Prediction of Parkinson's Disease Using Multimodal Machine Learning and Integration Techniques

Shreeya Pahune[1], Kushagra Agarwal[2], Bhaswar Ghosh[1], Nita Parekh[1*]

[1]Center for Computational Natural Sciences and Bioinformatics, International Institute of Information Technology - Hyderabad, Professor CR Rao Rd, Gachibowli, 500032, Hyderabad, India.
[2]School of Computer Science, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, 15213, Pennsylvania, USA.

*Corresponding author(s). E-mail(s): nita@iiit.ac.in;
Contributing authors: shreeya.pahune@research.iiit.ac.in; kagarwa2@andrew.cmu.edu; bhaswar.ghosh@iiit.ac.in;

Abstract

Parkinson's Disease (PD) is a severe neurodegenerative disorder which lacks reliable early diagnostic tests. In this study, we present a method for PD prediction that uses multimodal data from the Parkinson Progression Marker Initiative (PPMI), specifically clinico-demographic, biospecimen, and genetic data, to improve predictive accuracy and facilitate timely interventions via a multimodal machine learning approach. We evaluated data obtained from 598 participants (171 healthy controls and 427 PD patients) in three different modalities: 29 clinical features, five cerebrospinal fluid biomarkers, and 154 SNPs (single nucleotide polymorphisms), which were selected through a biology-driven feature selection method. We utilized three multimodal integration strategies—early, intermediate, and late—and trained various machine learning models, including LightGBM and Multilayer Perceptron (MLP), each of which was optimized by hyperparameter tuning and cross-validation. In early integration, we combined feature sets from all modalities into a single set, leveraging complementary information to increase predictive power. The intermediate integration method made use of autoencoders to encode features into a single 12-dimensional vector as the input into a Neural Network classifier. Late integration combined outputs from the top-performing models for each modality using ensemble techniques such as Voting Classifier and

Stacking. We found that early integration achieved the highest performance, with Support Vector Machines with 90% accuracy, 0.98 AUC-ROC, 0.99 precision, and 0.93 F1 score. Intermediate integration with the Neural Network classifier followed with an AUC-ROC of 0.84 and an F1 score of 0.85. Our feature importance analysis identified two clinical scores, namely UPSIT (related to sensory decline) and SCOPA (measuring autonomic dysfunction), along with two SNPs (chr4_90755939_A_G and chr4_90646886_G_A, both previously linked to PD susceptibility) as crucial predictors, emphasizing their well-established relevance in PD diagnosis. Early diagnosis and treatment of PD patients are facilitated by the integration of multiple data modalities, which also greatly increases the predicted accuracy for the disease while also providing us with a thorough understanding of its complexity.

Keywords: Parkinson's Disease prediction; Multimodal data; Machine Learning

1 Introduction

Parkinson's disease (PD) is a dopamine-receptor-based neurological disorder resulting from the loss of dopaminergic neurons in the substantia nigra pars compacta (SNpc) [1]. Common symptoms and indicators of Parkinson's disease are bradykinesia (slow movement), hypokinesia, rigidity, and rest tremor [1, 2]. However, each individual has their own unique experience with the disease. Patients' initial symptoms, such as tremor dominant (TD) or postural instability and gait disturbance (PIGD) [3, 4], have been used to categorize patients and predict their long-term clinical outcomes. PIGD-dominant patients exhibit more rapid illness progression and more non-motor symptoms, but their interpretation is difficult owing to long-term clinical outcomes instability. [3, 5, 6].

PD presents itself with substantial diagnostic challenges owing to its heterogeneity, and this is more prominent in its early stages. As the diagnosis of PD is typically based on the patient's medical history and physical examination [7], currently, there are no reliable tests that can differentiate between Parkinson's disease and other conditions with similar manifestations. Delay in the onset of motor symptoms that leads to PD diagnosis, as well as the associated neurodegeneration, represents a missed opportunity for early therapeutic intervention [8]. The potential for increased diagnostic accuracy has been highlighted by recent developments in multimodal evaluation techniques. The PD inpatient multidisciplinary treatment concept (Parkinson disease Multimodal Complex Treatment (PD-MCT)) [9] includes motor and physical activity as crucial treatment components and is found to have a critical positive treatment on motor function in PD patients [10]. Additionally, initiatives such as the Parkinson Progression Marker Initiative (PPMI) [11] have produced extensive datasets that integrate many modalities, allowing predictive models to become more accurate and reliable.

Clinical assessments [12] include the Movement Disorders Society Unified Parkinson Disease Rating Scale (MDS-UPDRS) and Hoehn and Yahr scales [13]. Montreal Cognitive Assessment (MoCA) may be a useful screening tool for cognitive function

in PD with numerous advantages [14, 15] and can be used to evaluate global cognition [14, 16]. Other cognitive tests include the Hopkins Verbal Learning Test-Revised (HVLT-R) to evaluate memory, the Benton Judgment of Line Orientation (BJLO) for visuospatial function, the Symbol-Digit Modalities Test (SDMT) for processing speed and attention, the Letter-Number Sequencing (LNS), and semantic (animal) fluency to evaluate working memory and executive abilities [17–21]. Neurobehavioral testing includes the Geriatric Depression Scale (GDS), State-Trait Anxiety Inventory (STAI), and Questionnaire for Impulsive-Compulsive Disorders (QUIP) [22–24]. Additional assessments like the Epworth Sleepiness Scale and REM sleep behaviour disorder (RBD) questionnaire assess sleep behaviour, Scales for Outcomes in Parkinson's Disease-Autonomic (SCOPA-AUT) is used to assess autonomic function, and the University of Pennsylvania Smell Identification Test (UPSIT) evaluates olfactory function [25–27].

There has also been a lot of work to understand the differential presence of biospecimen markers obtained from the cerebrospinal fluid in Healthy Controls (HC) and PD patients. The concentrations of α-syn, t-tau, and p-tau in the cerebrospinal fluid have been found to be lower in PD than in HC. However, there is a substantial overlap between the two groups [28]. Research suggests that PD patients may have α-syn aggregation in their central nervous systems, which could reduce CSF α-syn levels similar to how Alzheimer's disease lowers Aβ1-42. CSF tau and α-syn proteins exhibited a strong connection in both PD patients and healthy controls, suggesting complex protein interactions that may not be unique to PD [28].

Genetic factors, notably variations in PARK loci and adjacent areas, have also been increasingly identified as significant markers [29–31]. A study of SNPs in the SNCA and MAPT regions revealed that they are common risk factors for PD [32]. Interesting associations have also been found in the CAST and GAPDH genes [33, 34]. In the MIHG analysis [32], SNCA had the strongest relationship, and both SNCA and MAPT were discovered to be genome-wide important. SNPs in the CAST gene and the single rs1136666 SNP in the GAPDH gene, a high-risk marker of PD, are significantly associated with PD [33]. According to studies, several SNPs in the SNCA region, specifically rs2736990 and rs356219, consistently correlate with the risk of Parkinson's disease [34].

Current research on Parkinson's disease (PD) diagnosis relies on a single modality function, but different modalities can supplement knowledge from various angles [35, 36]. A significant amount of effort has been put into the development of machine-learning-based predictive modelling of Parkinson's disease diagnosis and severity [2, 3, 8, 36–45]. Deep learning has been used to evaluate unstructured data, such as speech and audio signals, as well as to diagnose and predict the severity of Parkinson's disease using voice data [37–41]. A study showed how incorporating multiple data modalities into modelling efforts can boost prediction efficiency thus better assessing the risk of Parkinson's disease [44]. This paper aims to combine the predictive power of individual data modalities, including clinico-demographic, cerebrospinal fluid, and genetic (SNP) data, using different integration approaches to aid the early diagnosis of PD.

2 Results

Across all three different data modalities, we had 598 participants in our analysis (171 healthy controls and 427 PD patients). The final dataset included 154 genetically relevant SNPs, five cerebrospinal fluid biomarkers, and 29 clinical features that were found using a biology-driven approach.

We explored the three integration approaches—early, intermediate, and late—to take advantage of the complementary information from clinical, biospecimen, and genetic sources from the consolidated dataset. While each strategy had specific advantages in terms of performance, early integration revealed itself to be the most effective method. The superior performance of the early integration method indicates that direct feature interactions between modalities, as opposed to those of learned representations or ensemble predictions, can significantly impact model performance.

2.1 Single-Modality Performance Analysis

To assess baseline performance and to understand the predictive potential of each modality independently, we first evaluated various machine learning models for each data modality. The results provided us with expectations for our following integration efforts and offered key insights into the predictive power of each data.

Table 1 Best Performing Models for Each Data Modality

Modality	Best Model	AUC-ROC	F1 Score	Precision	Recall
Clinical	Voting Classifier	0.939	0.922	0.936	0.907
Biospecimen	Gaussian NB	0.634	0.813	0.739	0.904
Genetic	MLP	0.533	0.759	0.774	0.746

Models consistently achieved high-performance metrics, with clinical data exhibiting the strongest independent prediction power (Table 1). Combining several base models, the Voting Classifier produced the best overall results (AUC-ROC: 0.939, F1: 0.922). Logistic Regression followed (AUC-ROC: 0.942, F1: 0.904), indicating that even linear models can successfully represent PD-related clinical trends. Notably, the MLP demonstrated its strength in detecting possible PD cases by achieving the highest recall (0.979) while maintaining reasonable precision (0.826).

Although biomarkers contain pertinent diagnostic information, models trained on just the biospecimen data suggested that they might not be enough independently to provide a reliable diagnosis of Parkinson's disease. The best recall (0.904) and balanced performance (F1: 0.813) were attained by the Gaussian Naive Bayes classifier.

Models trained solely with genetic data gave overall poor performance metrics when compared to other modalities. The MLP had the highest F1 score (0.759) along with balanced precision-recall values (precision: 0.774, recall: 0.746). These results suggest that while genetic markers can help diagnose Parkinson's disease (PD), their high dimensionality and intricate interactions make it particularly challenging to use them individually.

2.2 Early Integration Performance

The early integration strategy gave the best overall performance among the various machine learning techniques employed while taking the unified feature set across all three modalities as input. A summary of the outcomes is given in Table 2. The Voting Classifier obtained the highest AUC-ROC of 0.942 (F1-score: 0.89), while Logistic Regression showed low recall and good precision. This trade-off suggests that while some models perform well in eliminating false positives, others maintain a better balance between precision and recall, as demonstrated by the Random Forest model's high F1 score of 0.905.

Table 2 Performance Metrics for Early Integration Models

Model	AUC-ROC	AUC-PR	F1 Score	Precision	Recall
Random Forest	0.923	0.970	**0.905**	0.938	0.874
MLP	0.929	0.971	0.894	0.973	0.828
Voting Classifier	**0.942**	0.978	0.888	0.973	0.816
SVC	0.872	0.959	0.881	0.972	0.805
LightGBM	0.707	0.843	0.840	0.782	**0.908**
Gaussian NB	0.868	0.949	0.838	0.918	0.770
Logistic Regression	0.940	**0.979**	0.783	**1.000**	0.644

Bold values indicate the best performance for each metric.

Key hyperparameters for the best-performing models are detailed in Table 3. Details about model parameters are also mentioned in Supplementary Table 4.

Table 3 Optimal Hyperparameters for Top Early Integration Models

Model	Key Hyperparameters
Random Forest	class_weight={0: 0.986, 1: 0.013}, max_depth=19, n_estimators=250
MLP	hidden_layer_sizes=(87, 87, 87), activation='tanh', learning_rate='invscaling'
Voting Classifier	voting='soft', combination of RF, GNB, SVC, LR, LightGBM, and MLP

2.3 Intermediate Integration Analysis

The performance of the autoencoder-based intermediate integration strategy was moderate but consistent. The autoencoder architecture's dimensionality reduction (Table 4) shows that we were able to capture essential patterns (Table 5) while drastically decreasing the feature space from 188 to 12 dimensions. Interestingly, the base autoencoder configuration without SMOTE (synthetic samples for minority class to account for imbalance) maintained competitive performance across other metrics and achieved the highest recall (0.872), indicating that simpler architectures might be more resilient for this specific integration task.

Table 4 Autoencoder Architecture Details

Modality	Input Dimensions	Compressed Dimensions
Clinical	29	4
Biospecimen	5	4
Genetic	154	4

Table 5 Performance Metrics for Intermediate Integration Models

Model Configuration	AUC-ROC	AUC-PR	F1 Score	Precision	Recall
Base Autoencoder	0.837	**0.934**	**0.852**	0.833	**0.872**
With SMOTE	**0.843**	0.933	0.828	**0.915**	0.756

2.4 Late Integration Results

Table 6 shows that the late integration ensemble approaches performed well across the various metrics, complementing our earlier integration approaches. The best performance in the present case was the XGBoost Classifier, which achieved balanced values of precision (0.856) and recall (0.895) by ranking the highest on the majority of criteria. This suggests that model-level integration can effectively capture complex patterns while still maintaining generalisability.

Table 6 Performance Metrics for Late Integration Models

Model	AUC-ROC	AUC-PR	F1 Score	Precision	Recall
XGBoost	**0.892**	**0.961**	**0.875**	0.856	**0.895**
Random Forest	0.891	0.959	0.874	0.864	0.884
Stacking Classifier	0.872	0.955	0.862	0.852	0.872
Voting Classifier	0.866	0.953	0.843	**0.918**	0.779

2.5 Feature Importance Analysis

As observed in Table 7, feature importance analysis using XGBoost across modalities identified several important determinants for diagnosing Parkinson's disease. The two most important clinical features, UPSIT and SCOPA-Total, indicate the critical roles of olfactory function and autonomic dysfunction in the diagnosis of Parkinson's disease. In the Biospecimen data, we observe α-synuclein and the P_TAU/T_TAU ratio were the most significant predictors, as well as T_TAU, URATE, and Aβ42 all contributing significantly. Among the many SNPs, ones on chromosome 4 were found to be the top predictors, highlighting their potential role in the early detection of Parkinson's disease.

Table 7 Key Features Identified Through XGBoost Analysis Across Modalities

Data Modality	Important Features
Clinical Features	UPSIT, First_Fam_PD, SCOPA_GASTRO, SCOPA_TOTAL, Other_Fam_PD
Biospecimen Features	P_TAU/T_TAU, α-SYN, T_TAU, URATE, Aβ42
Genetic Markers	chr4_90645671_T_A, chr4_90697157_T_C, chr4_90635338_G_C, chr4_90668614_T_C, chr4_90754292_T_C

3 Methods

3.1 Study Data

Data across all three different modalities, namely, Clinico-demographic, Biospecimen, and Genetic, from the Parkinson Progression Marker Initiative (PPMI) [11] was used for this study. The detailed workflow used can be found in Figure 2.

For the clinical modality, a total of 682 participants were analyzed, containing 196 control cases and 486 PD patients. The mean age for the participants was 60.9 years, with a standard deviation of 10.2. The following features were considered for clinico-demographic analysis: features encompassing family history parameters (First_Fam_Num, First_Fam_PD, Other_Fam_Num, Other_Fam_PD, AGE, Socio Score), cognitive assessments (BJLO, HVLT measures, LNS), behavioural indices (ESS, GDS, QUIP sections), and autonomic function measures (SCOPA components). Common diagnostic indicators like MDS-UPDRS scores, Tremor scores, PIGD scores, and MoCA scores were not included in the analysis to prevent diagnostic circularity. Supplementary Table 1 provides a thorough explanation of the different parameters and how they are calculated.

Similarly, for the biospecimen modality, 641 participants were analyzed, with 185 HC and 456 PD cases. The features considered were Aβ42 (958.9 ± 384.5), α-synuclein (1614.2 ± 623.4), p-tau (15.7 ± 6.6), t-tau (182.2 ± 65.0), and urate (312.3 ± 73.5) obtained from the cerebrospinal fluid of the participant. The detailed methodology used to extract these features can be found in Supplementary Table 2.

Finally, for the genetic biomarkers modality, we considered 733 participants, with 217 belonging to HC and 516 to PD. Variant Calling Format (VCF) files containing SNPs (hg19) of each participant were analysed to give a total of 6899 unique SNPs in the PPMI data. A biology-based approach was applied to shortlist prognostic SNPs. To do this, a detailed survey of SNPs majorly lying in PD-related gene regions was done, and 88 prognostic SNPs were identified and annotated to hg19 reference. 51 of these 88 biologically relevant SNPs overlapped with our genetic dataset. Using these 51 SNPs, a correlation analysis was performed with the 6899 and top 20 of the most correlated SNPs (Chi-Squared test, p-value < 0.05) for each prognostic SNP, were selected. This resulted in a set of 335 biologically relevant SNPs, following which the XGBoost feature selection method was used to rank the SNPs by feature importance and a final set of 154 prognostic SNPs were obtained (Figure 1). These 154 SNPs, mentioned in Supplementary Table 3, were used for further genetic analysis.

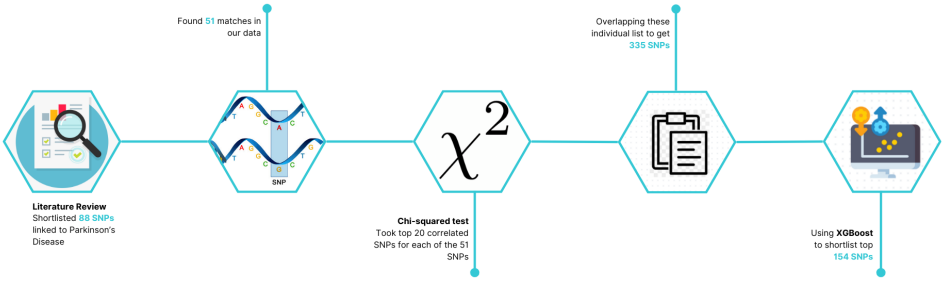

Fig. 1 Workflow for Genetic Data Feature Selection.

3.2 Data Preprocessing and Integration

The first step was to go over multiple time points for which participant data was collected in each dataset and take an average over them. For each modality, only those participants who did not have missing data for any of the features were considered. Additionally, we perform a dimensionality analysis using Principal Component Analysis (PCA) using Eigenvalue Decomposition and Singular Value Decomposition (SVD). Based on the correlation between features, PCA assisted in identifying trends in data. Finally, for the multimodal dataset, participants with information across all three individual modalities were considered (n = 598).

3.3 Machine Learning Framework

We built predictive models using supervised machine learning for PD diagnosis using each of the three data modalities individually to get baseline models and then applied the different integration approaches for multimodal data.

To leverage the complementary information provided by various modalities, we implemented three different integration approaches. We created a single dataset with 598 participants and 188 features (29 clinical, five biospecimen, and 154 genetic markers) by concatenating features from all modalities in the early integration phase. This combined dataset was used to train six machine learning models: Random Forest, Gaussian Naive Bayes, Support Vector Machines (SVM), Logistic Regression, LightGBM, and Multilayer Perceptron (MLP), all implemented using Python Scikit-learn.

For intermediate integration, we developed separate autoencoder networks for each modality to learn compressed representations. Biospecimen autoencoders compressed

five dimensions to four, clinical autoencoders lowered 29 dimensions to four, and genetic autoencoders brought down 154 dimensions to four. The Adam optimizer with learning rate decay was used to optimize each autoencoder, which used LeakyReLU activation functions. A neural network classifier with batch normalization and dropout layers was fed the concatenated compressed features. Binary cross-entropy loss and sigmoid activation were used for the final prediction. To address the class imbalance, we also applied SMOTE within our five-fold cross-validation framework to generate synthetic minority class samples, followed by StandardScaler normalization.

The late integration strategy involved training separate models for each modality and combining their predictions through ensemble methods. Based on validation performance, we first determined which model performed best for each modality (single-modality). Four ensemble methods—Random Forest, Voting Classifier, Stacking Classifier, and XGBoost—were then used to integrate these predictions. This method enabled distinct features of the data to be captured by each modality-specific model before we integrated them.

3.4 Model Training and Evaluation

To preserve class distribution for the standard machine learning models, we used an 80:20 train-test split with stratification. Five-fold cross-validation and randomized search were used to optimize across the hyperparameter grid. To prevent overfitting, we included early stopping based on validation loss for neural network components in the intermediate integration strategy, along with a 60:20:20 train-validation-test split. Since accuracy alone can be deceptive in class-imbalanced datasets like ours, the F1 score was the primary metric used to assess model performance.

The Keras framework was used to train the neural network classifier, which used ReLU and sigmoid activation functions. Hyperparameters were tuned using the Adam optimizer and binary cross-entropy loss. We used dropout and batch normalization for regularisation, setting the dropout rate at 0.3.

Our public GitHub repository (https://github.com/SP9144/MutliModal_Prediction_NLD) contains all implementation code, including trained models and their weights.

4 Discussion

In this study, we demonstrate that the diagnostic accuracy of Parkinson's disease can be considerably improved by combining various data modalities using machine learning techniques. We observe that our early integration strategy performs more accurately (AUC-ROC: 0.92, F1-score: 0.90) than intermediate (AUC-ROC: 0.83, F1-score: 0.85) and late integration (AUC-ROC: 0.89, F1-score: 0.86) approaches. Thus, we emphasize the idea that maintaining original feature associations across the data modalities allows us to capture complex disease patterns much better. One of the key findings was the complementary effect brought in by combining the different modalities. Genetic data on its own had a low performance of AUC-ROC = 0.65 (F1-score: 0.76). However, upon combining with clinical and biospecimen data, it contributed

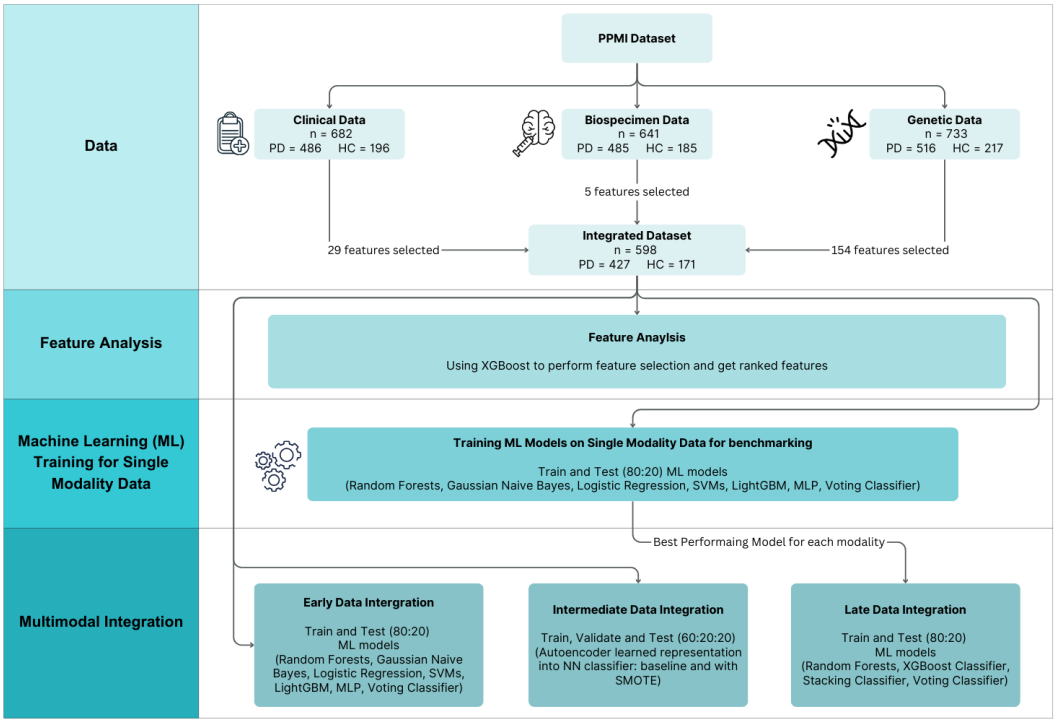

Fig. 2 Overview of the multimodal workflow used in the study.

two SNPs (chr4_90727088_C_T and chr4_90637010_A_G) among the top predictive features in the integrated model. This suggests that there may be linked effects that only show up when other biological markers are present [44].

Our feature importance analysis confirmed the diagnostic relevance of established clinical markers while revealing novel insights. UPSIT scores emerged as a top predictor, aligning with evidence of olfactory dysfunction as an early PD indicator [46]. Similarly, the importance of autonomic dysfunction assessment in the diagnosis of PD was confirmed by the presence of SCOPA-Total and SCOPA-Gastro scores as one of the top predictive characteristics. [23]. These findings support the clinical utility of these measures while suggesting their enhanced value when considered alongside genetic and biospecimen data.

Compared to previous studies, our approach demonstrates several advantages. While earlier work using speech signals achieved accuracies of 85-95% [2], and studies focusing on non-motor symptoms reported accuracies of 72-92% [47], our multimodal approach achieved 90% accuracy without relying on traditional clinical markers like MDS-UPDRS scores, thus avoiding diagnostic circularity. The effectiveness of the integration approaches employed was also highlighted when compared to the recent

multimodal models using PPMI data, which reported lower performance (accuracy: 75%, AUC: 0.85) [11].

Our innovative, biology-driven feature selection approach for genetic data reduces the dimensionality and solidifies our understanding of the disease by ensuring that biological relevance is maintained. Combined with our thorough analysis of integration techniques, it allowed us to provide valuable insights for future multimodal studies in PD and other complex diseases. The accuracy of our approach can be further improved by augmenting it with other modalities like speech data or neuroimaging of patients, among other attributes. Our multimodal approach's high-performance points to its potential clinical relevance for early PD diagnosis, especially when single-modality examinations yield conflicting results. Identifying important predictive characteristics across modalities also offers possible targets for future research and targeted clinical assessments.

5 Conclusion

In this paper, we propose a multimodal machine-learning model that can help diagnose Parkinson's disease. Our proposed model produces more accurate results (AUC-ROC: 0.92, F1-score: 0.91) as compared to existing methods. It was observed that the model trained on multimodal data shows better performance than those trained on single data modalities and highlights unexpected features that could be prognostic biomarkers of Parkinson's Disease. The multimodality also adds to the robustness as well as generalizability of the model with the various data sources compensating and supplementing each other, giving us a more comprehensive and holistic view. This paper tries to provide insight into the broader application of machine learning models on multimodal data sources to improve disease diagnosis in the domain of healthcare. This is especially crucial for diseases like cancer because of their complicated and heterogeneous patient manifestations. An integrated heterogeneous data source could further leverage machine learning techniques' ability to capture intricate correlations between characteristics in order to create predictive models that are even more accurate.

Supplementary Information. The following supplementary files are provided with this study:

Supplementary File 1: Detailed description of the various clinical parameters and their computations.

Supplementary File 2: Detailed methodology used to extract biospecimen features.

Supplementary File 3: List of the final 154 prognostic SNPs shortlisted using genetic data.

Supplementary File 4: Details of each model and corresponding hyperparameters.

References

[1] Lang, A.E., Lozano, A.M.: Parkinson's disease. New England Journal of Medicine **339**(16), 1130–1143 (1998)

[2] Bind, S., Tiwari, A.K., Sahani, A.K.: A survey of machine learning based approaches for Parkinson disease prediction (2015)

[3] Aleksovski, D., Miljkovic, D., Bravi, D., Antonini, A.: Disease progression in parkinson subtypes: the ppmi dataset. Neurological Sciences **39** (2018)

[4] Alves, G., Larsen, J.P., Emre, M., Wentzel-Larsen, T., Aarsland, D.: Changes in motor subtype and risk for incident dementia in parkinson's disease. Movement Disorders **21**(8), 1123–1130 (2006)

[5] Biundo, R., Weis, L., Antonini, A.: Cognitive decline in parkinson's disease: The complex picture. npj Parkinson's Disease **2**(1) (2016)

[6] Kelly, V., Johnson, C., McGough, E., Shumway-Cook, A., Horak, F., Chung, K., Espay, A., Revilla, F., Devoto, J., Wood-Siverio, C., *et al.*: Association of cognitive domains with postural instability/gait disturbance in parkinson's disease. Parkinsonism & Related Disorders **21**(7), 692–697 (2015)

[7] Michell, A., Lewis, S., Foltynie, T., Barker, R.: Biomarkers and parkinson's disease. Brain **127**(8), 1693–1705 (2004)

[8] Bowman, F.D., Drake, D.F., Huddleston, D.E.: Multimodal imaging signatures of parkinson's disease. Frontiers in Neuroscience **10** (2016)

[9] Knop, M., Samann, P.G., Keck, M.E.: Parkinson-komplextherapie am max-planck-institut für psychiatrie: Ein multimodales, flexibles stationäres therapieprogramm für kritische krankheitsphasen. Psychiatr. Neurol. **5**, 24–29 (2017)

[10] Hartelt, E., Scherbaum, R., Kinkel, M., Gold, R., Muhlack, S., Tonges, L.: Parkinson's disease multimodal complex treatment (pd-mct): Analysis of therapeutic effects and predictors for improvement. Journal of Clinical Medicine **9**(6), 1874 (2020)

[11] Marek, K., Jennings, D., Lasch, S., Siderowf, A., Tanner, C., Simuni, T., Coffey, C., Kieburtz, K., Flagg, E., Chowdhury, S., *et al.*: The parkinson progression marker initiative (ppmi). Progress in Neurobiology **95**(4), 629–635 (2011)

[12] Marek, K., Chowdhury, S., Siderowf, A., Lasch, S., Coffey, C.S., Caspell-Garcia, C., Simuni, T., Jennings, D., Tanner, C.M., Trojanowski, J.Q., *et al.*: The parkinson's progression markers initiative (ppmi) establishing a pd biomarker cohort. Annals of Clinical and Translational Neurology **5**(12), 1460–1477 (2018)

[13] Goetz, C.G., Tilley, B.C., Shaftman, S.R., Stebbins, G.T., Fahn, S., Martinez-Martin, P., Poewe, W., Sampaio, C., Stern, M.B., Dodel, R., *et al.*: Movement disorder society-sponsored revision of the unified parkinson's disease rating scale (mds-updrs): scale presentation and clinimetric testing results. Movement Disorders **23**(15), 2129–2170 (2008)

[14] Dalrymple-Alford, J., MacAskill, M., Nakas, C., Livingston, L., Graham, C., Crucian, G., Melzer, T., Kirwan, J., Keenan, R., Wells, S., *et al.*: The moca: well-suited screen for cognitive impairment in parkinson disease. Neurology **75**(19), 1717–1725 (2010)

[15] Skorvanek, M., Goldman, J.G., Jahanshahi, M., Marras, C., Rektorova, I., Schmand, B., Duijn, E., Goetz, C.G., Weintraub, D., Stebbins, G.T., *et al.*: Global scales for cognitive screening in parkinson's disease: Critique and recommendations. Movement Disorders **33**(2), 208–218 (2018)

[16] Nasreddine, Z.S., Phillips, N.A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., Chertkow, H.: The montreal cognitive assessment, moca: a brief screening tool for mild cognitive impairment. Journal of the American Geriatrics Society **53**(4), 695–699 (2005)

[17] Benedict, R.H., Schretlen, D., Groninger, L., Brandt, J.: Hopkins verbal learning test revised: Normative data and analysis of inter-form and test-retest reliability. The Clinical Neuropsychologist **12**(1), 43–55 (1998)

[18] Benton, A.L., Varney, N.R., Hamsher, K.D.: Visuospatial judgment: A clinical test. Archives of Neurology **35**(6), 364–367 (1978)

[19] Smith, A.: Symbol Digit Modalities Test. Western Psychological Services Los Angeles, ??? (1973)

[20] Wechsler, D.: Wechsler Adult Intelligence Scale, (1955)

[21] Gladsjo, J.A., Schuman, C.C., Evans, J.D., Peavy, G.M., Miller, S.W., Heaton, R.K.: Norms for letter and category fluency: demographic corrections for age, education, and ethnicity. Assessment **6**(2), 147–178 (1999)

[22] Weintraub, D., Oehlberg, K.A., Katz, I.R., Stern, M.B.: Test characteristics of the 15-item geriatric depression scale and hamilton depression rating scale in parkinson disease. The American journal of geriatric psychiatry **14**(2), 169–175 (2006)

[23] Knight, R.G., Waal-Manning, H.J., Spears, G.F.: Some norms and reliability data for the state-trait anxiety inventory and the zung self-rating depression scale. British Journal of Clinical Psychology **22**(4), 245–249 (1983)

[24] Weintraub, D., Hoops, S., Shea, J.A., Lyons, K.E., Pahwa, R., Driver-Dunckley,

E.D., Adler, C.H., Potenza, M.N., Miyasaki, J., Siderowf, A.D., *et al.*: Validation of the questionnaire for impulsive-compulsive disorders in parkinson's disease. Movement disorders: official journal of the Movement Disorder Society **24**(10), 1461–1467 (2009)

[25] Ondo, W.G., Vuong, K.D., Khan, H., Atassi, F., Kwak, C., Jankovic, J.: Daytime sleepiness and other sleep disorders in parkinson's disease. Neurology **57**(8), 1392–1396 (2001)

[26] Stiasny-Kolster, K., Mayer, G., Schafer, S., Moller, J.C., Heinzel-Gutenbrunner, M., Oertel, W.H.: The rem sleep behavior disorder screening questionnaire a new diagnostic instrument. Movement disorders **22**(16), 2386–2393 (2007)

[27] Visser, M., Marinus, J., Stiggelbout, A.M., Van Hilten, J.J.: Assessment of autonomic dysfunction in parkinson's disease: the scopa-aut. Movement disorders: official journal of the Movement Disorder Society **19**(11), 1306–1312 (2004)

[28] Kang, J.-H., Mollenhauer, B., Coffey, C.S., Toledo, J.B., Weintraub, D., Galasko, D.R., Irwin, D.J., Van Deerlin, V., Chen-Plotkin, A.S., Caspell-Garcia, C., *et al.*: Csf biomarkers associated with disease heterogeneity in early parkinson's disease: the parkinson's progression markers initiative study. Acta Neuropathologica **131**(6), 935–949 (2016)

[29] Chung, S.J.: Human genetic variation and parkinson's disease. Journal of Movement Disorders **3**(1), 1–5 (2010)

[30] Chung, S.J., Armasu, S.M., Biernacka, J.M., Lesnick, T.G., Rider, D.N., Lincoln, S.J., Ortolaza, A.I., Farrer, M.J., Cunningham, J.M., Rocca, W.A., Maraganore, D.M.: Common variants in park loci and related genes and parkinson's disease. Movement Disorders **26**(2), 280–288 (2010)

[31] Klein, C., Westenberger, A.: Genetics of parkinson's disease. Cold Spring Harbor Perspectives in Medicine **2**(1), 008888–008888 (2012)

[32] Edwards, T.L., Scott, W.K., Almonte, C., Burt, A., Powell, E.H., Beecham, G.W., Wang, L., Zuchner, S., Konidari, I., Wang, G., *et al.*: Genome-wide association study confirms snps in snca and the mapt region as common risk factors for parkinson disease. Annals of Human Genetics **74**(2), 97–109 (2010)

[33] Allen, A.S., Satten, G.A.: Snps in cast are associated with parkinson disease: A confirmation study. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics **9999B**, (2010)

[34] Campelo, C.L.d.C., Silva, R.H.: Genetic variants in snca and the risk of sporadic parkinson's disease and clinical outcomes: A review. Parkinson's Disease **2017**, 111 (2017)

[35] Lei, B., Chen, S., Ni, D., Wang, T.: Discriminative learning for alzheimer's disease diagnosis via canonical correlation analysis and multimodal fusion. Frontiers in Aging Neuroscience **8**, 77 (2016)

[36] Huang, Z., Lei, H., Zhao, Y., Zhou, F., Yan, J., Elazab, A., Lei, B.: Longitudinal and multi-modal data learning for parkinson's disease diagnosis. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 1411–1414 (2018). IEEE

[37] Grover, S., Bhartia, S., Yadav, A., Seeja, K.: Predicting severity of parkinson's disease using deep learning. Procedia Computer Science **132**, 1788–1794 (2018)

[38] Genain, N., Huberth, M., Vidyashankar, R.: Predicting Parkinson's disease severity from patient voice features (2014)

[39] Galaz, Z., Mzourek, Z., Mekyska, J., Smekal, Z., Kiska, T., Rektorova, I., Orozco-Arroyave, J.R., Daoudi, K.: Degree of parkinson's disease severity estimation based on speech signal processing. In: 2016 39th International Conference on Telecommunications and Signal Processing (TSP), pp. 503–506 (2016). IEEE

[40] Asgari, M., Shafran, I.: Extracting cues from speech for predicting severity of parkinson's disease. In: 2010 IEEE International Workshop on Machine Learning for Signal Processing, pp. 462–467 (2010). IEEE

[41] Tsanas, A., Little, M., McSharry, P., Ramig, L.: Accurate telemonitoring of parkinson's disease progression by non-invasive speech tests. Nature Precedings, 1–1 (2009)

[42] Nalls, M.A., McLean, C.Y., Rick, J., Eberly, S., Hutten, S.J., Gwinn, K., Sutherland, M., Martinez, M., Heutink, P., Williams, N.M., *et al.*: Diagnosis of parkinson's disease on the basis of clinical and genetic classification: a population-based modelling study. The Lancet Neurology **14**(10), 1002–1009 (2015)

[43] Faghri, F., Hashemi, S.H., Leonard, H., Scholz, S.W., Campbell, R.H., Nalls, M.A., Singleton, A.B.: Predicting onset, progression, and clinical subtypes of parkinson disease using machine learning. bioRxiv, 338913 (2018)

[44] Makarious, M.B., Leonard, H.L., Vitale, D., Iwaki, H., Sargent, L., Dadu, A., Violich, I., Hutchins, E., Safo, D., Bandres-Ciga, S., et al.: Multi-modality machine learning predicting parkinson's disease. Scientific Reports (2021)

[45] Li, W., Zhu, W., Dorsey, E.R., Luo, J.: Predicting parkinson's disease with multimodal irregularly collected longitudinal smartphone data. In: 2020 IEEE International Conference on Data Mining (ICDM), pp. 1106–1111 (2020). IEEE

[46] Wills, J., Jones, J., Haggerty, T., Duka, V., Joyce, J.N., Sidhu, A.: Elevated

tauopathy and alpha-synuclein pathology in postmortem parkinson's disease brains with and without dementia. Experimental neurology **225**(1), 210–218 (2010)

[47] Armananzas, R., Bielza, C., Chaudhuri, K.R., Martinez-Martin, P., Larranaga, P.: Unveiling relevant non-motor parkinson's disease severity symptoms using a machine learning approach. Artificial Intelligence in Medicine **58**(3), 195–202 (2013)