Analysis of Visually Guided Tracking Performance in Parkinson's Disease

Yi Liu[1,2], Chonho Lee[2], Bu-Sung Lee[2], James K.R. Stevenson[3], and Martin J. McKeown[3]

[1]Nanyang Institute of Technology in Health & Medicine, Interdisciplinary Graduate School,
Nanyang Technological University, Singapore
[2]School of Computer Engineering, Nanyang Technological University, Singapore
[3]Department of Neuroscience, University of British Columbia, Vancouver, BC, Canada

***Abstract*— Recent studies have suggested significant differences in motor performances of Parkinson's Disease (PD) patients who have L-dopa induced dyskinesias (LIDs), even when off of L-dopa medication. The pathophysiology of LIDs remains obscure, so applying data-mining techniques to the patients' motor performance may provide some heuristic insight. This paper investigated visually-guided tracking performance of PD patients using data mining techniques to reveal the differences between dyskinesia and non-dyskinesia patients. We found that K-means clustering of the root mean square (RMS) tracking error at faster tracking speeds and with ambiguous visual stimuli was able to effectively discriminate between the two groups with 77.8% accuracy. Decision tree classification was less accurate (68.4%) and determined that years since diagnosis was the best feature to distinguish between groups. Our results suggest that data mining methodologies may provide novel insights into features of the neurovegetative disease.**

***Index Terms*— Parkinson's disease, Data Mining, Dyskinesia, Tracking performance**

I. INTRODUCTION

As a concept that emerged twenty years ago, data mining helps to get deep insights and facilitate unprecedented understanding of large data sets. Data mining has potential in revealing the hidden meaningful correlations, patterns and trends of information extraction from biomedical data sets. Recently it has become popular in healthcare and medicine, such as identifying health insurance fraud, classifying at-risk people, predicting healthcare cost, discovering relationships between health conditions and disease, and supporting clinical decision making [1].

Parkinson's disease (PD) is a common neuro-degenerative disorder of the central nervous system [2] , and the main motor features of the disease include tremor, bradykinesia (slowness of movement), rigidity (stiffness of the limbs and trunk), and postural instability [3]. Levodopa (L-dopa) is the most effective drug improving the motor disorders markedly, but the long-term L-dopa administration might result in L-dopa-induced dyskinesias (LIDs), which are involuntary writhing movements [4], [5].

Increasingly, data mining techniques (classification, clustering, association) are being applied in PD. For example, focusing on speech articulation difficulty symptoms of PD, three classification methods, e.g., tree classifier, statistical classifier and supportive classifier, have been used to make prediction of PD [6]. Accurate assessment of motor performance in PD subjects with dyskinesia is required to give informed medication suggestions. Thus, a support vector machine (SVM) classifier has been implemented to evaluate the severity of tremor, and detect motor fluctuations using wearable sensors [7], [3]. Guided by these existing data mining techniques in PD, we try to find if data mining techniques can be applied into analysis of tracking performance.

Previous studies have indicated that PD subjects have difficulty performing visually-guided tasks, in addition to internally-guided tasks. Studies have also assessed motor performance of both non-dyskinesia and dyskinesia using statistic methods. It demonstrated not only the presence of LIDs clearly differentiate motor performance of dyskinesia from non-dyskinesia while on medication, but also subtle differences in tracking performance can be detected between non-dyskinesia and dyskinesia even when off of medication [8], [2], [9]. In this paper, we utilize two data mining techniques to analyse visually-guided tracking performance between non-dyskinesia and dyskinesia. Consistent with previous studies, we calculated the root mean square (RMS) tracking error of non-dyskinetic and dyskinetic patients' tracking performance under different conditions. However, in this paper we used the K-means method to find which conditions are effective to cluster them into two groups. We also applied the decision tree algorithm to classify non-dyskinetic and dyskinetic patients based upon five clinical characteristics (e.g., Age, Years Since Diagnosis).

The contributions of the paper are as follows:

- We investigate visually guided tracking performance of both dyskinesia and non-dyskinesia PD patients and verify that they can be differentiated using data mining technique based on their performance monitored under different tracking conditions.
- We attempt to reveal the most effective tracking conditions (i.e., dataset attributes) to classify the PD patients into two different groups, and demonstrate that the highest classification accuracy is when using dataset attributes of fast-speed and with-noise tracking conditions.
- We study how data mining and statistical analysis complement each other by looking at clustering results (e.g., distance between cluster centroids) and statistical significance (e.g., p-value on t-test) of tracking performance.

978-1-4799-6644-8/14/$31.00 ©2014 IEEE

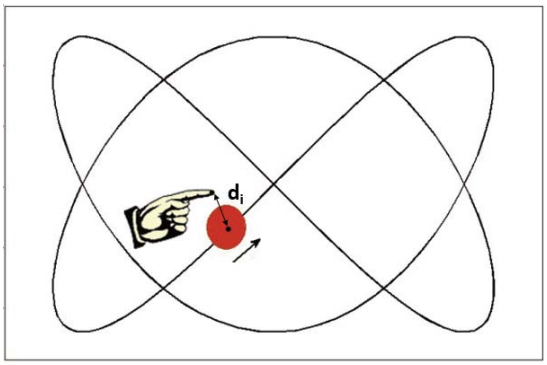

Fig. 1. The closed Lissajous figure measured 1.62 m by 1.22m. At time point i, the relative distance between the index finger and the target is d_i.

II. METHODOLOGY

A. Data Collection

In this study, we recruited nineteen clinically defined PD subjects (Hoehn and Yahr stage: I-III) [10], ten non-dyskinetic PD (NDPD) subjects and nine dyskinetic PD (DPD) subjects. Before the study all patients had stopped the medications of L-dopa overnight for at least 12 hours, and 18 hours for dopamine agonists, and thus they were clinically off medication. The motor severity of this disease was rated based on the Unified Parkinson's Disease Rating Scale (UPDRS) motor score in the off medication state (Table II).

A closed Lissajous curve was shown on a screen (Width: 1.62 m, Height: 1.22 m) with a red circular target (the diameter is 12 cm). The target moved through the Lissajous curve (Fig. 1). Subjects stood about half a meter in front of the screen, and tracked the moving visual circular target with the index finger [9]. A six degrees-of-freedom electromagnetic tracking system (Polhemus Fastrak [Polhemus, Colchester, VT, USA]) was used to record subjects' tracking. Data were recorded at 10 Hz, and then for each trial of each subject, a robust linear regression analysis was performed to determine the optimum affine transformation. The transformation was to map the sensor data coordinates to Lissajous figure coordinates in 2-dimension [9].

B. Attributes Selection

In this task, there was one no noise (No-N) and three visual ambiguity (i.e. "noise" (With-N)) conditions, i.e., small noise (SN), medium noise (MN) and large noise (LN), as shown in Fig. 2. There were two tracking speeds for each of above conditions, a slow speed (SS) and a fast speed (FS), so totally eight conditions. For each condition, each subject performed six individual trials at one state of either Off medication (OFF-M) and On medication (ON-M). Finally there were 96 trials in total for each subject. However, one dyskinetic subject did not complete all trials, and we omitted this subject in all subsequent calculations (but later for using decision tree, we did not omit this subject).

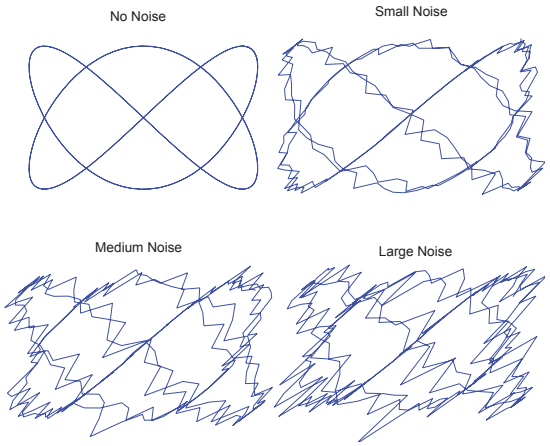

Fig. 2. The target trajectory in different noise.

TABLE I
K-MEANS CLUSTERING ALGORITHM. [11]

Choose K points from all points as the initial centroids randomly .
repeat
  Assign every point to the closest centroid to form K clusters .
  Recompute each cluster's centroid.
until The centroids would not change

For each individual trial, we got the value of root mean square (RMS) tracking error, which was calculated by squaring the distance between index finger position and target position for each time point, averaging the squared values, and finally taking the square root of the result.

$$X_{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} d_i^2} \quad (1)$$

In Equation (1), d_i is the relative distance between two points (the index finger and the target) at time point i (Fig. 1).

We used root mean square (RMS) tracking error as attributes, totally 96 attributes for each subject.

C. Application of Data Mining Techniques

In this paper, our main purpose was to support the previous statistical results, and find how data mining techniques and statistical methods complement each other, so we applied basic data mining techniques. We tried to find whether it could provide a new heuristic insight into clinically relevant topics. We did not focus on the data mining algorithm itself or comparing different algorithms.

We utilized two representative data mining techniques, K-means and decision tree, which are basic techniques in the data mining field. Using the K-means clustering algorithm (Table I), we need to set the initial value of K. We initially set k = 3, but the derived clusters did not appear meaningful. In this paper, since we mainly tried to find which conditions are effective for separating NDPD and DPD, we set two clusters, i.e. k = 2. Thus our expectation was that all non-dyskinetic subjects would be clustered into a cluster, and

likewise all dyskinetic subjects would be clustered into the other cluster.

Since the K-means clustering algorithm is unsupervised, there would be no label in clustering results. But we simply modified the K-means algorithm to a supervised-like algorithm with labeling clusters based on subjects distribution from the results. This scenario is that we labeled a cluster as non-dyskinesia if the majority of the nodes are non-dyskinetic subjects. Likewise we labeled the other group as "dyskinesia". The performance of the algorithm is then evaluated based on classification accuracy. Then we extracted the conditions, i.e., SS, FS, With-N, No-N, OFF-M, ON-M, to find out which conditions were more effective for clustering non-dyskinesia and dyskinesia. Meanwhile Principal Components Analysis (PCA) was also conducted to verify the results.

In previous studies, we did not find significant differences between non-dyskinetic and dyskinetic subjects in each clinical characteristics using statistical methods (i.e. t-test). In this paper, we applied the decision tree algorithm, and tried to combine the characteristics to find whether it could separate the non-dyskinetic and dyskinetic subjects. Meanwhile we also observe the structure of the distribution. For the splitting criterion, GainRATIO which is well known in C4.5 [12] was used in this paper.

$$GainRATIO_{split} = \frac{Gain_{split}}{SplitINFO} \tag{2}$$

$$SplitINFO = -\sum_{i=1}^{k} \frac{n_i}{n} \log \frac{n_i}{n} \tag{3}$$

$$Gain_{split} = Entropy(p) - (\sum_{i=1}^{k} \frac{n_i}{n} Entropy(i)) \tag{4}$$

$$Entropy(i) = -\sum_{a} p(a|i) \log p(a|i) \tag{5}$$

In Equation (2-5), the parent node is p that is split into k sub-nodes, and n_i is the number of records in sub-node i. While p(a|i) is the relative frequency of label a at node i.

According to the GainRATIO-based decision tree, we could get which factor(characteristic) was at the first-level node (root node) that means it is the optimal split characteristic among these five characteristics. So roughly we could know which characteristic is more effective for classifying NDPD and DPD.

III. RESULTS AND DISCUSSION

A. K-means Clustering

We used root mean square (RMS) tracking error as attributes described in Section II.B. Fig. 3 shows the results of K-means clustering. X-axis stands for subjects, and first ten subjects (1-10) are NDPD, last eight (11-18) are DPD. Y-axis stands for the cluster label. Since subjects with dyskinesia or non-dyskinesia are known in advance, the accuracy of clustering can be calculated. First, Fig. 3(a) shows the result of using all attributes. In NDPD (1-10), NDP subject 2, 3 are misclustered into a cluster which is more like a DPD cluster. In DPD (11-18), DPD subject 17, 18 are also misclusterted into the other cluster which is more like a NDPD cluster. So the accuracy of clustering equals 77.8%.

TABLE II
SUBJECTS' CHARACTERISTICS. [9]

ID	Subject	Age	Years Since Diagnosis	*UPDRS	Converted Daily L-dopa Dosage(mg)	L-dopa Equivalent Daily Dose (mg)
NDPD						
1	ND1	63	5	8	320	620
2	ND2	68	4	19	400	400
3	ND3	64	9	69	860	860
4	ND4	59	9	14	740	740
5	ND5	45	4	11	780	780
6	ND6	65	9	51	640	1003.3
7	ND7	63	10	54	800	1000
8	ND8	66	7	22	640	673.3
9	ND9	62	5	31	400	400
10	ND10	59	12	47	400	775
DPD						
11	D1	65	22	67	650	750
12	D2	64	7	42	880	1173.3
13	D3	68	13	51	660	880
14	D4	65	15	57	720	960
15	D5	66	5	45	1020	1020
16	D6	64	4	22	1280	1580
17	D7	51	7	37	800	1000
18	D8	55	13	40	640	665
19	**D9	75	8	47	600	600
	P-value	0.75	0.16	0.14	0.047	0.062

*UPDRS, Unified Parkinson's Disease Rating Scale
**D9, this subject did not complete all trials

TABLE III
ATTRIBUTES SET

	Fast speed		Slow speed	
	No noise	Noise	No noise	Noise
Off medication	①	②	③	④
On medication	⑤	⑥	⑦	⑧

In order to find out which attributes are more contributive for clustering NDPD and DPD, we selected 6 different attribute sets respectively. In Table III, dataset attributes of slow speed (SS) condition are ③④⑦⑧, likewise fast speed (FS) condition ①②⑤⑥, noise (With-N) condition ②④⑥⑧, no noise (No-N) condition ①③⑤⑦, off medication condition (OFF-M) ①③③④, on medication condition (ON-M) ⑤⑥⑦⑧.

Fig. 3(b) and (c) show the results of using SS and FS conditions, and the clustering accuracy is 66.7% and 77.8%, respectively. Fig. 3(d) and (e) show the results of using With-N and No-N conditions, and the clustering accuracy is 77.8% and 55.6%, respectively. Fig. 3(f) and (g) show the results of using OFF-M and ON-M conditions, and the clustering accuracy is 66.7% and 66.7%, respectively. From above results, the accuracy of using FS and With-N conditions is best (Accuracy = 77.8%), i.e. clinically it is relatively easier to separate NDPD and DPD subjects under FS and With-N conditions. This result is consistent with previous studies which reported that subjects with dyskinesia are generally

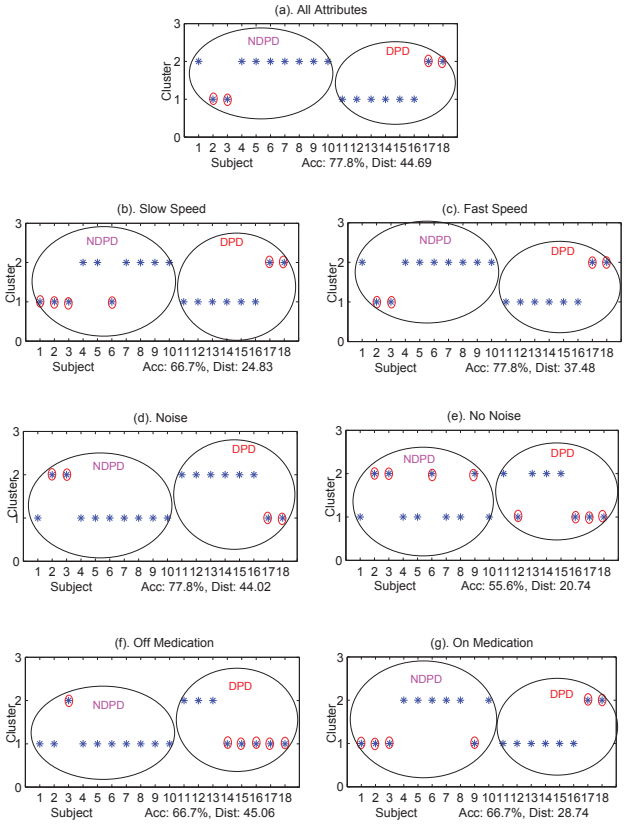

Fig. 3. The results of clustering using tracking root mean square (RMS) errors. "ACC" is the accuracy of clustering, "Dist" is squared Euclidean distance between centroids of two clusters (DCC).

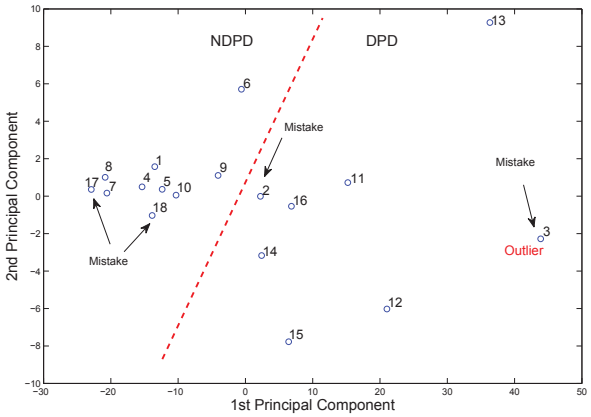

Fig. 4. Subspace with Principal Components

TABLE IV
EIGENVECTORS OF THE FIRST THREE PRINCIPAL COMPONENTS.

speed	noise	attribute ID	pc1	pc2	pc3
FS	LN	87	0.287	-0.235	0.243
FS	LN	91	0.243	-0.07	0.154
FS	LN	85	0.235	-0.357	0.161
FS	SN	47	0.231	0.023	-0.468
FS	LN	89	0.22	-0.284	-0.028
FS	MN	67	0.212	0.056	-0.01
FS	LN	95	0.208	0.066	-0.101
FS	MN	71	0.207	0.073	0.229
FS	LN	93	0.186	-0.155	-0.103
FS	MN	63	0.179	-0.044	0.025
FS	SN	39	0.175	-0.153	0.14
FS	MN	69	0.166	-0.031	-0.018
SS	LN	79	0.164	0.231	0.032
SS	MN	51	0.161	0.095	0.282
SS	LN	81	0.157	0.3	-0.027

more responsive to dynamic changes in motor task in which causes larger RMS tracking error.

As for Off-M and On-M conditions, although they had the same accuracy, for the distance between centroids of two clusters (DCC), Off-M case is better than On-M case because the DCC of Off-M is larger, which means the two clusters are separated slightly better, and this results also support our previous statistical conclusion [13] that L-dopa medication is able to improve tracking performance so that it is harder to classify under the On-M case. In addition, we also made some combinations, such as a combination of FS and With-N, FS and No-N, or Off-M and With-N and so on, but all the results are not better than before (accuracy = 77.8%), so attributes shown in Fig. 3 are representative for clustering.

In the results of clustering, subject 3, 17, 18 always are misclustered. NDPD subject 3 is always clustered into the DPD cluster, inversely DPD subject 17, 18 are always clustered into the NDPD cluster, i.e. tracking performance of NDPD subject 3 is more like DPD, and tracking performance of DPD subject 17, 18 is more like NDPD. However, corresponding UPDRS scores (Table II) of those three subjects could be a reason why they are misclustered. In NDPD group, the UPDRS score subject 3 is much higher than others, which means the subject would have more serious motor disorder though he is diagnosed as NDPD. Likewise, subject 17, 18 have relatively lower UPDRS score in the DPD group. Diagnosing NDPD and DPD is a clinically subjective decision based on patients' symptoms, and this result might assist doctors to make a decision more objectively.

In order to support the conclusion of clustering from the perspective of data, we also conducted Principal Components Analysis (PCA) to find the effective attributes to distinguish NDPD from DPD. During PCA, we found that there is an outlier, i.e., one of the NDPD subject (ND3) as shown in Figure 4. We removed the subject from the dataset and performed K-mean algorithm in a subspace with principal components.

Table IV shows eigenvectors of the first three principal components. The first principal component has more weight on attributes of FS and LN attributes compared to the others. Table V shows the clustering accuracy and DCC in the subspace. In order to see how the data mining technique supports statistical analysis, we compared the DCC cluster distance and p-value by t-test on the average RMS tracking error of both NDPD and DPD subjects. For example, the difference of RMS tracking error of NDPD and DPD is not

TABLE V
P-VALUE AND K-MEANS.

Condition	# of Attributes	Distance	Accuracy	P-value
All	96	41.47	76.47%	
Slow	48	23.53	70.59%	0.041
Fast	48	35.14	82.35%	0.049
No-N	24	9.12	64.71%	0.429
With-N	72	41.38	82.35%	0.026
OFF-M	48	31.71	76.47%	0.028
ON-M	48	29.85	76.47%	0.132

significant (p-value > 0.05) in No-N and ON-M cases; thus, the DCC is relatively smaller than the other. On the other hand, the difference of RMS tracking error of NDPD and DPD is significant (p-value ≤ 0.5) in SS, FS, With-N and OFF-M cases. The DCC is relatively larger than the other, which leads to the higher clustering accuracy. Besides, we observed the higher clustering accuracy in FS and With-N cases compared to other cases. This result indicates that NDPD and DPD patients have different ability to respond to visual tracking cues. Note that there is an exception in all case that the DCC is larger than the other because the number of principal components (e.g., 95% of variance contribution) is bigger than that of the other cases.

B. Decision Tree Classification

Next, we applied the decision tree classification algorithm to classify NDP and DP based on subjects characteristics (Table II). Although previous studies demonstrated that there was no significant difference (p-value > 0.05) in each characteristic except converted daily L-dopa dosage (p-value= 0.047), we still jointly combined all characteristics, "Age, "Years Since Diagnosis", "UPDRS", "Converted Daily L-dopa Dosage", "L-dopa Equivalent Daily Dose", to indicate whether this combination could classify these two groups or not. Here, we did not use a clustering algorithm, because of different units, if we applied K-means, these attributes (characteristics) need to be normalized so that attributes would not have original meaning. To preserve original meaning of characteristics, decision tree is one of the most appropriate algorithms in healthcare and medicine fields.

As a supervised method, decision tree would label DPD and NDPD automatically, so we know which node is NDP or DP from the result shown in Fig. 5. The decision tree classify NDPD and DPD perfectly based on subjects characteristics. Then we used the leave-one-out cross validation, (i.e. iteratively using 18 subjects to build the model and 1 subject to test the model) the accuracy is 68.4%. Table VI demonstrates rules of the decision tree.

According to the structure of the decision tree (roughly indicates attributes ranking), the first level is "years since diagnosis", which indicates that in the model the years since diagnosis is the most crucial factor for classification of NDPD and DPD. The "L-dopa Equivalent Daily Dose" and "Age" are not in the model, which means they are not the main factors in classifying between NDPD and DPD.

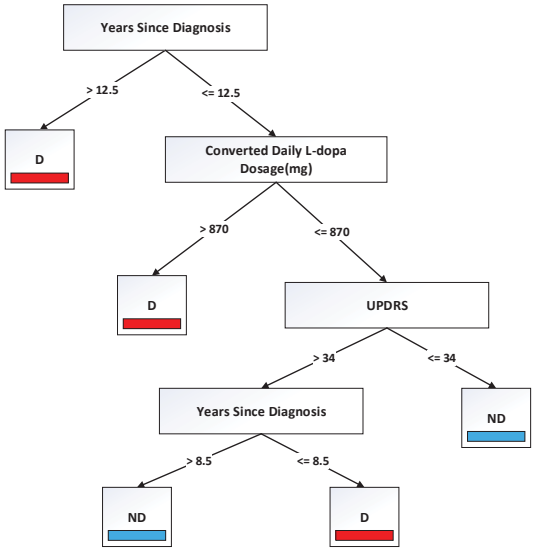

Fig. 5. Decision Tree. The split criterion is GainRATIO. "D" stands for DPD subjects, and "ND" stands for NDPD subjects. The result shows NDPD and DPD are perfectly classified by using three of five characteristics.

TABLE VI
THE DECISION TREE RULES.

Years since diagnosis > 12.500: D {ND=0, D=4}
Years since diagnosis ≤ 12.500
| Converted Daily L-dopa Dosage(mg) > 870: D {ND=0, D=3}
| Converted Daily L-dopa Dosage(mg) ≤ 870
| | UPDRS > 34
| | | Years since diagnosis > 8.500: ND {ND=4, D=0}
| | | Years since diagnosis ≤ 8.500: D {ND=0, D=2}
| | UPDRS ≤ 34: ND {ND=6, D=0}

IV. CONCLUSIONS

Our results indicate that data mining techniques have potential in healthcare and biomedicine. They allow us to analyze visually-guided tracking performance and identify which conditions are effective to cluster PD subjects into non-dyskinesia and dyskinesia subgroups, and which characteristics are primary and secondary to classify. We utilized two data mining techniques, K-means clustering algorithm and Decision tree classification algorithm. For K-means, we used RMS tracking error as attributes to cluster two groups in unsupervised fashion, in which we found that under fast speed and with noise conditions, it is effective to cluster these into two groups. Moreover, we studied how data mining and statistical analysis complement each other by looking at clustering result (e.g., DCC) and statistical significance (e.g., p-value of t-test) of the tracking performance dataset.In addition, For the decision tree, we used subjects characteristics to classify two groups with NDPD/DPD label, and the accuracy is 68.4% with leave-one-out cross validation which is lower than the accuracy of tracking performance clustering, 77.8%, so clinically tracking performance assessment is better than demographic characteristics while differentiating NDPD and DPD. Thus, we can use results of data mining to support

decision making on patients with NDPD or DPD. The small number of studied subjects is a limitation, and in future work we will extend the pool of researched subjects to achieve more general conclusions.

ACKNOWLEDGMENT

This work is a collaboration with the Joint NTU-UBC Research Centre of Excellence in Active Living for the Elderly (LILY).

REFERENCES

[1] Illhoi Yoo, Patricia Alafaireet, Miroslav Marinov, Keila Pena-Hernandez, Rajitha Gopidi, Jia-Fu Chang, and Lei Hua. Data mining in healthcare and biomedicine: a survey of the literature. *Journal of medical systems*, 36(4):2431–2448, 2012.

[2] Meeko MK Oishi, Ahmad Ashoori, and Martin J McKeown. Mode detection in switched pursuit tracking tasks: Hybrid estimation to measure performance in parkinson's disease. In *Decision and Control (CDC), 2010 49th IEEE Conference on*, pages 2124–2130. IEEE, 2010.

[3] Shyamal Patel, Konrad Lorincz, Richard Hughes, Nancy Huggins, John Growdon, David Standaert, Metin Akay, Jennifer Dy, Matt Welsh, and Paolo Bonato. Monitoring motor fluctuations in patients with parkinson's disease using wearable sensors. *Information Technology in Biomedicine, IEEE Transactions on*, 13(6):864–873, 2009.

[4] Roshan Cools, Roger A Barker, Barbara J Sahakian, and Trevor W Robbins. L-dopa medication remediates cognitive inflexibility, but increases impulsivity in patients with parkinson's disease. *Neuropsychologia*, 41(11):1431–1441, 2003.

[5] S Fahn, D Oakes, Ira Shoulson, K Kieburtz, A Rudolph, A Lang, CW Olanow, C Tanner, K Marek, et al. Levodopa and the progression of parkinson's disease. *The New England journal of medicine*, 351(24):2498, 2004.

[6] Geeta Yadav, Yugal Kumar, and GADADHAR Sahoo. Predication of parkinson's disease using data mining methods: A comparative analysis of tree, statistical and support vector machine classifiers. In *Computing and Communication Systems (NCCCS), 2012 National Conference on*, pages 1–8. IEEE, 2012.

[7] Paolo Bonato, Delsey M Sherrill, David G Standaert, Sara S Salles, and Metin Akay. Data mining techniques to detect motor fluctuations in parkinson's disease. In *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, volume 2, pages 4766–4769. IEEE, 2004.

[8] Wing-Lok Au, Ni Lei, Meeko MK Oishi, and Martin J McKeown. l-dopa induces under-damped visually guided motor responses in parkinsons disease. *Experimental brain research*, 202(3):553–559, 2010.

[9] James KR Stevenson, Pouria Talebifard, Edna Ty, Meeko MK Oishi, and Martin J McKeown. Dyskinetic parkinsons disease patients demonstrate motor abnormalities off medication. *Experimental brain research*, 214(3):471–479, 2011.

[10] Margaret M Hoehn and Melvin D Yahr. Parkinsonism: onset, progression, and mortality. *Neurology*, 50(2):318–318, 1998.

[11] Gang Li and Fei Liu. Application of a clustering method on sentiment analysis. *Journal of Information Science*, 38(2):127–139, 2012.

[12] John Ross Quinlan. *C4. 5: programs for machine learning*, volume 1. Morgan kaufmann, 1993.

[13] James KR Stevenson, Chonho Lee, Bu-Sung Lee, Pouria TalebiFard, Edna Ty, Kristina Aseeva, Meeko MK Oishi, and Martin J McKeown. Excessive sensitivity to uncertain visual input in l-dopa-induced dyskinesias in parkinsons disease: further implications for cerebellar involvement. *Frontiers in neurology*, 5, 2014.