Course Title: Biomedical Signal Processing
Course no – EEE 6205

A Review
On

High-Accuracy Detection of Early Parkinson's Disease through Multimodal Features and Machine Learning

Submitted to –

Dr. Sheikh Anwarul Haque Fattah

Submitted by-

Khan Shaikhul Hadi

ID: 0417062258

Md. Ajoad Hasan

ID: 0417062244

Abstract

In this paper, we deal with the analysis of a good set of motor and non-motor biomarkers for early prediction of PD. Then, we model this classifier using different machine learning algorithms. Finally, we comparing their performance.

Introduction

Parkinson's disease (PD) is a progressive neurodegenerative dis-order immensely affecting the quality of lives of millions of people worldwide. As of now, there is only symptomatic treatment available for PD and initiation of treatment at a later stage is of little help as the deterioration becomes extensive. Hence, early detection of PD is crucial for early management and for allowing neuro-protective strategies, to be administered earlier in the disease process, when available.

Pathophysiological Perspective

Parkinson's disease is a chronic, degenerative neurological disorder. The symptoms include motor symptoms like, resting tremor, bradykinesia, postural instability etc. and non-motor Symptoms, like cognitive impairment, sleep difficulties, loss of sense of smell etc. Some of the pathologies are-

- α-synuclein deposition in the brainstem.
- Ubiquinated protein deposits, termed as Lewy bodies and thread-like proteinaceous inclusions, termed as Lewy neurites in nerve cells.

The main pathophysiological cause lies in the loss of cells from various parts of the brain, especially Substantia Nigra. This region produces Dopamine, which acts as a chemical messenger for transmitting neuro-signals within the brain, thereby co-ordinating movements.

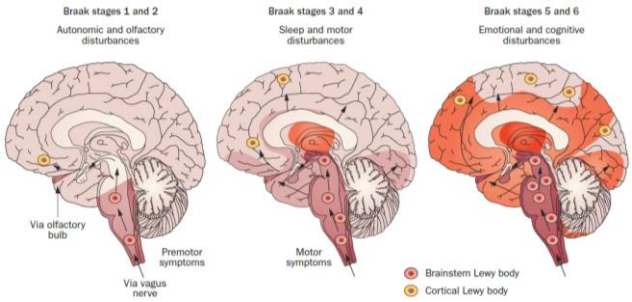

Figure: Different stages of progression of PD through brain [2]

Analysis Flowchart

Data Collection

- Parkinson's Progression Markers Initiative

Feature Extraction

- Non-motor features
- Cerebrospinal Fluid markers
- Dopaminergic Imaging Markers

Feature Statistical Analysis

- Box plot
- Wilcoxon Rank Sum test

Machine Learning Classifiers

- Naïve Bayes
- Support Vector Machine
- Boosted Trees
- Random Forest

Results Analysis

- Accuracy
- Sensitivity
- Specificity
- AUC

Data Collection

The data from Parkinson's Progression Markers Initiative (PPMI) database was obtained. (www.ppmi-info.org) [1]; the number of Normal Patients and Early Parkinson's patients is 183 and 401, respectively in this study. The data at baseline observation are considered among various visits, such as baseline, screening, 1st, 2nd visit etc. And 13 feature parameters were obtained in total for the provision in classifiers.

Feature Extraction

Neurodegenerative Biomarkers

1. Olfactory measurement: University of Pennsylvania Smell Identification Test (UPSIT)
2. REM sleep analysis: REM sleep Behaviour Disorder Screening Questionnaire (RBDSQ)
3. Cerebrospinal Fluid measurement (CSF)
4. Neuroimaging: DatScan SPECT imaging

As During the premotor phase in PD (5-20 years), the subject mostly shows non-motor symptoms such as REM sleep Behavior Disorder (RBD) and olfactory loss. But these does not carry sufficient sensitivity that can be used for screening. So, these are used in conjunction with other potential biomarkers such as CSF measurements and dopamine transporter imaging.

1. Olfactory Loss

Odor loss is an early marker for PD. The prime pathophysiology of olfactory loss is the deposition of α-synuclein in the olfactory bulb and the associated anterior olfactory nucleus (AON) at initial stages. Another cause can be due to the main projection neurons of the olfactory bulb are the mitral and tufted cells. The structures receiving these cells, such as the piriform and entorhinal cortex exhibit large numbers of Lewy bodies and Lewy neurites. Also the cortical nucleus of the amygdala, which receives the primary olfactory bulb projections, exhibits considerably more α- synuclein pathology and neuronal loss.

UPSIT

University of Pennsylvania Smell Identification Test (UPSIT) is developed for quantifying odor loss through a convenient screening [2].

The maximum total score of UPSIT is 40 points, out of 1 parameter. In this test, a subject is provided with 4 different 10 page booklets, each containing different odor. For each of this pages, there exists a question with 4 options. Based on the right/wrong answers, the UPSIT score is provided.

2. REM Sleep Disorder

REM Sleep Disorder is the elaborated motor activity that accompanies REM sleep with dream mentation. This is a strong predictor of neurodegenerative diseases, including PD.

The pathophysiology of REM sleep behavior disorder (RBD) is due to the α-synuclein-mediated degeneration of sleep-regulating nuclei and dysfunction of brainstem structures, especially in the pontine tegmentum that modulates REM sleep.

RBDSQ

A short RBD screening questionnaire (RBDSQ) is developed in order to quantify sleep disorder through an easily applicable diagnostic screening tool [3].

The maximum total score of the RBDSQ is 13 points, out of 10 questions, treated as 1 parameter.

- Questions 1 to 4 address the frequency and content of dreams and their relationship to nocturnal movements and behavior.
- Question 5 asks about self-injuries and injuries of the bed partner.
- Question 6 comprises 4 sub-items assessing nocturnal motor behavior sp.

- Questions 7 and 8 deal with nocturnal awakenings.
- Question 9 focuses on disturbed sleep in general.
- Question 10 on the presence of any neurological disorder.

3. Cerebrospinal Fluid

Cerebrospinal Fluid (CSF) is a colorless body fluid found in brain and spinal cord, which has more physical contact with the brain as compared to any other fluid [4]. Due to the close proximity with the brain, any protein or peptide which is related to the brain specific disease (i.e. PD) are diffused into CSF.

xMAP/Luminex Multiplex Immunoassay
CSF biomarkers are extracted using xMAP/Luminex multiplex immunoassay. Values of 4 biomarkers are directly extracted from the CSF fluid [5]-

- α-synuclein
- Aβ1-42 (amyloid beta (1-42))
- T-tau (total tau)
- P-tau$_{181}$ (tau phosphorylated at threonine)

T-tau and α-synuclein are associated with severity of motor dysfunction in early PD. Apart from these, 3 other parameters [6] (In total, 7 parameters) are provided-

- T-tau/ Aβ1-42
- P-tau$_{181}$/ Aβ1-42
- P-tau$_{181}$/ T-tau

4. Neuroimaging

Neuroimaging of the dopaminergic cells can act as a biomarker for early diagnosis of PD [7]. Dopaminergic imaging identifies presynaptic dopaminergic deficits and reduction in binding in the caudate and putamen. These two regions long been associated with motor processes due to its role in Parkinson's disease.

DatScan SPECT
DATSCAN SPECT (Single-photon emission computed tomography) imaging is a common neuroimaging technique in order to quantify the dopaminergic transporter loss through the value of striatal binding ratio in the substantia nigrata.

Striatal binding ratio (SBR) is a measure of count density of these regions, which can be quantified as-

SBR = (striatal region/reference region) −1

The SBR value are extracted from the SPECT images according to following steps. Firstly, Raw SPECT images are collected and reconstructed. Next, Attenuation correction. Then, Gaussian filter are applied and normalized. Last of all, values of 4 biomarkers (4 parameters) are extracted and calculated.

- right caudate SBR
- left caudate SBR
- right putamen SBR
- left putamen SBR

Machine Learning Classifiers

- **Linear Classifier**

Naïve Bayes requires small amount of training data to estimate the necessary parameters and therefore, can be extremely fast [8].

Logistic Regression provides probabilistic outputs [8].

- **Non-linear Classifier**

SVM is fast compared to others but requires more parameters to tune and required normalized data.

Random Forests & Boosted trees gives higher accuracy in most cases, requires no normalized data but slower than trivial methods [9].

Our Work

We collected data with 147 Normal patient and 388 Early PD patient from updated Parkinson's Progression Markers Initiative (PPMI) database. Then we analyzed the boxplots of the feature; statistically analyzed all 13 features mentioned in the work to evaluate their claims. Hence we tried to recreated result of SVM & Boosted Tree algorithm. Last of all, we tried to analyze the reason behind the mismatch and implement Logistic Regression considering them.

Statistical Analysis of Feature

In order to visualize the spread and distribution between normal and early PD groups, Box-plot is provided for each feature. Statistically these plots are analyzed for significance using Wilcoxon rank sum test. Only statistically significant features (p-value < 0.05) were used.

Feat.	Normal	Early PD	Z-stat.	*P*-value	Expected *p*-value
RBDSQ	2.81±2.325	3.2279±2.66	-1.2483	0.21193	0.016
UPSIT	34.22±4.61	22.353±8.29	14.502	≈0	≈0
SBR-RC	2.92±0.59	1.9853±0.59	13.526	≈0	≈0
SBR-LC	2.98±0.61	1.9871±0.58	13.967	≈0	≈0
SBR-RP	2.13±0.56	0.84441±0.36	17.141	≈0	≈0
SBR-LP	2.11±0.54	0.80933±0.35	17.357	≈0	≈0
Aβ1-42	379.15±113.42	372.32±100.32	0.99827	0.31815	0.501
p-Tau	18.871±12.12	15.76±10.17	3.6239	0.00029	0.002
T-tau	52.101±24.64	44.75±18.2	3.0995	0.001938	0.002
T-tau/ Aβ1-42	0.16317±0.19	0.12597±0.06	2.0481	0.040548	0.023
P-tau/ Aβ1-42	0.057898±0.07	0.043667±0.03	2.8142	0.00489	0.02
P-tau/T-tau	0.38209±0.2	0.3716±0.23	1.2351	0.21681	0.737
α-syn	190.73±57.29	186.75±50.1	1.1068	0.26836	≈0

Machine Learning Algorithms

We used Classifier Learner Toolbox, MATLAB 2017a, where we have provided clinical data of 147 Normal and 388 Early PD patients.

Algorithm		Result		Expected Result	
		Accuracy (%)	AUC (%)	Accuracy (%)	AUC (%)
SVM	Train	96.2±0.57	98.3±0.48	97.14±0.4	99.27±0.1
	Test	96.35±1	95.76±1.2	96.4±1.08	98.88±0.6
Boosted Tree	Train	96.5±0.77	97.2±0.42	100±0	100±0
	Test	95.4±0.42	94.11±1.8	95.08±1.2	98.23±0.8
Logistic Regression	Train	96.29±0.4	95.26±0.5	96.5±0.6	99.20±0.2
	Test	95.8±1.06	94.95±1.6	95.63±1.2	98.66±0.7

Mismatch and Limitations

The database are continuously updated, therefore lesser amount of data than the ones presented in the paper. Also there was lesser degree of freedom to tune parameters in Classifier Learner Toolbox.

The analysis involved early stage PD patients and age-matched healthy normal, and not any premotor PD subjects (subjects who is at risk of PD but not diagnosed as PD due to absence of classic motor symptoms).

Conclusion

This work is an extension of the work by Kang et al. where they used CSF measurements and observed a low diagnostic utility. Here, relevant and significant pre-clinical features corresponding to non-motor and imaging markers are used and a superior accuracy is obtained. To mention further, SVM produced a near perfect classification.

References

[1] K. Marek, D. Jennings, S. Lasch, A. Siderowf, C. Tanner, T. Simuni, C. Coffey, K. Kieburtz, E. Flagg, S. Chowdhury, W. Poewe, B. Mollenhauer, P.-E. Klinik, T. Sherer, M. Frasier, C. Meunier, A. Rudolph, C. Casaceli, J. Seibyl, S. Mendick, N. Schuff, Y. Zhang, A. Toga, K. Crawford, A. Ansbach, P. De Blasio, M. Piovella, J. Trojanowski, L. Shaw, A. Singleton, K. Hawkins, J. Eberling, D. Brooks, D. Russell, L. Leary, S. Factor, B. Sommerfeld, P. Hogarth, E. Pighetti, K. Williams, D. Standaert, S. Guthrie, R. Hauser, H. Delgado, J. Jankovic, C. Hunter, M. Stern, B. Tran, J. Leverenz, M. Baca, S. Frank, C.-A. Thomas, I. Richard, C. Deeley, L. Rees, F. Sprenger, E. Lang, H. Shill, S. Obradov, H. Fernandez, A. Winters, D. Berg, K. Gauss, D. Galasko, D. Fontaine, Z. Mari, M. Gerstenhaber, D. Brooks, S. Malloy, P. Barone, K. Longo, T. Comery, B. Ravina, I. Grachev, K. Gallagher, M. Collins, K.L. Widnell, S. Ostrowizki, P. Fontoura, T. Ho, J. Luthman, M.v.d. Brug, A.D. Reith, P. Taylor, The Parkinson Progression Marker Initiative (PPMI), Prog Neurobiol, 95 (2011) 629-635.

[2] R.L. Doty, P. Shaman, M. Dann, Development of the University of Pennsylvania Smell Identification Test: a standardized microencapsulated test of olfactory function, Physiol Behav, 32 (1984) 489-502

[3] The REM Sleep Behaviour Disorder Screening Questionnaire A new Diagonostic Instrument, Karin Stiasny-Kolster, Geert Mayer, Hephata Klinik Schwalmstadt-Treysa, Germany,Sylvia Schfer, Wolfgang H Oertel

[4] Cerebrospinal fluid biomarkers in parkinsonian conditions: an update and future directions Nadia Magdalinou, Andrew J Lees, Henrik Zetterberg.

[5] A 1-42, t-tau and p-tau181 measurements in 120 PPMI CSF samples using xMAP/Luminex multiplex immunoassay, Ju Hee Kang, John Q Trojanowski and Leslie M Shaw, Department of Pathology & Laboratory Medicine and Institute onAging, Center for Neurodegenerative Disease Research, Perelman School of Medicine, University of Pennsylvania

[6] J.H. Kang, D.J. Irwin, A.S. Chen-Plotkin, A. Siderowf, C. Caspell, C.S. Coffey, T. Waligrska, P. Taylor, S. Pan, M. Frasier, K. Marek, K. Kieburtz, D. Jennings, T. Simuni, C.M. Tanner, A. Singleton, A.W. Toga, S. Chowdhury, B. Mollenhauer, J.Q. Trojanowski, L.M. Shaw, Parkinson's Progression Markers Initiative, Association of cerebrospinal fluid -amyloid 1-42, T-tau, P-tau181, and -synuclein levels with clinical features of drugnaive patients with early Parkinson disease, JAMA Neurol, 70 (2013) 1277-1287.

[7] J.L. Cummings, C. Henchcliffe, S. Schaier, T. Simuni, A. Waxman, P. Kemp, The role of dopaminergic imaging in patients with symptoms of dopaminergic system neurodegeneration, Brain, 134 (2011) 3146-3166.

[8] A.Y. Ng, M.I. Jordan, On Discriminative vs. Generative Classifiers: Acomparison of Logistic Regression and Naive Bayes, in: T.G. Dietterich, S.Becker, Z. Ghahramani (Eds.), Advances in Neural Information ProcessingSystems (NIPS), MIT Press, 2002, pp. 841–848.

[9] Y. Freund, R.E. Schapire, A Short Introduction to Boosting, Journal of JapaneseSociety for Artificial Intelligence 14 (1999) 771–780.