Available online at www.sciencedirect.com
ScienceDirect
journal homepage: www.elsevier.com/locate/bbe

Original Research Article

A hybrid intelligent system for the prediction of Parkinson's Disease progression using machine learning techniques

Q1 *Mehrbakhsh Nilashi* *, *Othman Ibrahim, Hossein Ahmadi, Leila Shahmoradi* *, *Mohammadreza Farahmand*

Faculty of Computing, Universiti Technologi Malaysia, Johor, Malaysia

ARTICLE INFO

Article history:
Received 7 June 2017
Received in revised form
5 August 2017
Accepted 18 September 2017
Available online xxx

Keywords:
Healthcare
Parkinson Disease diagnosis
UPDRS
Clustering
Dimensionality reduction
ISVR

ABSTRACT

Parkinson's Disease (PD) is a progressive degenerative disease of the nervous system that affects movement control. Unified Parkinson's Disease Rating Scale (UPDRS) is the baseline assessment for PD. UPDRS is the most widely used standardized scale to assess parkinsonism. Discovering the relationship between speech signal properties and UPDRS scores is an important task in PD diagnosis. Supervised machine learning techniques have been extensively used in predicting PD through a set of datasets. However, the most methods developed by supervised methods do not support the incremental updates of data. In addition, the standard supervised techniques cannot be used in an incremental situation for disease prediction and therefore they require to recompute all the training data to build the prediction models. In this paper, we take the advantages of an incremental machine learning technique, Incremental support vector machine, to develop a new method for UPDRS prediction. We use Incremental support vector machine to predict Total-UPDRS and Motor-UPDRS. We also use Non-linear iterative partial least squares for data dimensionality reduction and self-organizing map for clustering task. To evaluate the method, we conduct several experiments with a PD dataset and present the results in comparison with the methods developed in the previous research. The prediction accuracies of method measured by MAE for the Total-UPDRS and Motor-UPDRS were obtained respectively MAE = 0.4656 and MAE = 0.4967. The results of experimental analysis demonstrated that the proposed method is effective in predicting UPDRS. The method has potential to be implemented as an intelligent system for PD prediction in healthcare.

Q2 * *Corresponding authors at*: Faculty of Computing, Universiti Technologi Malaysia, Johor, Malaysia.
E-mail addresses: nilashidotnet@hotmail.com (M. Nilashi), leilashahmoradi@hotmail.com (L. Shahmoradi).

http://dx.doi.org/10.1016/j.bbe.2017.09.002

Please cite this article in press as: Nilashi M, et al. A hybrid intelligent system for the prediction of Parkinson's Disease progression using machine learning techniques. Biocybern Biomed Eng (2017), http://dx.doi.org/10.1016/j.bbe.2017.09.002

1. Introduction

Parkinson's Disease (PD) [1–3] is a progressive degenerative disease of the nervous system that affects movement control. This disease affects approximately 1% of the population over 60 years [4] with a prevalence of approximately 250 per 100,000 persons. The average age at onset lies between 55 and 65 years. The evidence shows that some environmental factors (e.g., consumption of well water, rural living and pesticide exposure) can increase an individual's risk of developing PD [5]. The PD burden is considerable and its diagnosis is clinical in nature. This disease includes four cardinal features which are: resting tremor (usually in a frequency of 4–7 Hz), rigidity, bradykinesia (slowness of movement), and impaired postural instability [6]. Three hierarchical levels have been defined by the World Health Organization (WHO) which a disease can be assessed. 'Impairment' reflects the basic level and in PD consists of clinical signs, such as tremor, bradykinesia and rigidity. 'Disability' depends on overlapping PD features that cause a functional decrease of activities of daily life. A multitude of assessment scales has been developed for the assessment of PD. The need for scales arose after the introduction of an effective therapy, so many were developed in the sixties and seventies. In 1987 the Unified Parkinson's disease rating scale (UPDRS) [7] was introduced, and in the subsequent years this scale has evolved into the most widely used scale for the assessment of PD. The UPDRS is a useful way to maintain an ongoing record of patient function and to assess disability [8]. According to Rascol et al. [9], UPDRS is the most widely used standardized scale to assess parkinsonism. UPDRS is made up of the (1) Motor sections, (2) ADL, (3) Mentation, Behavior and Mood. Motor-UPDRS and Total-UPDRS of UPDRS which refer respectively to the full range of the metric and the motor section of the UPDRS are used in the diagnosis of PD. The structure and metric properties of the UPDRS has been investigated in several studies on PD [10,11]. An example of a UPDRS chart measuring the change from baseline (mean ± SE) is shown in Fig. 1.

PD is a very complex disorder in which individual motor features vary in their presence and severity over time. The early diagnosis of PD is important. The problem of an early diagnosis of PD has raised an interest of numerous researchers. In addition, the researchers in the previous studies have emphasized that the main challenge in the diagnosis of PD is the correctly recognition of PD affected subjects at the early stage [12]. The early diagnosis of PD can affect on the progression of the disease and the quality of life the patients [13]. It has been emphasized that the main medical challenge is to correctly recognize the Parkinson's disease affected subjects at the early stage [12,14,15].

As early as 1997, the potential of data mining for improving the problems in the medical domain had been identified by World Health Organization (WHO) [16]. The usefulness of knowledge detection from medical data repositories has been emphasized by WHO as it benefits medical diagnosis and prediction.

Data mining is a process of discovering useful knowledge from database to build a structure (i.e., model or pattern) that can meaningfully interpret the data. It has been defined is as a process of discovering interesting patterns and knowledge from large amount of data [17]. It uses the machine learning techniques to discover hidden pattern in the data. These techniques can be in the three main categories which are supervised learning techniques, unsupervised learning techniques and semi-supervised learning techniques [18]. Expert systems developed by machine learning techniques can be used to assist physicians in diagnosing and predicting diseases [19]. Due to diseases diagnosis importance to mankind, several studies have been conducted on developing methods for their classification (see Table 1). Although these techniques can be used to predict the PD through a set of real-world datasets, however the most methods developed by supervised prediction techniques in the previous researches do not support the incremental updates of the data for PD prediction. Furthermore, standard supervised techniques cannot be used for the incremental learning and therefore they require to recompute all the training data to build prediction models. As the computation time and accuracy are two important criteria for the assessment of the medical diagnosis systems, in this research a new method using Incremental Support Vector Regression (ISVR) [20,21], an effective regression machine learning technique, is proposed in order to improve the predictive accuracy and reduce the computation time of PD prediction. The ISVR have been successfully used to solve real-time prediction problems. To the best knowledge of the

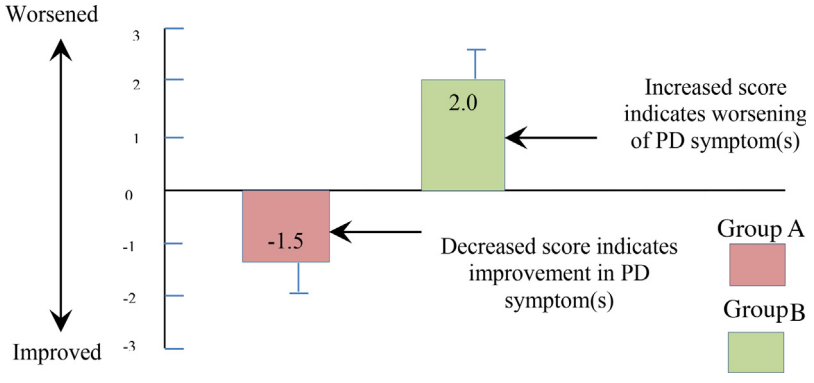

Fig. 1 – The effect of increasing and decreasing score on PD symptom(s).

Please cite this article in press as: Nilashi M, et al. A hybrid intelligent system for the prediction of Parkinson's Disease progression using machine learning techniques. Biocybern Biomed Eng (2017), http://dx.doi.org/10.1016/j.bbe.2017.09.002