Beyond Diagnosis: Enhancing Parkinson's Disease Classification and Symptom Profiling Using Wearable Data

Authors

Charlotte Wong | University of Toronto

Kiarash Kianidehkordi | University of Toronto

Motivation

Parkinson's Disease (PD) is a progressive neurodegenerative disorder affecting approximately 1 in 500 Canadians, primarily impairing motor control through symptoms like tremor, rigidity, and bradykinesia. As the disease advances, patients suffer reduced mobility, independence, and quality of life. Traditional diagnostic methods are limited in their ability to monitor symptom variability in real time. There is a growing need for innovative, data-driven tools to enhance early detection and symptom monitoring for development of smart drug delivery systems. Wearable technology, combined with machine learning, offers a promising solution for continuous, objective assessment of PD symptoms.

Objectives

- Replicate a supervised machine learning pipeline (BOSS + SVM) for classifying Parkinson's disease implemented by the original paper.
- Extend prior research by applying unsupervised learning to uncover tremor-related symptom patterns from the accelerometer data.

Data

Parkinson's Apple Watch Dataset (PADS)

- Real-world data from 469 participants wearing wrist-mounted Apple Watches on both arms.
- Data include 3-axis accelerometer signals and Questionnaire on symptom severity and patient metadata.

Supervised classification of PD

The Bag-of-Symbolic Fourier Approximation Symbols (BOSS) algorithm was used to convert each 72-channel wrist accelerometer time series with 1789 time steps per sample into a symbolic representation. A Support Vector Machine (SVM) classifier was then trained to classify Parkinson's Disease versus Healthy Controls (PD vs HC) and Parkinson's Disease versus Differential Diagnoses (PD vs DD). Model evaluation was conducted using nested 5-fold cross-validation.

Unsupervised Symptoms profiling

- Each subject's 3-axis acceleration data was first transformed into a single-dimensional signal by computing the Euclidean magnitude across axes.
- The resulting signal was bandpass filtered between 3–12 Hz using a 6th-order Butterworth filter to isolate tremor-related activity.
- Feature extraction was performed on the filtered signal to obtain mean acceleration, mean envelope amplitude (based on local maxima and minima), peak power in the tremor band, and the area under the power spectral density (PSD) curve.
- These four features were computed for each wrist and concatenated into an 8-dimensional feature vector per subject. The full feature matrix was standardized using Z-score normalization.
- Clustering was performed using a Gaussian Mixture Model (GMM), with the optimal number of clusters selected via Bayesian Information Criterion (BIC).
- Cluster quality was evaluated using silhouette score and the Calinski-Harabasz index.
- Dimensionality reduction using t-distributed Stochastic Neighbor Embedding (t-SNE) was applied to visualize the clustering results in two dimensions. t-SNE plots were colored by GMM cluster assignments, patient diagnoses, and the absolute difference between patient age and age at diagnosis (used as a proxy for disease duration).

Results

Our replicated BOSS + SVM pipeline on the PADs dataset achieved results and values that align with the previously published results .

- In the PD vs. HC task, our model achieved a balanced accuracy of 78.0% ± 5.0%, with an F1 score of 0.75 ± 0.05, precision of 0.75 ± 0.07, and recall of 0.77 ± 0.07.
- For the PD vs. DD task, the balanced accuracy was 76.1% ± 3.0%, F1 score was 0.63 ± 0.05, precision was 0.67 ± 0.14, and recall was 0.63 ± 0.08.

The clustering solution demonstrated strong internal validity, with a silhouette score of 0.688 and a Calinski-Harabasz index of 1718.96, indicating well-separated and compact clusters.

Limitations

- The Gaussian Mixture Model (GMM) assumes that the underlying data distribution is a mixture of multivariate Gaussian components. This assumption may not hold for real-world tremor data
- Although t-SNE is effective for visualizing high-dimensional data, it is a non-linear technique that is sensitive to initialization and hyperparameter choices such as perplexity.
- Clustering is inherently constrained by the input features derived from wrist-worn devices.
- Use of an unsupervised learning framework means that cluster labels are not guided by predefined diagnostic categories. While this allows for unbiased discovery of data-driven groupings, it also makes the interpretation of clusters more challenging.