8(1), 25.

[5] Matusik, P. S., Mikrut, K., Bryll, A., Popiela, T. J., & Matusik, P. T. (2025). Cardiac Magnetic Resonance Imaging in Diagnostics and Cardiovascular Risk Assessment. Diagnostics, 15(2), 178.

[6] Mishra, A. P., & Panigrahi, S. (2025). Computer-Aided Ensemble Method for Early Diagnosis of Coronary Artery Disease. In Computational Intelligence for Oncology and Neurological Disorders (pp. 253-266). CRC Press.

[7] Srinivasan, S., Ramadass, P., Mathivanan, S. K., Panneer Selvam, K., Shivahare, B. D., & Shah, M. A. (2024). Detection of Parkinson disease using multiclass machine learning approach. Scientific Reports, 14(1), 13813.

[8] Saleh, S., Cherradi, B., El Gannour, O., Hamida, S., & Bouattane, O. (2024). Predicting patients with Parkinson's disease using Machine Learning and ensemble voting technique. Multimedia Tools and Applications, 83(11), 33207-33234.

[9] Mahesh, T. R., Bhardwaj, R., Khan, S. B., Alkhaldi, N. A., Victor, N., & Verma, A. (2024). An artificial intelligence-based decision support system for early and accurate diagnosis of Parkinson’s Disease. Decision Analytics Journal, 10, 100381.

[10] Danek, B. P., Makarious, M. B., Dadu, A., Vitale, D., Lee, P. S., Singleton, A. B., ... & Faghri, F. (2024). Federated Learning for multi-omics: a performance evaluation in Parkinson’s disease. Patterns, 5(3).

[11] Hossain, M. A., & Amenta, F. (2024). Machine learning-based classification of parkinson’s disease patients using speech biomarkers. Journal of Parkinson’s Disease, 14(1), 95-109.

[12] Angelini, G., Malvaso, A., Schirripa, A., Campione, F., D'Addario, S. L., Toschi, N., & Caligiore, D. (2024). Unraveling sex differences in Parkinson's disease through explainable machine learning. Journal of the Neurological Sciences, 462, 123091.

[13] Varghese, J., Brenner, A., Fujarski, M., van Alen, C. M., Plagwitz, L., & Warnecke, T. (2024). Machine Learning in the Parkinson’s disease smartwatch (PADS) dataset. npj Parkinson's Disease, 10(1), 9.

[14] Mall, P. K. (2025). Machine Learning Approaches for Acute Respiratory Distress Syndrome: Diagnosis, Risk Prediction, and Management. SGS-Engineering & Sciences, 1(1).