

Prediction of Parkinson's Disease using Machine Learning Techniques

¹ A. Rajiv Kannan,
Head of the Department,
Department of CSE,
K.S.R. College of Engineering,
Tiruchengode, India.

² M. M. Kiruthiga, ³ C. Praveen Kumar,
⁴ A. Dhenaskumar, ⁵ S. Manoj,
^{2,3,4,5} UG Students, Department of CSE,
K.S.R. college of Engineering,
Tiruchengode, India.

Abstract—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 as a process of discovering interesting patterns and knowledge from large amount of data. 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. Expert systems developed by machine learning techniques can be used to assist physicians in diagnosing and predicting diseases. Due to diseases diagnosis importance to mankind, several studies have been conducted on developing methods for their classification. 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. K-mean clustering, standard supervised techniques cannot be used for the incremental learning and therefore they require recomputing all the training data to build prediction models. The method proposed in this study has been evaluated by a public datasets from UCI which have input and output parameters for PD diagnosis. In addition, compared to the bighealthcare data, the nature of the data in these datasets is not complex. In addition, in case of big healthcare data which can be complex datasets with unique characteristics, the future studies need to consider this issue in the development of new method in order to overcome the challenges of data processing time and take advantage of big data. Bayesian classification, as big healthcare data include multi-spectral, heterogeneous, imprecise and incomplete observations (e.g., diagnosis) which are derived from different sources, therefore new methods are needed and relying solely on conventional machine learning techniques may include some shortcomings in predicting the disease.

Keywords— PD Dataset, PCA Feature Extraction , Data Mining Classification Model, Bayesian Classification.

I. INTRODUCTION

Data mining is a dominant technology with prodigious potential to help the organization. The data mining tools predict the future trends, behaviors, knowledge driven decision. Data mining is a process of extracting the valuable information from large amounts of data. In other argument's data mining is

mining the knowledge from data. The classification of data mining system is classified based on different criteria such as types of data and data models. Data mining makes classification models by using already classified data and finds the predicted pattern. The classification problems are used to identify the features of group in each case of class.

The cause of Parkinson's disease (PD) is unknown, however re- search has shown that a degradation of the dopaminergic neurons affect the dopamine production to decline. Dopamine is used by the body to control movement, hence the less dopamine that is in circulation the more difficult the person has to control the movements and may experience tremors and numbness in extremities. As a direct cause of reduced control of motor-neurons in the central nervous system, the ability of articulating vocal phonetics is reduced. In this case the symptom (the inability to articulate words) is related to the presence of Parkinson's disease and is described as Dysphonia, a reduced functionality of the vocal cords. One of the immediate effects of vocal Dysphonia is that the voice is experienced as more course by fellow listeners. The features used in the prediction of Parkinson's disease in this study have been obtained from vocal records of people.

The field of speech processing and development of speech recognition systems have received considerable attention during the last decades. Separation of voice and background noise is important issues. With the emerge of portable phones and studio recording microphones analyzing methods involving traditional digital signal processing approaches such as hidden Markov models, Kalman filter, short- time frequency analysis and wavelet analysis have been success- fully used for both speech enhancement and speech recognition application.

II. RELATED WORKS

Luukka et al [1] describes a feature selection and plays an important role in classification for several reasons. First it can simplify the model and this way computational cost can be reduced and also when the model is taken for practical use fewer inputs are needed which means in practice that fewer measurements from new samples are needed. Second by removing insignificant features from the data set one can also make the model more transparent and more comprehensible,

providing better explanation of suggested diagnosis, which is an important requirement in medical applications. Feature selection process can also reduce noise and this way enhances the classification accuracy. In this article, feature selection method based on fuzzy entropy measures is introduced and it is tested together with similarity classifier. Model was tested with four medical data sets which were dermatology, Pima-Indian diabetes, breast cancer and Parkinson's data set. With all the four data sets, we managed to get quite good results by using fewer feature that in the original data sets. Also with Parkinson's and dermatology data sets, classification accuracy was man- aged to enhance significantly this way. Mean classification accuracy with Parkinson's data set being 85.03% with only two features from original 22. With dermatology data set, mean accuracy of 98.28% was achieved using 29 features instead of 34 original features. Results can be considered quite good.

Akin Ozcift et al [2] describe a machine learning algorithms and vital in designing high performance computer-aided diagnosis (CADx) systems. Researches' have shown that a base classifier performance might be enhanced by ensemble classification strategies. In this study, we construct rotation forest (RF) ensemble classifiers of 30 machine learning algorithms to evaluate their classification performances. In proposed system, first the feature dimension of three datasets is reduced using correlation based feature selection (CFS) algorithm. Second, classification performances of 30 machine learning algorithms are calculated for three datasets. Third, 30 classifier ensembles are constructed based on RF algorithm to assess performances of respective classifiers with the same disease data. All the experiments are carried out with leave-one-out validation strategy and the performances of the 60 algorithms are evaluated using three metrics; classification accuracy (ACC), kappa error (KE) and area under the receiver operating characteristic (ROC) curve (AUC). Base classifiers succeeded 72.15%, 77.52% and 84.43% average accuracies for diabetes, heart and Parkinson's datasets, respectively. As for RF classifier ensembles, they produced average accuracies of 74.47%, 80.49% and 87.13% for respective diseases. RF, a newly proposed classifier ensemble algorithm, might be used to improve accuracy of miscellaneous machine learning algorithms to design advanced CADx systems.

Freddie Åström et al [3] describe a parallel feed-forward neural network structure is used in the prediction of Parkinson's disease. The main idea of this paper is using more than a unique neural network to reduce the possibility of decision with error. The output of each neural network is evaluated by using a rule-based system for the final decision. The proposed prediction system is based on using parallel neural networks and evaluating the outputs to find the best prediction result. It is known that in parallel systems the reliability increases. In the same way it is evidently observed that the performance of the prediction has been increased in this paper compared to the use of a unique network. Using unlearned data in the next neural network also gave profound impact on the robustness of the system. It has also been shown that after a certain number of parallel networks, the accuracy of the prediction does not improve anymore.

Hui- Ling Chen et al [4] describe a swarm intelligence technique based support vector machine classifier (PSO _SVM) is proposed for breast cancer diagnosis. In the proposed PSO - SVM, the issue of model selection and feature

selection in SVM is simultaneously solved under particle swarm (PSO optimization) framework. A weighted function is adopted to design the objective function of PSO, which takes into account the average accuracy rates of SVM (ACC), the number of support vectors (SVs) and the selected features simultaneously. Furthermore, time varying acceleration coefficients (TVAC) and inertia weight (TVIW) are employed to efficiently control the local and global search in PSO algorithm. The effectiveness of PSO - SVM has been rigorously evaluated against the Wisconsin Breast Cancer Dataset (WBCD), which is commonly used among researchers who use machine learning methods for breast cancer diagnosis. The proposed system is compared with the grid search method with feature selection by F - score. The experimental results demonstrate that the proposed approach not only obtains much more appropriate model parameters and discriminative feature subset, but also needs smaller set of SVs for training, giving high predictive accuracy.

Musa Peker et al [5] describe a new approach for accurately diagnosing PD that can help medical personnel to make better and faster decisions. The proposed approach is capable of automatically analyzing data related to PD to develop prediction/diagnostic models with a high degree of accuracy in a relatively short time. The main novelty of the proposed study relates to the use of a hybrid methodology herein referred to as mRMR + CVANN, which integrates an effective feature selection method and a strong classifier. In this methodology, an effective feature set was obtained using an mRMR algorithm. Application of this algorithm resulted in a smaller feature set by eliminating less relevant features. Complex numbered features were then obtained from the optimally selected/reduced feature set. The complex-valued feature combinations produced and used in this study are among the most important contributions/innovations of the proposed method. A CVANN algorithm with high functionality and a very good classification capability was designed and developed during the classification stage of the proposed method. The prediction results obtained were very promising. Thus, a prediction system that can be used as a part of a computer-aided diagnosis system was developed. This system has the capability and potential to help doctors and other medical professionals in the diagnostic related decision processes for different diseases.

Hui- Ling Chen et al [6] explore the potential of extreme learning machine (ELM) and kernel ELM (KELM) for early diagnosis of Parkinson's disease (PD). In the proposed method, the key parameters including the number of hidden neuron and type of activation function in ELM, and the constant parameter C and kernel parameter γ in KELM are investigated in detail. In their study, Support Vector Machine (SVM) with Gaussian kernel functions in combination with the feature selection approach was taken to predict PD.

In this work, is to develop an efficient hybrid method, mRMR - KELM, for addressing PD diagnosis problem. The core component of the proposed method is the KELM classifier, whose key parameters are explored in detail. With the aid of the feature selection techniques, especially the mRMR filter, the performance of KELM classifier is ameliorated with much smaller features. The promising performance obtained on the PD dataset has proven that the proposed hybrid method can distinguish well enough between patients with PD and healthy persons.

Bo Yang et al [7] describe a parallel time variant particle swarm optimization (TVPSO) algorithm to simultaneously perform the parameter optimization and feature selection for SVM, termed PTVPSO-SVM. It is implemented in a parallel environment using Parallel Virtual Machine (PVM). In the proposed method, a weighted function is adopted to design the objective function of PSO, which takes into account the average classification accuracy rates (ACC) of SVM, the number of support vectors (SVs) and the selected features simultaneously. Furthermore, mutation operators are introduced to overcome the problem of the premature convergence of PSO algorithm. In addition, an improved binary PSO algorithm is employed to enhance the performance of PSO algorithm in feature selection task. The performance of the proposed method is compared with that of other methods on a comprehensive set of 30 benchmark data sets. The empirical results demonstrate that the proposed method cannot only obtain much more appropriate model parameters, discriminative feature subset as well as smaller sets of SVs but also significantly reduce the computational time, giving high predictive accuracy.

Alaa Tharwat et al [8] describe an important step in drug development. Nevertheless, the current experimental methods used to estimate the drug toxicity are expensive and time-consuming, indicating that they are not suitable for large-scale evaluation of drug toxicity in the early stage of drug development. The proposed model consists of three phases. In the first phase, the most discriminative subset of features is selected using rough set-based methods to reduce the classification time while improving the classification performance. In the second

phase, different sampling methods such as Random Under-Sampling, Random Over-Sampling and Synthetic Minority Oversampling Technique (SMOTE), Border Line SMOTE and Safe Level SMOTE are used to solve the problem of imbalanced dataset. In the third phase, the Support Vector Machines (SVM) classifier is used to classify an unknown drug into toxic or non-toxic. In this paper, Whale Optimization Algorithm (WOA) has been proposed to optimize the parameters of SVM, so that the classification error can be reduced. The experimental results proved that the proposed model achieved high sensitivity to all toxic effects. Overall, the high sensitivity of the WOA + SVM model indicates that it could be used for the prediction of drug toxicity in the early stage of drug development.

Cuicui Yang et al [9] describe a new swarm intelligence algorithm for structural learning of Bayesian networks, BFO-B, based on bacterial foraging optimization. In the BFO-B algorithm, each bacterium corresponds to a candidate solution that represents a Bayesian network structure, and the algorithm operates under three principal mechanisms: chemotaxis, reproduction, and elimination and dispersal. The chemotaxis mechanism uses four operators to randomly and greedily optimize each solution in a bacterial population, then the reproduction mechanism simulates survival of the fittest to exploit superior solutions and speed convergence of the optimization. Finally, an elimination and dispersal mechanism controls the exploration processes and jumps out of a local optima with a certain probability. We tested the individual contributions of four algorithm operators and compared with two states of the art swarm intelligence based algorithms and seven other well-known algorithms on many benchmark

networks. The experimental results verify that the proposed BFO-B algorithm is a viable alternative to learn the structures of Bayesian networks, and is also highly competitive compared to state of the art algorithms.

M. Hariharan et al [10] describe a hybrid intelligent system using Model-based clustering (Gaussian mixture model), feature reduction/selection using principal component analysis (PCA), linear discriminant analysis (LDA), sequential forward selection (SFS) and sequential backward selection (SBS), and classification using three supervised classifiers such as least-square support vector machine (LS-SVM), probabilistic neural network (PNN) and general regression neural network (GRNN). PD dataset was used from University of California-Irvine (UCI) machine learning database. The strength of the proposed method has been evaluated through several performance measures. The experimental results show that the combination of feature pre-processing, feature reduction/selection methods and classification gives a maximum classification accuracy of 100% for the Parkinson's dataset. The proposed integration of feature weighting method, feature reduction/selection method and classifiers gives a very promising classification accuracy of 100% which is closer to the results published in the literature. From the simulation results, we can also conclude that the proposed method may be instrumental to the physicians in detecting PWP accurately. In the future, the proposed method will be applied to other medical datasets to enhance the discriminatory power of the clinical features.

III. SYSTEM DESIGN

1. DATASET COLLECTION

In this module Parkinson's dataset add in the classification. In this module, the dataset attribute details like,

- 1) ID number
- 2) Outcome (R = recur, N = nonrecur)
- 3) Time (recurrence time if field 2 = R, disease-free time if field 2 = N)
- 4) Ten real-valued features are computed for each cell nucleus:
 - a) radius (mean of distances from center to points on the perimeter)
 - b) texture (standard deviation of gray-scale values)
 - c) perimeter
 - d) area
 - e) smoothness (local variation in radius lengths)
 - f) compactness ($\text{perimeter}^2 / \text{area} - 1.0$)
 - g) concavity (severity of concave portions of the contour)
 - h) concave points (number of concave portions of the contour)

2. PCA Feature Extraction

In this module, is used to extraction of feature using principal Component Analysis (PCA) analysis and to find Parkinson's size dimensionality reduction technique for classification for dataset. If all features in this feature vector were statistically independent, one could simply eliminate the

least discriminative features from this vector. The least discriminative features can be found by various greedy feature selection approaches.

The Parkinson's features depend on each other or on an underlying unknown variable. A single feature could therefore represent a combination of multiple types of information by a single value. Removing such a feature would remove more information than needed. In the next paragraphs, we introduce PCA as a feature extraction solution to this problem, and introduce its inner workings from two different perspectives.

3. KMEAN CLUSTERING (KMC)

Though points with highest hubness scores are without doubt the prime candidates for cluster centers, there is no need to disregard the information about hubness scores of other points in the data.

```

initializeClusterCenters();
Cluster[] clusters = formClusters();
float t = t0; initialize temperature
repeat
float  $\theta$  getProbFromSchedule(t);
for all Cluster  $c \in$  clusters do
if randomFloat(0,1) <  $\theta$  then
DataPoint h = findClusterHub(c);
setClusterCenter(c, h);
else
for all DataPoint  $x \in c$  do
setChoosingProbability(x,  $N_2 k(x)$ );
end for
normalizeProbabilities();
DataPoint h = chooseHubProbabilistically(c);
setClusterCenter(c, h);
end if
end for
clusters = formClusters();
t = updateTemperature(t);
until noReassignments
return cluster

```

It is nearly identical to HPC, the only difference being in the deterministic phase of the iteration, as the configuration cools down during the annealing procedure: instead of reverting to K-hubs, the deterministic phase executes K-means updates.

```

initializeClusterCenters();
Cluster[] clusters = formClusters();
float t = t0; {initialize temperature}
repeat

```

```

float  $\theta$  = getProbFromSchedule(t);
for all Cluster  $c \in$  clusters do
if randomFloat(0, 1) <  $\theta$  then
DataPoint h = findClusterCentroid(c);
setClusterCenter(c, h);
else
for all DataPoint  $x \in c$  do
setChoosingProbability(x,  $N_2 k(x)$ );
end for
normalizeProbabilities();
DataPoint h = chooseHubProbabilistically(c);
setClusterCenter(c, h);
end if
end for
clusters = formClusters();
t = updateTemperature(t);
until noReassignments
return clusters

```

4. SHARED-NEIGHBOR CLUSTERING

This method finds the similarity between individual data points using the nearest neighbor concept. This particular clustering algorithm can handle several issues related to clusters simultaneously like for e.g., it finds clusters of different sizes, shapes and densities from very large and high dimensional data sets. This algorithm first finds the list of nearest neighbors for each point and then redefines the similarity between points by the number of common neighbors between them. The shared neighbor similarity takes the sum of the similarity of the points nearest neighbors as a measure of density.

5. BAYESIAN NETWORK CLASSIFIERS

Bayesian network classifiers are used in many real Parkinson's Disease fields and one common class of classifiers are naive Bayesian classifiers. In this paper, we introduce an approach for reasoning about Bayesian network classifiers in which we explicitly convert them into Ordered Decision Diagrams (ODDs), which are then used to reason about the Parkinson's Disease properties of these classifiers. Specifically, we present an algorithm for converting any naive Bayesian classifier into an ODD, and we show simulation that this algorithm can give us an ODD that is tractable in size even given an intractable number of instances.

Since ODDs are tractable representations of classifiers, proposed algorithm allows us to efficiently test the equivalence of two naive Bayesian classifiers and characterize discrepancies between them. In proposed system also show a number of additional Parkinson's Disease results including a count of distinct classifiers that can be induced by changing some CPT in a naive Bayesian classifier, and the range of allowable changes to a CPT which keeps the current classifier unchanged.

A Bayesian network is a compact, graphical model of a probability distribution which assigns a probability to every event of interest. For example, in the rice Parkinson's Disease detection, a Bayesian network can be used to compute the probability of any particular Parkinson's Disease prediction given the season displayed by a rice Parkinson's Disease dataset.

Classification is a basic task in data analysis and pattern recognition that requires the construction of a classifier, that is, a function that assigns a class label to instances described by a set of attributes. The induction of classifiers from data sets of pre-classified instances is a central problem in machine learning. Numerous approaches to this problem are based on various functional representations such as decision trees, decision lists, neural networks, decision graphs, and rules. given C whenever $\Pr(A|B, C) = \Pr(A|C)$ for all possible values of A, B and C, whenever $\Pr(C) > 0$.

When represented as a Bayesian network, a naive Bayesian classifier has the simple

structure depicted in Figure 3.3.1. This network captures the main assumption behind the naive Bayesian classifier, namely, that every attribute (every leaf in the network) is independent from the rest of the attributes, given the state of the class variable (the root in the network).

$$LL(\text{Disease_B} | \text{Disease_D}) = N \sum_{i=1} \log PB(\text{Disease_ci} | \text{Disease_ai1}, \dots, \text{Disease_ain}) + N \sum_{i=1} \log PB(\text{Disease_ai1}, \dots, \text{Disease_ain})$$

This restriction was motivated mainly by computational considerations: these networks can be induced in a provably effective manner. This raises the question whether to achieve better performance at the cost of computational efficiency. One straightforward approach to this question is to search the space of all augmented naive Bayesian networks (or the larger space of Bayesian multinets) and select the one that minimizes the MDL score.

In this proposed system to examine a greedy search procedure. Such a procedure usually finds a good approximation to the minimal Parkinson's disease prediction scoring network. In this proposed system Parkinson's disease data set generated from a parity function and captured by augmenting the naive Bayes structure with a complete subgraph. However, the greedy procedure returned the naive Bayes structure, which resulted in a poor classification rate. The greedy procedure learns this network because Parkinson's Disease attributes are independent of each other given the class. As a consequence, the addition of any single edge did not improve the score, and thus, the greedy procedure terminated without adding any edges.

In this paper, we have analyzed the direct application of the Parkinson's disease prediction method to learning unrestricted Bayesian networks for classification tasks. The proposed system showed that, although the prediction method presents strong asymptotic guarantees, it does not necessarily optimize the classification accuracy of the learned networks. A proposed system analysis suggests a class of scoring functions that may be better suited to this task. These scoring functions appear to be computationally intractable, and we therefore plan

to explore effective approaches based on approximations of these scoring functions.

BAYESIAN NETWORKS ALGORITHM

- Step1: Read the Parkinson's disease dataset.
- Step 2: Create the data list from Parkinson's disease dataset and feature extraction for prediction Parkinson's Disease details.
- Step 3: Create the Bayesian net using Neuralnet package
- Step 4: To create a Bayesian network base net using model2network function in R-Studio.
- Step 5: To read the dataset and assign the data into CA, CS, Ck, Cw, CB, CL and CE object variable. The object contains Parkinson's disease year, Parkinson's disease production, Parkinson's disease Area, Parkinson's disease size, Parkinson's disease range, details is connected the Bayesian network.
- Step 6: The Parkinson's disease classification rule and probability values assign the Bayesian net.
- Step 7: To create custom Bayesian net using Bayesian theory in Parkinson's disease.
- Step 8: To check the Bayesian Rule for affect Parkinson's disease and return the accuracy values.
- Step 9: Repeat the process Step 3 to Step 8.
- Step 10: To accuracy calculate the TP, TN, FP and FN values.

MEASURES OF ACCURACY

Some measures of model accuracy like mean absolute error (MAE), mean absolute percentage error (MAPE), symmetric mean absolute percentage error (SMAPE), mean squared error (MSE) and root mean squared error (RMSE).

```
r.lm <- lm(Fertility ~ ., data=swiss)
```

```
MAE(r.lm)
```

```
# the same as:
```

```
MAE(predict(r.lm), swiss$Fertility)
```

```
MAPE(r.lm)
```

```
MSE(r.lm)
```

```
RMSE(r.lm)
```

IV. EXPERIMENTAL RESULTS

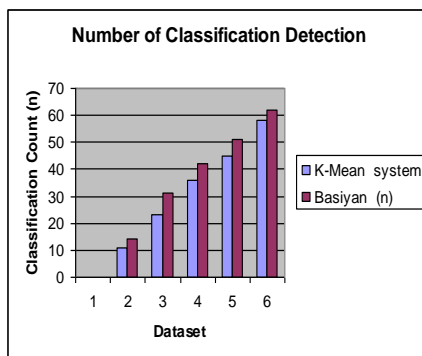


Fig 1 Number of Classification Detection

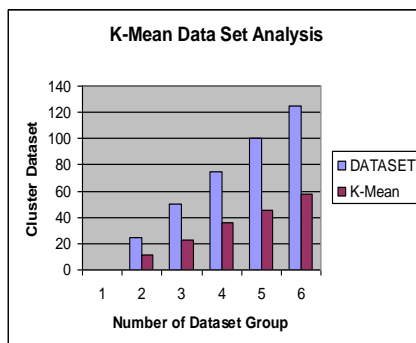


Fig 2 K-Mean Clustering

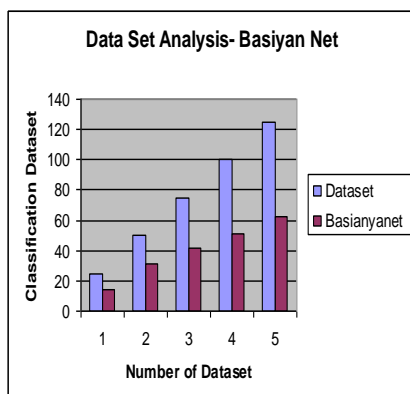


Fig 3 Bayesian Net Classification

Conclusion

The main findings of this paper are that the prediction method is integrated with dimensionality reduction and clustering techniques improved the accuracy prediction of PD and reduced the computation time. The superiority of the present method can be explained by the fact that our model support incremental updates of the data. In addition, the proposed method in this study supports incremental updates and re-learning of data which is more efficient in memory requirement. It is worth noting that our proposed method achieved the best performance on the PD dataset.

In this paper, plan to evaluate the proposed method on additional PD datasets and in particular on large datasets which includes other attributes for PD diagnosis to show the

effectiveness of the method for computation time of large data. In addition, our future work will investigate that how the proposed method can be extended to be applicable to the other types of datasets in medical domain

The testing application if developed as web service, then many applications can make use of it. The new system is designed such that those enhancements can be integrated with current modules easily with less integration work. The new system becomes useful if the above enhancements are made in future. The new system become useful if the below enhancements are made in future.

- The application if developed as web site can be used from anywhere.
- The factors used in the algorithm can be generalized so that default values produce the generic classification.
- The algorithm should segment only one image at time. In future, append concept for classification the multiple text file at same time.
- In future, the algorithm can be applied for pattern recognition. For identifying the similar pattern efficiently.

REFERENCES

- [1] Postuma RB, Montplaisir J. Predicting Parkinson's disease – why, when, and how? *Parkinsonism Relat Disord* 2009;15: S105–9.
- [2] Armañanzas R, Bielza C, Chaudhuri KR, Martinez-Martin P, Larrañaga P. Unveiling relevant non-motor Parkinson's disease severity symptoms using a machine learning approach. *Artif Intell Med* 2013;58(3):195–202.
- [3] Dunnewold RJW, Jacobi CE, Van Hilten JJ. Quantitative assessment of bradykinesia in patients with Parkinson's disease. *J Neurosci Methods* 1997;74(1):107–12.
- [4] Nutt JG, Wooten GF. Diagnosis and initial management of Parkinson's disease. *N Engl J Med* 2005;353(10).
- [5] Priyadarshi A, Khuder SA, Schaub EA, Priyadarshi SS. Environmental risk factors and Parkinson's disease: a metaanalysis. *Environ Res* 2001;86(2):122–7. 1021–7.
- [6] Renfroe JB, Bradley MM, Okun MS, Bowers D. Motivational engagement in Parkinson's disease: preparation for motivated action. *Int J Psychophysiol* 2016;99:24–32.
- [7] Daneault JF, Carignan B, Sadikot AF, Duval C. Are quantitative and clinical measures of bradykinesia related in advanced Parkinson's disease? *J Neurosci Methods* 2013;219(2):220–3.
- [8] Olanow CW, Watts RL, Koller WC. An algorithm (decision tree) for the management of Parkinson's disease (2001): treatment guidelines. *Neurology* 2001;56(Suppl. 5):S1–88.
- [9] Rascol O, Goetz C, Koller W, Poewe W, Sampaio C. Treatment interventions for Parkinson's disease: an evidence based assessment. *Lancet* 2002;359(9317): 1589–98.
- [10] Martignoni E, Franchignoni F, Pasetti C, Ferriero G, Picco D. Psychometric properties of the unified Parkinson's disease rating scale and of the short Parkinson's evaluation scale. *Neurol Sci* 2003;24(3):190–1.
- [11] Van Hilten JJ, Van Der Zwan AD, Zwinderman AH, Roos RAC. Rating impairment and disability in Parkinson's disease: evaluation of the Unified Parkinson's Disease Rating Scale. *Mov Disord* 1994;9(1):84–8.
- [12] Farnikova K, Krobot A, Kanovsky P. Musculoskeletal problems as an initial manifestation of Parkinson's disease