Parkinson's disease detection using olfactory loss and REM sleep disorder features

R. Prashanth-*EMBS Member*, S. Dutta Roy-*IEEE Member*, and P. K. Mandal, S. Ghosh

***Abstract*— In Parkinson's disease, there exists a prodromal or a premotor phase characterized by symptoms like olfactory loss and sleep disorders, which may last for years or even decades before the onset of motor clinical symptoms. Diagnostic tools based on machine learning using these features can be very useful as they have the potential in early diagnosis of the disease. In the paper, we use olfactory loss feature from 40-item University of Pennsylvania Smell Identification Test (UPSIT) and Sleep behavior disorder feature from Rapid eye movement sleep Behavior Disorder Screening Questionnaire (RBDSQ), obtained from the Parkinson's Progression Marker's Initiative (PPMI) database, to develop automated diagnostic models using Support Vector Machine (SVM) and classification tree methods. The advantage of using UPSIT and RBDSQ is that they are quick, cheap, and can be self-administered. Results show that the models performed with high accuracy and sensitivity, and that they have the potential to aid in early diagnosis of Parkinson's disease.**

I. Introduction

Parkinson's disease (PD) is usually accompanied by premotor or prodromal phase, that can last for years or even decades, between the onset of neurodegeneration and manifestation of the classic clinical motor symptoms [1-5]. Among the premotor PD symptoms, Rapid eye movement sleep Behavior Disorder (RBD) [6] and olfactory loss [7] are the common ones. Diagnostic tools based on machine learning are important as they can aid in the early diagnosis of the disease.

Existing literature shows that combination of premotor features (RBD and olfactory loss) is a promising approach in the preclinical diagnosis of PD [2, 4, 5, 8]. Hence, further analysis on these premotor symptoms is necessary [9]. As of now, there are only few studies which developed prediction models based on non-motor features for PD risk estimation [10-12]. In [10, 11], the authors use odor identification test data to develop a logistic model to obtain a probability of PD for a subject. However, this study did not use any other potential markers such as the REM behavior. Armañanzas et al. 2013 [12] use the non-motor features of cognitive impairment, psychiatric complications, autonomic dysfunction and sleep disturbance for classifying subjects into 3 severity categories namely, mild, moderate and severe. However, this study did not use olfactory loss feature which is an established pre-clinical marker for PD. Further, most of the sample subjects were under levodopa medication. Their classification accuracy is also low: 66.85% to 72.51% for predicting the Hoehn and Yahr severity index; and 74.05% to 80.00% for predicting the clinical impression of severity index for PD.

In this paper, we use the olfactory loss and REM features to develop prediction models using data from the Parkinson's Progression Marker's Initiative (PPMI) database [13].

R. Prashanth* (corresponding author) and Sumantra Dutta Roy are with the Department of Electrical Engineering, Indian Institute of Technology Delhi, India, (email: {*eez108051, sumantra}@ee.iitd.ac.in).

Pravat K. Mandal is with the Neuroimaging and Neurospectroscopy Laboratory, National Brain Research Centre, India (pravat@nbrc.ac.in).

Shantanu Ghosh is with the Martinos Center for Biomedical Imaging, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA (shantanu@nmr.mgh.harvard.edu).

II. Materials and Methods

A. Database and Study Cohort details

Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (www.ppmi-info.org/data). PPMI [13] is a comprehensive, observational, international, multi-center and the first large-scale study that mainly recruits *de-novo* PD subjects and age matched normal controls to identify and explore PD progression biomarkers.

Among the measures to evaluate olfactory function, identification tests are widely used where the subject has to identify the stimulus from a list of choices. The 40-item University of Pennsylvania Smell Identification Test (UPSIT) is among the most reliable (test-retest $r = 0.94$) and accurate smell identification test available [14]. UPSIT consists of four booklets and each booklet contains 10 odors, one per page. Stimuli (or odor) are contained in plastic microcapsules on each page and the subject scrapes the strip with a pencil, which releases the odor. Subject then, identifies the smell and marks the option that best describes the odor [15]. The number of odors correctly identified results in a total score which can be a maximum of 40.

The RBD Screening Questionnaire (RBDSQ) is a specific screening questionnaire to assess Rapid eye movement sleep Behavior Disorder (RBD) and was developed by Stiasny-Kolster et al. in 2007 [16]. Researchers studied the utility of RBDSQ and have found that it performed with high sensitivity and reasonable specificity [17-19]. We use the total score from items 1 to 9 (with a maximum total score of 12) from the RBDSQ [16] for the present study.

For our study, we use UPSIT and RBDSQ data from the PPMI database. We downloaded the data on 29th May 2013 and as on this date, it had RBDSQ and UPSIT data from 195 normal controls and 423 early PD (Hoehn and Yahr (HY) Stage 1 and 2 with mean ± standard deviation as 1.56 ± 0.50) subjects. Fig. 1 shows the box plots of UPSIT and RBDSQ scores for normal and early PD groups. Table 1 shows the

mean ± standard deviation of UPSIT and RBDSQ scores for normal and early PD population.

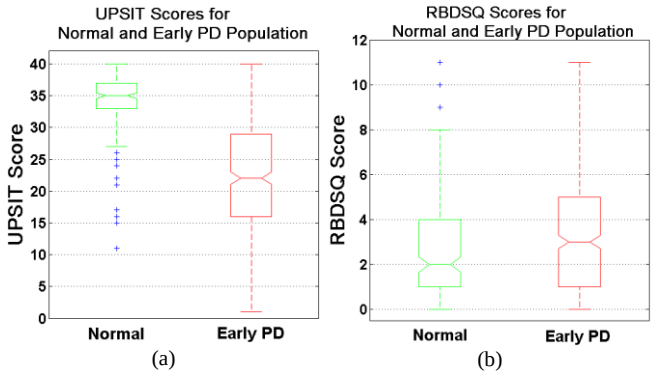


Figure 1. Box plots of (a) UPSIT score, and (b) RBDSQ score.

TABLE I. MEAN ± STANDARD DEVIATION OF UPSIT AND RBDSQ SCORES FOR NORMAL AND EARLY PD GROUPS.

Group	UPSIT	RBDSQ
Normal	34.05 ± 4.87	2.64 ± 2.14
Early PD	22.25 ± 8.22	3.23 ± 2.65

B. Prediction models for distinguishing Early PD from Normal Controls

A. Given a test sample $\mathbf{x}_k$; $\mathbf{x}_k = [UPSIT, RBDSQ]$ and $\mathbf{x}_k \in R^2$, the aim is to estimate the class membership y_k; $y_k \in [Early\ PD = 1, Normal = 0]$ based on the training data set D; $D = [(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_n, y_n)]$. Support Vector Machine (SVM) and classification tree are among the most widely used techniques for model building in biomedicine [20]. Classification tree technique models posterior probability $p(y_k \mid \mathbf{x}_k)$ of class y_k (which is a nominal variable), given input $\mathbf{x}_k$. On the other hand, SVMs are powerful classification tools which produce dichotomous outcomes without providing any information regarding class membership probability [20]. We used the LIBSVM (Library for Support Vector Machines) [21] for SVM modeling and the Statistics toolbox in MATLAB for classification tree modeling. 70% of the data is used for training and the rest 30% for testing the models. (Both the training and test sets have roughly the same class proportions as in the set of class labels).

1. Support Vector Machine

Support Vector Machine (SVM) is a supervised classification technique that finds a linear separating hyperplane by mapping input data to a higher-dimensional feature space through linear or nonlinear kernel functions [22]. The optimization problem in SVM is given by

$$\min_{\mathbf{w},b,\xi} \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{y_k=1}\xi_k \qquad (1)$$

where C is the regularization parameter and ξ_k is the slack variable used. The optimal $\mathbf{w}$ obtained after solving, satisfies

$$\mathbf{w} = \sum_{k=1}^{n} y_k \alpha_k \phi(\mathbf{x}_k) \qquad (2)$$

and the decision function is given by

$$sign(\mathbf{w}^T\phi(\mathbf{x}_k) + b) = sign\left(\sum_{k=1}^{n} y_k \alpha_k K(\mathbf{x}, \mathbf{x}_k) + b\right) \qquad (3)$$

$K(\mathbf{x}, \mathbf{x}_k)$ is the kernel function defined by $K(\mathbf{x}, \mathbf{x}_k) = \phi^T(\mathbf{x})\phi(\mathbf{x}_k)$; and α_i are the Lagrange multipliers. By using different kernel functions, varying degrees of nonlinearity and flexibility can be included in the model. For our analysis, we used the linear kernel defined by $K(\mathbf{x}, \mathbf{x}_k) = \mathbf{x}^T\mathbf{x}_k$, and the non-linear kernel of Radial Basis Function (RBF) defined by $K(\mathbf{x}, \mathbf{x}_k) = \exp(-\gamma \|\mathbf{x} - \mathbf{x}_k\|^2); \gamma > 0$ for the SVM classifier.

2. Classification Tree

Classification trees are non-parametric classifiers that repeatedly applies a split criterion that maximizes the separation of the data. This results in a tree-like structure which can be expressed as a set of "if-then" rules [20, 23, 24]. Non-significant predictors (branches) can be pruned from the final tree and removed from the analysis. The splitting criteria used in the study is the Gini diversity index (also called Gini impurity index). The steps carried out to build a classification tree are the following:

- Start with all input data, and examine all possible binary splits on every predictor.
- Select the split that minimizes the impurity of a node, given by the Gini diversity index (*gdi*) of the node. The *gdi* of a node *j* is given by

$$gdi(j) = 1 - \sum_{i=0}^{1} P(i \mid j)^2 \qquad (4)$$

$P(i \mid j)$ is the conditional probability of a class i; $i \in [Early\ PD = 1, Normal = 0]$, given the node *j*. A pure node (node with just one class) has a Gini index of 0, otherwise the Gini index is positive.

- Impose the split on the parent node to obtain two child nodes (leaves). The first parent node is the root node.
- Repeat recursively for the two child nodes.
- Stop splitting if the node is pure, or there are fewer than specified observations in the parent node, or any split imposed on the parent node would produce children (leaf node) with fewer than specified observations.

III. RESULTS AND DISCUSSION

Table II shows the performance of SVM and classification tree classifiers used in this study. The positive

predictive values (PPV) were comparable for the classifiers. But, comparing the accuracies and the negative predictive values (NPV) obtained for the test set, we observe that the SVM classifier using the RBF kernel performed the best with more than 85 % accuracy and more than 78% NPV for test data classification. SVM-RBF's ability to incorporate nonlinearity in the data may be the reason that it scores over other models. Fig. 2 shows the ROC plots for the classifiers used.

TABLE II. PERFORMANCES MEASURES FOR SVM CLASSIFIER WITH LINEAR KERNEL (SVM-LINEAR), SVM CLASSIFIER WITH RBF KERNEL (SVM-RBF), AND CLASSIFICATION TREE (CT) MODEL

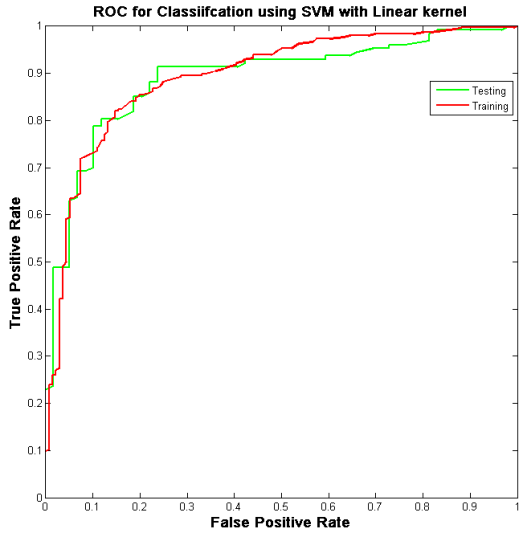
Performance measures	**Training**			**Testing**		
	SVM-Linear	***SVM-RBF***	***CT***	***SVM-Linear***	***SVM-RBF***	***CT***
True Positive	253	260	248	112	115	106
False Positive	28	32	20	13	15	13
True Negative	108	104	116	46	44	46
False Negative	43	36	48	15	12	21
PPV (%)	90.04	89.04	92.54	89.60	88.46	89.08
NPV (%)	71.52	74.29	70.73	75.41	78.57	68.66
Accuracy (%)	83.56	84.26	84.26	84.95	85.48	81.72
Sensitivity (%)	85.47	87.84	83.78	88.19	90.55	83.46
Specificity (%)	79.41	76.47	85.29	77.97	74.58	77.97
AUC (%)	89.39	89.12	79.87	88.93	88.22	86.81

70% (296 early PD + 136 normal = 432 observations) of the data was used for training and the rest 30% (127 early PD + 59 normal = 186 observations) was used for testing the models. True Positive, True Negative, False Positive, False Negative are in number of samples, and positive predictive value (PPV), negative predictive value (NPV), accuracy, sensitivity, specificity and area under the ROC curve (AUC) are in percentage.

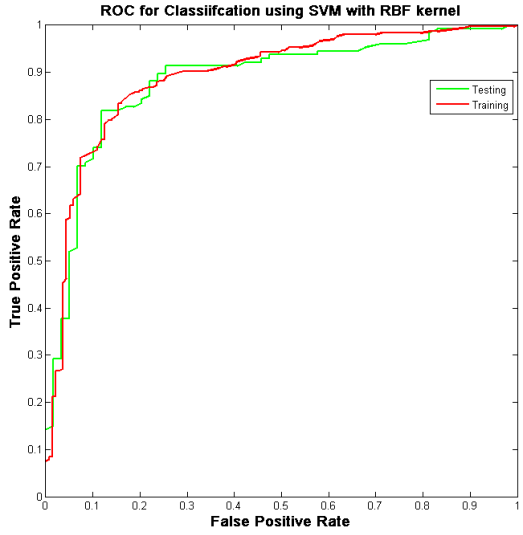
A. Limitations and future work

Studies show that up to 60% patients diagnosed with RBD develop neurodegenerative disorders such as PD [25]. However, RBDSQ itself does not reflect the complete spectrum of RBD features. (It was intended as a screening tool for RBD, and not for its clinical diagnosis). This may be one of the reasons for the lower discriminatory value of the RBDSQ score as compared to the UPSIT score. Figure 2 demonstrates this. It is observed that UPSIT score had higher separation, between normal and early PD, than that for RBDSQ score. As these proposed diagnostic models using both UPSIT and RBDSQ scores show the potential in discriminating early PD from normal control, they can be used as a primary screen for PD; and polysomnography (PSG) which is the gold standard for revealing loss of REM-related muscle atonia in RBD and dopaminergic imaging using Single Photon Emission Computed Tomography (SPECT) with ^{123}I-Ioflupane (DaTSCAN) which has shown to be sensitive to early PD [4] can be used as a secondary screen for PD. The advantage of using UPSIT and RBDSQ is that they are cheap, quick and they can be self-administered; unlike the PSG and SPECT imaging which are expensive and time taking process.

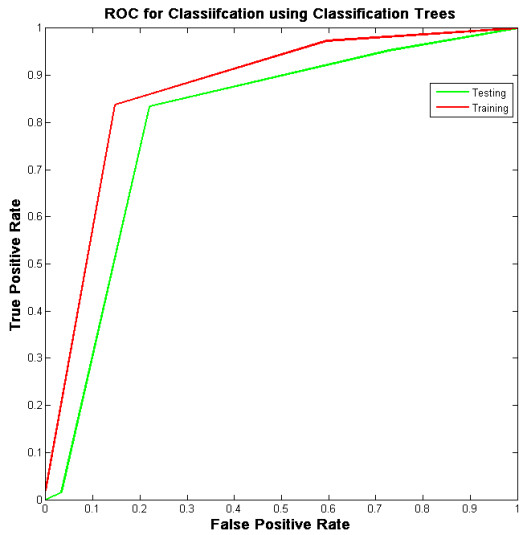
Along with differentiating PD from normal, these features also show to be useful in discriminating other disorders which have almost similar clinical symptoms like in PD such as multiple system atrophy, progressive supranuclear palsy, corticobasal degeneration and essential



(a)



(b)



(c)

Figure 2. ROC curves for classification using (a) SVM with linear kernel (b) SVM classifer with RBF kernel and (c) classification tree.

tremor [26]. Proper diagnosis of these disorders is important as misdiagnosis as PD can lead to unnecessary medical examinations and therapies and associated side-effects. As a future work, validity of the developed models for the diagnosis of PD can be carried out. Dopaminergic transporter imaging have shown to be sensitive marker in PD even in the early stages of the disease. A combination of premotor features (RBD and olfactory loss) with dopaminergic imaging to develop models can lead to encouraging results in the preclinical diagnosis of PD [2, 4, 5, 8] which we are looking as a possible future work. A comparison of the present approach with other valid test instruments, such as the Movement Disorder Society-Unified Parkinson's Rating Scale (MDS-UPDRS) that is used clinically and for research purposes in the PD detection problem, is another possible future work.

IV. Conclusion

Non-motor features like the olfactory loss and REM sleep behavior disorder predates, by years or even decades, the occurrence of first clinical symptoms in PD. Further analysis and diagnostic tools using machine learning techniques based on them have immense potential which can help in the early diagnosis of PD. In this work, we developed prediction models using machine learning techniques of logistic regression, classification trees and support vector machine based on these statistically significant features and observe that the performance to be high. These models have the potential to be used in screening for PD as they can estimate the risk of PD with high accuracy, thereby these models can aid in the early diagnosis of PD.

Acknowledgment

PPMI – a public-private partnership – is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, including [list of names of all of the PPMI funding partners at www.ppmi-info.org/ fundingpartners].

References

[1] A. Gaenslen, I. Swid, I. Liepelt-Scarfone, J. Godau, and D. Berg, "The patients' perception of prodromal symptoms before the initial diagnosis of Parkinson's disease," *Mov Disord*, vol. 26, no. 4, pp. 653-8, Mar, 2011.

[2] H. W. Berendse, and M. M. Ponsen, "Diagnosing premotor Parkinson's disease using a two-step approach combining olfactory testing and DAT SPECT imaging," *Parkinsonism Relat Disord*, vol. 15 Suppl 3, pp. S26-30, Dec, 2009.

[3] E. Tolosa, Y. Compta, and C. Gaig, "The premotor phase of Parkinson's disease," *Parkinsonism Relat Disord*, vol. 13 Suppl, pp. S2-7, Sep, 2007.

[4] J. L. Cummings *et al.*, "The role of dopaminergic imaging in patients with symptoms of dopaminergic system neurodegeneration," *Brain*, vol. 134, no. Pt 11, pp. 3146-66, Nov, 2011.

[5] K. Stiasny-Kolster *et al.*, "Combination of 'idiopathic' REM sleep behaviour disorder and olfactory dysfunction as possible indicator for alpha-synucleinopathy demonstrated by dopamine transporter FP-CIT-SPECT," *Brain*, vol. 128, no. Pt 1, pp. 126-37, Jan, 2005.

[6] R. B. Postuma *et al.*, "Quantifying the risk of neurodegenerative disease in idiopathic REM sleep behavior disorder," *Neurology*, vol. 72, no. 15, pp. 1296-300, Apr 14, 2009.

[7] M. M. Ponsen *et al.*, "Idiopathic hyposmia as a preclinical sign of Parkinson's disease," *Ann Neurol*, vol. 56, no. 2, pp. 173-81, Aug, 2004.

[8] W. G. Meissner *et al.*, "Priorities in Parkinson's disease research," *Nat Rev Drug Discov*, vol. 10, no. 5, pp. 377-93, May, 2011.

[9] R. B. Postuma, and J. Montplaisir, "Predicting Parkinson's disease - why, when, and how?," *Parkinsonism Relat Disord*, vol. 15 Suppl 3, pp. S105-9, Dec, 2009.

[10] L. Silveira-Moriyama *et al.*, "The use of smell identification tests in the diagnosis of Parkinson's disease in Brazil," *Mov Disord*, vol. 23, no. 16, pp. 2328-34, Dec 15, 2008.

[11] L. Silveira-Moriyama *et al.*, "The use of a color coded probability scale to interpret smell tests in suspected parkinsonism," *Mov Disord*, vol. 24, no. 8, pp. 1144-53, Jun 15, 2009.

[12] R. Armananzas, C. Bielza, K. R. Chaudhuri, P. Martinez-Martin, and P. Larranaga, "Unveiling relevant non-motor Parkinson's disease severity symptoms using a machine learning approach," *Artif Intell Med*, vol. 58, no. 3, pp. 195-202, Jul, 2013.

[13] K. Marek *et al.*, "The Parkinson Progression Marker Initiative (PPMI)," *Prog Neurobiol*, vol. 95, no. 4, pp. 629-635, 2011.

[14] R. L. Doty, P. Shaman, and M. Dann, "Development of the University of Pennsylvania Smell Identification Test: a standardized microencapsulated test of olfactory function," *Physiol Behav*, vol. 32, no. 3, pp. 489-502, Mar, 1984.

[15] M. A. Fornazieri, R. Pinna Fde, T. F. Bezerra, M. B. Antunes, and R. L. Voegels, "Applicability of the University of Pennsylvania Smell Identification Test (SIT) in Brazilians: pilot study," *Braz J Otorhinolaryngol*, vol. 76, no. 6, pp. 695-9, Nov-Dec, 2010.

[16] K. Stiasny-Kolster *et al.*, "The REM sleep behavior disorder screening questionnaire--a new diagnostic instrument," *Mov Disord*, vol. 22, no. 16, pp. 2386-93, Dec, 2007.

[17] T. Nomura, Y. Inoue, T. Kagimura, Y. Uemura, and K. Nakashima, "Utility of the REM sleep behavior disorder screening questionnaire (RBDSQ) in Parkinson's disease patients," *Sleep Med*, vol. 12, no. 7, pp. 711-3, Aug, 2011.

[18] T. Miyamoto *et al.*, "The REM sleep behavior disorder screening questionnaire: validation study of a Japanese version," *Sleep Med*, vol. 10, no. 10, pp. 1151-4, Dec, 2009.

[19] Y. Nihei *et al.*, "REM sleep behavior disorder in Japanese patients with Parkinson's disease: a multicenter study using the REM sleep behavior disorder screening questionnaire," *J Neurol*, vol. 259, no. 8, pp. 1606-12, Aug, 2012.

[20] S. Dreiseitl, and L. Ohno-Machado, "Logistic regression and artificial neural network classification models: a methodology review," *J Biomed Inform*, vol. 35, no. 5-6, pp. 352-9, Oct-Dec, 2002.

[21] C.-C. Chang, and C.-J. Lin, "LIBSVM: A Library for Support Vector Machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1--27, 2011.

[22] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in Proceedings of the 5th Annual Workshop on Computational Learning Theory (COLT'92), 1992, pp. 144-152.

[23] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*, New York: Chapman & Hall (Wadsworth, Inc.), 1984.

[24] R. J. Lewis, "An Introduction to Classification and Regression Tree (CART) Analysis," in Annual Meeting of the Society for Academic Emergency Medicine, San Francisco, California, 2000.

[25] S. Fulda, "Idiopathic REM sleep behavior disorder as a long-term predictor of neurodegenerative disorders," *EPMA J*, vol. 2, no. 4, pp. 451-8, Dec, 2011.

[26] G. K. Wenning *et al.*, "Olfactory function in atypical parkinsonian syndromes," *Acta Neurol Scand*, vol. 91, no. 4, pp. 247-50, Apr, 1995.