ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2018) 000–000

Procedia
Computer Science

www.elsevier.com/locate/procedia

CENTERIS - International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies, CENTERIS/ProjMAN/HCist 2018

Machine Learning Application to Quantify the Tremor Level for Parkinson's Disease Patients

Taígo Ítalo Pedrosa[a], Felipe F. Vasconcelos[a], Leonardo Medeiros[a] , Leandro Dias Silva[b]

[a]*Federal Institute of Alagoas, St. Ferroviário - 530, Maceió 57020-600, Brazil*
[b]*Federal University of Alagoas, St. Lourival Melo Mota, Maceió 57072-900, Brazil*

Abstract

Parkinson's disease (PD) is a neurodegenerative disorder with progressive nature. It causes motor symptoms such as resting tremor, bradykinesia and others movement disorders. Because of its progressive nature, this disease needs a continuous monitoring of motor symptoms. Health Monitoring Systems are widely used to monitor the disease progress, improving the treatment and minimizing the drug side effects. In this research, we developed two predictive models using a supervised machine learning approach. These models can classify the Parkinson disease’s rest tremor between high or low frequencies, showing the intensity of Parkinson’s motor symptom. This classification allows the detailed monitoring of the medication's effectiveness and the disease progress. In our validation, we applied leave-one-out cross-validation methods to classify the level of the PD tremor. In our results, we reached a classification accuracy of 92.8%. Therefore, this work proposes a new approach to classifying PD tremor and improving the the patients quality´s live on treatment, using non-invasive health monitoring systems improved by machine learning classification algorithms.

Keywords: Machine learning; Parkinson disease; resting tremor; KNN classifier; Linear SVM classifier

---

1. Introduction

Parkinson's disease is a progressive, chronic, neurodegenerative disease that is mainly related to the degeneration of dopaminergic neurons at the substantia nigra of the ventral midbrain [1] and can happen in other regions of the brain [2]. It has a higher incidence in people over 50 years and can occur around the age of 30-40 years, when it is named as juvenile Parkinson. The principal Parkinson's motor symptoms are rest tremors, muscular rigidity, bradykinesia and postural reflex abnormalities. The combination of at least two of the related symptoms cited is enough to serve as the disease diagnostic.

The disease usually reaches its peak in about 10 years and can bring a major social, financial and personal impact to the patient’s life, since the patients need people to take care of them and the medicines are expensive. The annual cost with Parkinson’s treatment reach in average US$ 6,000 per person only in São Paulo, Brazil [3]

Between 1% and 3% of the entire world population suffer from PD [3], and this number continues to increase as the average age of the world's population is increasing progressively due to improved life expectancy, according to studies by the World Health Organization (WHO) [4] very soon we will have more seniors than children. For this reason, considering that the elderly population has a higher incidence of chronic diseases [5], it is necessary to improve the monitoring of the health status of this population. Therefore, because of the increase in the number of chronic patients and the imminent reduction in the number of hospital beds available for treatment and monitoring of chronic diseases, in addition to the insufficiency of specialized professionals to the demand, it is necessary to transpose the monitoring services from hospital beds to their own homes [6].

Health Monitoring Systems (HMS) are applied to the monitoring of diseases and to get fully functional the patient need to add them to their routines [7]. However, many people complain that the HMS are invasive, diminishing their privacy [7]. For this reason, in recently years, the creation of computational technologies for health monitoring systems in a non-invasive approach has been a relevant and recurrent theme in science computing research [8], [9]. In relation to Parkinson's disease, it is known that the efficacy of drug dosage declines over time and, because of the progressive aspect of the disease, the symptoms tend to increase [10]. Therefore, the HMS can help medics to make decisions because of its ability to increase clinical understanding of the development of the disease.

There is not a specified clinical exam for PD; the diagnosis is based on the development of the symptoms, the medical history and some clinical exams to eliminate the possibility of other diseases. So, in this research, we aimed to develop classifier model based to quantify the level of PD rest tremor. Then, this classifier was able to classify the patient between two groups: High amplitude tremor; Low amplitude tremor. It is also able to show in numbers how the disease is developing throughout time.

So, in our research, we intent to apply this classifier in a embedded non-invasive HMS. Then , this system can be used by a doctor to monitor the progress of the disease and to identify when the medication start to become less efficient. In the end of the research, we will provide a non-invasive HMS integrated into to user´s daily routines to monitor Parkinson’s disease tremor motor symptom.

2. Related works

Several non-invasive systems have been and are being developed to increase the range of information provided to medics and to allow comfortable insertion of HMS into the patient's routine. In addition, the dissemination of these systems aligned with the use of computational solutions can provide a wide range of information, increasing clinical understanding of the disease progression [8].

Most of them are focus on bradykinesia or other Parkinson´s symptoms. In fact, they show that monitoring of the symptoms is something to be study. Stuebner *et. al.* [9] especially focus on blood pressure and heart rate, which are related to PD because advanced stages patients can have an atypical blood pressure, which can tell if they have a new symptoms or some other problem. We also tried methods to help developing a new diagnosis, but we used K-NN (K-Nearest Neighbor’s) classifier [11], which is a nearest neighbor’s supervised classifier, and others machine learnings scripts. Ramani *et. al.* [12] shows an approach related to classifier models for PD, but they try to find the most accurate classifier to a vocal range dataset. Our works differs by using rest tremor as main symptoms, also by using a cheaper hardware: a smartphone and its accelerometer. Due to using only certain moments of the day for the recording, we are able to present a greater precision, because we do not have to deal with non-pathological tremors.

3. Procedures

The PD tremor classification application was developed to help neurologists to have more information about patients. From the data entry of the tremors, it processes and classifies between high and low amplitude. The data processing takes place through the normalization of the data and the feature selection.

3.1. Instruments

The application is developed in *Python* programming language where data processing techniques and machine learning algorithms were applied. A low-intensity laser that was directed to a reflective piece of paper in the subject's finger for 60 seconds acquired the data. The database contains several cases of record such as when the patients receive Deep Brain Stimulation (DBS), when they are using medication, when the DBS was off and some others cases [13]. However, in the Machine Learning process we used the purest case of PD without DBS and medication. To be consider without medication, the patients did not take any medication in a period of 12 hours.

Deep Brain Stimulation is a neurosurgical procedure that involves the implementation of a device which sends electrical impulses to a certain part of the brain. DBS is able to reduce some of the symptoms, for this reason we used a case without DBS.

3.2. Data normalization

After the data was collected from *PhysioBank* Database *https://www.physionet.org/physiobank/database/tremordb/*, an analysis and the respective processing were performed. First, it was analyzed that the first second recorded of all subjects was discrepant, the tremor would go to very high levels, causing the graph and algorithm to have errors. Because of this, the first second record was removed in all cases. It was also noticed that some subjects presented their data in meters per second, while others presented in millimeters per second. After becoming aware of the problem, a simple algorithm was made to normalize all data in millimeters per second.

3.3. Methods

After the data were normalized, it was applied the Fast Fourier Transform (FFT) to extract the power spectrum (PWS). Discrete Fourier Transform (DFT) converts a discrete time sequence to the discrete frequency domain. The DFT itself brings a great computational cost, to solve this problem the FFT was developed, which is, basically, the fast form of the DFT.

FFT decomposes an N point time domain signal into N time domain signals each composed of a single point. After this, it is necessary calculate the N frequency spectra corresponding to these N time domain signals. Finally, the N spectra are synthesized into a single frequency spectrum. The second stage decomposes the data into four signals of 4 points. This pattern continues until there are N signals composed of a single point. An interlaced decomposition is use each time a signal is broken in two, that is, the signal is separated into its even and odd numbered samples. The FFT time domain decomposition is carry out by a bit reversal-sorting algorithm. The last step in the FFT is to combine the N frequency spectra in the exact reverse order that the time domain decomposition took place. This is where the decomposition is done in reverse, from one point to 4 points [20]. Power Spectrum (PWS) is a positive real function of a variable frequency associated with a stochastic process, or a deterministic function of time, which has energy or force dimension per Hertz.

Finally, after data processing, we applied them on a supervised machine learning approach to identify patterns on the data to distinguish between the groups HAT (High Amplitude Tremor) and LAT (Low Amplitude Tremor) and to be able to classify the test data.

4. Experiments and discussion

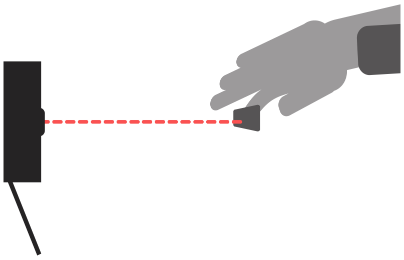

Figure. 1. Illustration of the recording process.

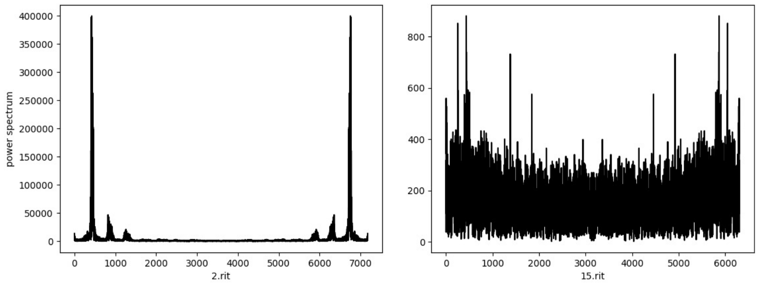

Figure 2. Power Spectrum graphics of individuals 2 (HAT) and 15 (LAT)

4.1. Data collection

This work uses data collected from the research of Beuter *et. al.* [13]. A total of 16 individuals were recorded to the database, all diagnosed with PD; the subjects were classified into two groups, HAT and LAT, each of the groups were composed of 8 individuals, belonging to the HAT group individuals 1 to 8 and the LAT group individuals from 9 to 16.

The database has 8 records; These records also cover other variables, such as presence of Deep Brain Stimulation and the presence of medication. In the record used, the subjects were not exposed to any type of medication or another stimulus, the record was chosen for showing the tremor without any external factor that can mask its real nature. In the course of the research, not all subjects were recorded for every possible situation. The chosen recording session was composed of 14 files; each file containing the tremor readings has approximately 1 minute of recording. All readings were processed to millimeter per second. Individuals used similar equipment during recordings. In this record, the high amplitude tremor group was composed of 6 individuals and the low amplitude tremor group by 8. The data was recorded by performing the following steps and is illustrated in the Figure 1:

- The subject stands at a determined distance from the low-intensity velocity-transducer laser
- The laser was directed to a piece of reflective paper on the subject’s index finger tip.
- During one minute, the velocity-transducer laser captures raw values outputting voltage proportional to the velocity of the finger.
- After one minute the recording process stops.

4.2. Machine learning feature selection

During the data processing some features that clearly differentiate the individuals belonging to group 1 (HAT, High Amplitude Tremor) from the individuals belonging to group 2 (LAT, Low Amplitude Tremor) were chosen, they are: Signal peak; Average power spectral density; Standard deviation of the power spectrum.

The Figure 2 shows the differences between individuals belonging to the HAT and LAT groups, comparing their power spectrum graphics. The individual "2.rit" belongs to HAT group and the individual "15.rit" belongs to LAT group.

With the purpose to validate our work, we applied LOOCV (Leave-one-out cross validation) [11] to evaluate the accuracy of the predictive models. The LOOCV is a type of cross-validation statistic method that separates one instance of the dataset for being the test data and use the rest of them as training data.

In this work, we use two data groups, the training group and the test group; in the training group were provided the determinant characteristics (processed data) and the class corresponding to each individual; in the test group, only the characteristics were provided, the classification of the group that the individual belongs to was made by the algorithm. We used two machine-learning algorithms: K-NN Classifier [11]; Linear Support Vector Machine (SMV) Classifier [11].

4.3. K-Nearest Neighbor's (KNN) classifier

The K-NN is a classifier algorithm by supervised machine learning based on the K nearest neighbor, as show in the Figure 3 [11]; in this application, its behavior is considered as the determinant characteristics mentioned above, connecting the test subject to the five nearest individuals. The value of K was determined based on the accuracy returned by the respective test; the five nearest neighbors were considered in the classification.

This algorithm was able to predict to which group the individual belongs to with a 92.8% of accuracy.

4.4. Linear Support Vector Machine (SVM) classifier

The Linear SVM is also a classifier algorithm by supervised machine learning that proposes to find the dividing line that determines which class an individual belongs [11]. In this application, its behavior considered the determinant characteristics mentioned above, providing to the algorithm the necessary information to draw the line and through RMSE (Root Mean Square Error) adjust the line to best classify the individuals.

As K-NN, SVM was able to distinguish to which group the individual belongs to 92.8% of accuracy.

4.5. Classification metrics

The classification performance of the model is best presented in a two-class confusion matrix [14] consisting of a 2x2 matrix, with True Positives (TP), False Negatives (FN), False Positives (FP) and True Negatives (TN) described in Table 1 and their metrics presented in Table 4.

Table 1. Confusion Matrix Of K-NN Classification and Linear SVC With One Leave Out Cross Validation.

		Predicted class	
		HAT	LAT
Actual	HAT	5	1
class	LAT	0	8

As the two algorithms tied in the metrics, the mean time in 10 implementations was measured, the K-NN averaged 0.135s and the SVC averaged 0.131s. The measured times are very close, so it was considered another tie.

Table 2. Results of classification

Subjects	Class predicted by K-NN	Class predicted by SVC
HAT 2	HAT	HAT
HAT 4	HAT	HAT
HAT 5	LAT	HAT
HAT 6	HAT	HAT
HAT 7	HAT	HAT
HAT 8	HAT	LAT
LAT 1	LAT	LAT
LAT 2	LAT	LAT
LAT 3	LAT	LAT
LAT 4	LAT	LAT
LAT 5	LAT	LAT
LAT 6	LAT	LAT
LAT 7	LAT	LAT
LAT 8	LAT	LAT

Table 3. Results of processed data

Subjects	Signal peak	Mean of Power Spectral Density	Absolute deviation of PWS
HAT 2	399757	5883	3430
HAT 4	201822	3861	2202
HAT 5	81773	1100	653
HAT 6	379978	7985	4498
HAT 7	129416	1381	765
HAT 8	196937	4320	2395
LAT 1	6678	27	147
LAT 2	3153	273	156
LAT 3	2923	351	182
LAT 4	851	211	117
LAT 5	2222	336	111
LAT 6	1155	234	136
LAT 7	880	160	90
LAT 8	6205	244	144

Table 4. Performance of PD and Control group classification on both algorithms

Classifier metrics	
Accuracy	92,85%
Precision	100%
TP rate	83.33%

5. Conclusion

In this research, a new approach for the classification of Parkinson's disease was developed, according to the intensity of the tremor. The processing of the information that constitutes the basis of tests of this work allowed the learning algorithms to classify what makes an individual belong to a certain group in the great majority of the tests. The predictive models created in this paper are a basis for the development of noninvasive monitors of the advancement of Parkinson's disease. The database for the predictive models consisted of recordings of the tremors of patients diagnosed, without medication, composing the tremor groups of high and low amplitude. The data recorded were readings of the tremor velocity at rest, the database was used to verify the efficiency of the predictive models and thus prove its validity. We applied a K-NN Classifier and Linear SVM Classifier to identify whether the subject has high amplitude tremor or low amplitude tremor, reaching with both algorithms an accuracy of 92.8%.

References

[1] L. C. Triarhou. “Dopamine and parkinson’s disease" *Madame Curie Bioscience Database*: 2000-2013.

[2] M. da Saude. (2010) “Protocolo clinico e diretrizes terapeuticas - doença de parkinson - portaria sas/ms n 228” [Clinical protocol and therapeutic guidelines- sas/ms n 228] *Sistema unico de Saude, Tech. Rep.* [in Portuguese]

[3] Bovolenta, Tânia M., et al. "Average annual cost of Parkinson’s disease in São Paulo, Brazil, with a focus on disease-related motor symptoms." Clinical interventions in aging **12** (2017): 2095.

[4] W. H. Organization. (2011) “Global health and ageing,” *US National Institute of Aging, Tech. Rep.*

[5] C. Leite-Cavalcanti, M. da Conceição Rodrigues-Gonçalves, L. S. Rios-Asciutti and A. Leite-Cavalcanti. (2009) “Prevalencia de doenças cronicas e estado nutricional em um grupo de idosos brasileiros” [Prevalence of chronic disease and nutritional status in a group of elderly Brazilians] *Revista de Salud Publica* **11**. [in Portuguese]

[6] E. Thum, L. A. Facchini, G. Wyshak and P. Campbell. (2011) “The utilization of home care by the elderly in brazil’s primary health care system” American *Journal Of Public Health* **101**.

[7] H. Alemdar, C. Tunca and C. Ersoy. (2015) “Daily life behaviour monitoring for health assessment using machine learning: Bridging the gap between domains” *Personal Ubiquitous Computing* **19**.

[8] S. Smith, M. Lones, M. Bedder, J. Alty, J. Cosgrove, R. Maguire, M. Pownall, D. Ivanoiu, C. Lyle, A. Cording and C. J. H. Elliott. (2015) “Computational approaches for understanding the diagnosis and treatment of parkinson’s disease” *IET Systems Biology* **9** (**6**): 226-233.

[9] E. Stuebner, E. Vichayanrat, D. A. Low, C. J. Mathias, S. Isenmann and C.-A. Haensch (2013) “Twenty-four hour noninvasive ambulatory blood pressure and heart rate monitoring in Parkinson’s disease” *Frontiers in neurology* **4**.

[10] N. I. for Health and C. E. G. Britain. (2006) "Parkinson’s Disease: Diagnosis and Management in Primary and Secondary Care, ser. NICE clinical guideline." *National Institute for Health and Clinical Excellence.*

[11] Jake VanderPlas. (2016) "Python Data Science Handbook: Essential Tools for Working with Data" *O'Reilly Media, Inc.*

[12] R. G. Ramani and G. Sivagami. (2011) “Parkinson disease classification using data mining algorithms,” *International journal of computer applications* **32** (**9**): 17-22.

[13] A. Beuter, M. S. Titcombe, F. Richer, C. Gross and D. Guehl. (2001) “Effect of deep brain stimulation on amplitude and frequency characteristics of rest tremor in parkinson’s disease,” *Thalamus and Related Systems* **1** (**3**): 203-211.

[14] M. Kantardzic. (2011) "Data Mining: Concepts, Models, Methods, and Algorithms, 2nd ed." *John Wiley & Sons.*