International Journal of Quantitative Structure-Property Relationships
Volume 7 • Issue 2

Classification of Parkinson's Disease Using Motor and Non-Motor Biomarkers Through Machine Learning Techniques

Anirudh Reddy Cingireddy, Jackson State University, USA
Robin Ghosh, Jackson State University, USA
Venkata Kiran Melapu, Jackson State University, USA
Sravanthi Joginipelli, Jackson State University, USA
Tor A. Kwembe, Jackson State University, USA

ABSTRACT

Parkinson's disease (PD) is the second most neurodegenerative disease in the United States of America after Alzheimer's disease. The Parkinson's disease patients and scans without evidence for dopaminergic deficit (SWEDD) patients will share the same symptoms, and it's hard to differentiate the PD, SWEDD patients, and healthy controls in the progression of PD. In this research, the authors classified PD patients, SWEDD patients, and healthy controls by considering motor and non-motor biomarkers, namely MDS-UPDRS part 1, SCOPA score, and QUIP-RS, from the PPMI database by using supervised and unsupervised machine learning algorithms, namely KNN, logistic regression, XGBooting, naive Bayes, decision tree, random forest, support vector machine, multilayer perceptron, and K-means clustering, respectively. Random forest scored 98% accuracy among all these algorithms and can identify and differentiate PD, SWEDD, and healthy control patients by motor and non-motor biomarkers.

KEYWORDS

Alzheimer's Disease, Decision Tree, K-Means Clustering, KNN, Logistic Regression, Multilayer Perceptron, Naïve Bayes, Parkinson's Disease, Random Forest, Support Vector Machine, SWEDD

INTRODUCTION

Among neurological diseases, Parkinson's disease (PD) is the second most common disease after Alzheimer's disease, and it affects around 1 million Americans (Marras et al., 2018). PD was first medically described as a neurological syndrome by James Parkinson in 1817 as a "shaking palsy." It is a chronic, progressive neurodegenerative disease characterized clinically, *i.e.* both motor and non-motor features characterized by Parkinsonism and widespread Lewy body pathology in the central nervous system (CNS), peripheral nervous system (PNS), and autonomic nervous system (ANS). Also, the loss of dopamine-releasing neurons of substantia nigra is the primary underlying pathology of PD. The risk factors of PD are genetic, environmental factors (Dick et al., 2007), age, and gender. At the age of 60, people who fall under risk factors are more prone to PD (DeMaagd & Philip, 2015). Hallmark motor symptoms of this disease include masklike facial expression, bradykinesia, resting tremor, rigidity, and festinating gait. Tremors are involuntary and oscillatory twitching of limbs.

DOI: 10.4018/IJQSPR.290011

Rigidity is the stiffness in movements caused by increased muscle tone, which can be painful. The slowness of movement (bradykinesia) is another characteristic feature of PD. Postural instability leads to impaired balance and sudden falls.

Approximately 10% of patients clinically diagnosed with early PD have normal dopaminergic functional imaging called scans without evidence for dopaminergic deficit (SWEDD). SWEDD alludes to the non-appearance, as opposed to the presence of an imaging anomaly in patients clinically diagnosed to have PD. The term SWEDD has since been generally utilized in clinical writing, even as a demonstrative mark. Both SWEDD and PD share the same motor and non-motor symptoms, and it is challenging to identify the patients with PD or SWEDD. So, it is essential to classify the patients with PD and SWEDD, to avoid unnecessary medical examinations and therapies and their associated financial costs, side effects, and safety risks.

The PD biomarkers are classified into biochemical, genetic, imaging, and clinical (Delenclos, Jones, McLean, & Uitti, 2016) (Emamzadeh & Surguchov, 2018). A biomarker is defined as a marker of a specific illness state or a specific condition of an organic entity. A boundary can be utilized to assess the advancement of infection or the impacts of treatment. The biochemical biomarkers show which chemicals like pesticides and herbicides are affecting the PD. The genetic biomarkers show which genes are causing PD and linked to the progression of PD, some of the names of the genes are PARK1, PARK2, PARK3, etc. Imaging biomarkers are the image techniques that are used to identify the PD patients. The well-known techniques are transcranial B-mode sonography (TCS), susceptibility-weighted imaging (SWI), diffusion-weighted imaging (DWI), positron emission tomography (PET) scan and, single-photon emission computed tomography (SPECT) scan are frequently used to identify the progression of PD patients. The clinical biomarkers are like motor and non-motor symptoms or risk factors observed in thousands of patients at the time of doctor's visits. Also, several research studies have proved that these motor and non-motor symptoms help identify and relate to PD patient's progression (Sharma et al., 2013).

These are some known documents-based questioner tests to evaluate the progression of PD and SWEDD based on the motor and non-motor symptoms (Noyce et al., 2012), namely Movement Disorder Society-sponsored Unified Parkinson's Disease Rating Scale (MDS-UPDRS)(Holden, Finseth, Sillau, & Berman, 2018). Scales for Outcomes in Parkinson's Disease-Autonomic questionnaire (SCOPA-AUT)(Visser, Marinus, Stiggelbout, & Hilten, 2004)(Bostantjopoulou et al., 2016). Questionnaire for Impulsive-Compulsive Disorders in Parkinson's Disease–Rating Scale (QUIP-RS) (Martínez-Martín et al., 2015). The MDS-UPDRS was created to assess different features of Parkinson's disease, including motor and non-motor encounters of everyday living and motor problems. It incorporates a motor assessment and portrays the degree and weight of infection across different populaces. The scale can be utilized in a clinical setting, just as in research (www.movementdisorders.org). It consists of 4 parts. Part I relates to ''non-motor experiences of daily living,'' Part II be about ''motor experiences of daily living,'' Part III is continuing to have the ''motor examination,'' and Part IV relates to ''motor complications.''. In this research, we considered only MDS-UPDRS part 1. SCOPA is used to assess the autonomic symptoms in PD. The scale is self-finished by patients and comprises 25 aspects evaluating the accompanying spaces: urinary (6), gastrointestinal (7), pupillomotor (1), cardiovascular (3), thermoregulatory (4), and sexual (2 items for men and 2 items for women). The QUIP-RS is a concise, self-finished, or rater-managed rating scale to survey the seriousness of manifestations of Impulse control disorders and related actions reported to happen in PD.

These tests are time-consuming and are challenging to perform. We need a machine learning model that can predict and classify in less time with high accuracy. For this reason, we collected the motor and non-motor symptoms data from Parkinson's Progression Markers Initiative (PPMI) and developed machine learning algorithms to classify the patients between PD, SWEDD, and Healthy Controls. The machine learning (ML) methods are a form of artificial intelligence (AI). Artificial intelligence algorithms allow computers to learn, train, and use various cases to predict the results. Machine learning algorithms can be used from the earliest stage and can be used to examine clinical datasets.