Selection of voice parameters for Parkinson´s disease prediction from collected mobile data

Monica Giuliano
Engineering Departament
Universidad Nacional de La Matanza
San Justo, Argentina
mgiuliano@unlam.edu.ar

Alfonsa García-López
Escuela Técnica Superior de Ingeniería
Universidad Politécnica de Madrid
Madrid, Spain
alfonsa.garcia@etsisi.upm.es

Silvia Pérez
Engineering Departament
Universidad Nacional de La Matanza
San Justo, Argentina
sperez@unlam.edu.ar

Fransisco Díaz Pérez
Escuela Técnica Superior de Ingeniería
Universidad Politécnica de Madrid
Madrid, Spain
fdiaz@etsisi.upm.es

Osvaldo Spositto
Engineering Departament
Universidad Nacional de La Matanza
San Justo, Argentina
spositto@unlam.edu.ar

Julio Bossero
Engineering Department
Universidad Nacional de La Matanza
San Justo, Argentina
jbossero@unlam.edu.ar

***Abstract*— Voice disorders, which can help in the diagnosis of Parkinson's disease (PD), can be measured with acoustic tools. In this work, demographic data and vocal phonation records /a/ from the available mPower database were analyzed to identify PD patients. A parsimonious model was then found that achieved a reduction from 62 to 5 phonation characteristics, which were considered in addition to gender and age. Multilayer Perceptron (MLP) and Logistic Regression (LR) neural networks were used to obtain a model with high prediction capacity (area under receiver operating characteristic curve, AUC-ROC, over 0.82). This work contributes to the monitoring of EP patients from the recording of a few phonation features collected by means of a mobile phone.**

Keywords—Parkinson disease, voice parameters, mobile data, classification models.

I. INTRODUCTION

Parkinson's Disease (PD), described in 1817 as “shaking palsy" by Doctor James Parkinson, is a neurodegenerative disorder of the central nervous system that affects the control mechanism of the human body [1]. In addition to the fact that traditional clinical assessment for monitoring Parkinson´s progression is expensive, PD patients may have trouble attending medical visits for control. Thus, there is a need for an inexpensive and objective clinical tool to track PD symptom progression. The growing developments in information technology and telecommunication offer a good opportunity for remote monitoring.

Voice disorders can be measured by acoustic tools and there is a great deal of research aiming at the automatic evaluation of pathological voice patterns [2, 3] and their relationship with PD [4, 5, 6, 7].

In this context, it has been usual to collect voice recordings and extract different features. The following features of speech signals have been measured in order to determine their influence on PD: Fundamental Frequency or Pitch [4, 8, 5, 9]; Shimmer, Jitter, HNR [4, 5, 8, 9]; Mel Frequency Cepstral Coefficients (MFCC) [10, 7] and non linear dysphonia measures such as: Pitch Period Entropy (PPE), Detrended Fluctuation Analysis (DFA), Recurrence Period Density Entropy (RPDE) [4, 9, 11]. Some features are highly correlated with each other, which suggested the use of features selection methods [9, 11, 12].

The current study utilized an open-access dataset named “mPower” available at https://www.synapse.org/. Synapse is a platform that allows scientific content sharing (data, code, results). The Synapse mPower (*Mobil Parkinson Disease Study*) is an observational smartphone-based ongoing study that pilots new approaches to monitoring key indicators of PD progression and diagnosis. It supplements traditional behavioral symptom measurements with novel metrics from sensor-rich mobile devices. As a scalable, inexpensive, and non-invasive method for frequent measurement and tracking of symptoms, the Parkinson´s mPower app has been able to survey a large, longitudinal cohort of volunteers with PD and healthy controls, who provide their voice recordings and demographic data, and do several memory, tapping and walking activities [12]

Demographic data and recordings of the /a/ vowel phonation were analyzed from the mPower database. We sought to identify the diagnosis of Parkinson's disease (PD) and non Parkinson's disease (non-PD) from the acoustic analysis of phonation. We searched for a parsimonious model with reduction of the number of parameters collected for classification of people with and without PD diagnosis.

II. METHODOLOGY

A. The data

Coded study data, consisting of survey responses and mobile sensor measurements, were stored in Synapse for controlled distribution to researchers. From the data base available in the mPower-Synapse platform, 62 voice parameters were extracted using OpenSmile software. These features were attached to other data: age, gender, PD diagnosis (true/false), years since diagnosis, years of medication. From phonation recordings, the audios and the medication time point (med time-point) at the time of the recording were gathered.

A total of 2253 people over 35 years of age were identified. For each, a recording of the phonation of the vowel /a/ was selected. PD diagnosis (true/false) was regarded as a classification variable. Afterwards, cases with inconsistencies in diagnosis, year since diagnosis, start time of medication and medication time point were eliminated resulting in 2222 cases, 933 PD and 1289 non-PD.

B. Statistical methods

Parametrization of each recording was performed in 62 linear features using OpenSmile software. Statistical analyses were performed to identify the relevant variables that predict

978-1-7281-1491-0/19/$31.00 ©2019 IEEE

PD diagnosis through these parameters and some demographic variables. For such purpose, the following techniques were used: Principal Component Analysis (PCA), Analysis of Variance (ANOVA), neural networks Multilayer Perceptron (MLP) and Logistic Regression (LR). To check the performance of the classification model, the Area Under the Receiver Operating Characteristic (AUC-ROC) Curve was used. Such method measures discrimination, that is, the ability of the test to correctly classify those with and without the disease. In addition, accuracy was calculated for each model, with a cut value 0.5, to give the percentage of well-predicted cases.

III. RESULTS

The audios from the mPower database were parametrized with OpenSmile, derived into 62 specific variables of voice parameters which were then joined to some demographic variables for subsequent analysis. The following demographic data were used: age, sex, professional diagnosis, years since diagnosis and medication start year. From the phonation recordings, the audios and the medication time point at the time of the recording were analyzed.

Initially, there were 51420 different voice recordings. We identified 2253 people with different health codes and 35 years of age or older. For each person, a recording of the /a/ vowel phonation was selected.

Professional diagnosis (true-false) of Parkinson's was taken as a class variable. Then, the inconsistent cases between diagnosis, year since diagnosis (diagnosis-year), medication start year and medication time point (med time-point) at the time of the recording were eliminated from the base. Finally, a total of 2222 cases remained, 933 PD and 1289 non-PD.

A. Reduction of variables

We performed cyclical analysis of variable reduction from the initial 62 parameters of the recorded phoneme.

At first, Principal Components Analysis (PCA) was performed with 76.1% of the variance and we went from 62 to 33 components, considering 18 factors with eigenvalues greater than 1, the unexplained variables and the most used in the bibliography (1 or 2 variables per factor plus those not explained by the model).

PCA was then carried out with 69.5% of the variance with 9 factors considering the greatest variance explained and the significant contribution of the eigenvalues with respect to the previous ones (disregarding low contributions of the last two factors with eigenvalue greater than 1). Nine variables remained, one for each factor, the one with the highest factorial load or, in case of having to decide between mean and deviation of a variable, the mean was chosen.

K-means clustering was applied on 9-factor PCA and 6 groups remained, one of them with only 4 cases that were eliminated, leaving a total of 2218 cases and 5 groups.

ANOVA was then applied to the 9 diagnostic variables in the 5 groups and variables without differences in mean values were eliminated, leaving 5 variables, namely:

- d1: Mean of logarithmic F0 on a semitone frequency scale.
- d2: Mean of the ratio of the energy of the spectral harmonic peak at the first formant's center frequency to the energy of the spectral peak at F0 in voiced regions
- d3: Coefficient of variation of the ratio of the energy of the spectral harmonic peak at the first formant's center frequency to the energy of the spectral peak at F0 in voiced regions
- d4: Mean of linear regression slope of the logarithmic power spectrum within 0–500 Hz band entropy.
- d5: Mean Jitter of the deviations in individual consecutive F0 period lengths.

With the results obtained, the 62 initial parameters were reduced to 5 significant parameters for the differentiation of PD and non-PD diagnosis.

Multilayer neural networks analysis was performed with 70% training and 30% sample testing. AUC-ROC curves were compared as shown in Table 1. The predictive capacity of the diagnostic models with 62, 33, 9 and 5 parameters, was similar according to the values of AUC-ROC. This fact validates the selection of 5 parameters to build our model.

TABLE I. AUC-ROC AREA ACCORDING TO THE NUMBER OF VARIABLES ANALYZED IN NEURAL NETWORKS.

Number of voice variables /parameters	Neural network. AUC
62	0.666
33	0.687
9	0.673
5	0.681

B. Prediction of PD and non-PD diagnosis

We used open software R (www.r-project.org) to build models. It was used for training and testing experiments. The sample was separated in 70% training and 30% testing data.

A two-layer MLP network is a fully-connected feed-forward neural network consisting of an input layer (which is not counted since its neurons are only for representation and thus, do no processing), a hidden layer, and an output layer (PD or non-PD)

The ROC curve for MLP model with 5 voice parameters, in addition to age and gender is observed in Fig. 1. The area under that curve is AUC = 0.826. ACCURACY= 0.768 (cut value=0.5).

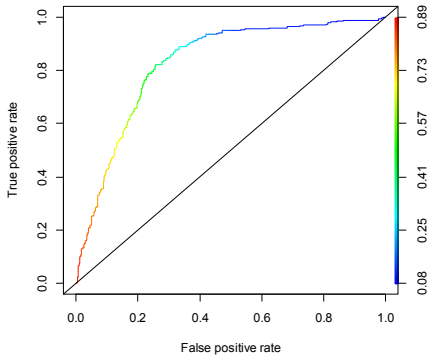

Fig. 1. MLP model: ROC curve

LR was carried out with the same sample. The ROC curve for LR model with 5 voice parameters, in addition to age and gender is observed in Fig. 2. The area under that curve is AUC = 0.832. ACCURACY= 0.773 (cut value=0.5).

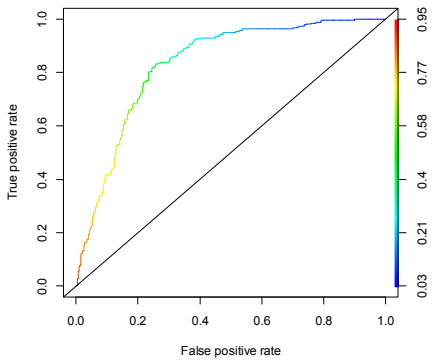

Fig. 2. L R Model

The MLP model with 5 variables, age, gender and medication point is observed with AUC = 0. 972. However, the result doesn´t converge with LR. AUC-ROC curve according to the number of variables analyzed considering MLP and LR are summarized in Table II.

TABLE II. SUMMARY OF THE RESULTS CONSIDERING MLP AND RL

Number of voice variables /parameters	Independent variables	AUC MLP	AUC LR
5	sex / age	0.826	0.832
5	Sex / age /medication time point	0.972	No converge

IV. CONCLUSIONS

The contribution of this work lies in the selection of five simple-to-measure linear parameters of the voice from the free software OpenSMILE. Besides the selection of 5 variables from the original 62, the study analyzed the statistical relevance of individuals´ age and gender. Predictive capacities superior to 0.82 were obtained for the identification of PD and non-PD diagnosis comparing neural networks and logistic regression models.

With the mPower App, voice recordings were collected by volunteer participants with their own Smartphone, and this data were then added to the prediction model of PD diagnosis. Mobile phone recordings both allow for monitoring of Parkinson's patients at a low cost and detection of alterations that require attendance to the medical consultation.

ACKNOWLEDGMENT

This work has been funded by PROINCE-C199 Project, Universidad Nacional de La Matanza, Argentina.

REFERENCES

[1] J. Langston, Parkinson's disease: Current and future challenges, Neurotoxicology 23 (4-5) (2002) 443{450.

[2] D. G. Silva, L. C. Oliveira, M. Andrea, Jitter estimation algorithms for detection of patological voices, EURASIP Journal of Advances in Signal Processing (2009) 1-9.

[3] J. P. Teixeira, C. Oliveira, C. Lopes, Vocal acoustic analysis- Jitter, shimmer and HNR parameters, Procedia Techonology 9 (2013) 1112-1122.

[4] A. Tsanas, M. A. Little, P. E. McSharry, L. O. Ramig, Accurate telemonitoring of Parkinson's disease progression by noninvasive speech test, IEEE Transactions of Biomedical Engineering 57 (4) (2010) 884-893.

[5] Dixit, V. Mittal, Y. Sharma, Discrimination of People with Parkinson (PWP) disease on the basis of voice parameter analysis, International Journal of Computer Applications 94 (13) (2014) 9-14.

[6] R. A. Shirvan, E. Tahami, Voice analysis for detecting Parkinson's disease using genetic algorithm and KNN classi_cation method, in: 18th Conference on Biomedical Engineering, Theran, IEEE, 2011, pp. 278-283.

[7] A. Benba, A. Jilbab, A. Hammouch, Discriminating between patients with Parkinson and neurological diseases using cepstral analysis, IEEE Transactions on Neural Systems and Rehabilitation Engineering 24 (10) (2016) 1100-1108.

[8] B. E. Sakar, M. E. Isenkul, C. O. Sakar, A. Sertbas, F. Gurgen, S. Delil,H. Apaydin, O. Kursun, Collection and analysis of a Parkinson speech dataset with multiple tipes of sound recordings, IEEE Journal of Biomedical and Health Informatics 17 (4) (2013) 828-834.

[9] M. Peker, B. Sen, D. Delen, Computer-aided diagnosis of Parkinson's disease using complex-valued neural networks and mRMR feature selection algorithm, Journal of Healthcare Engineering 6 (3) (2015) 281-302.

[10] A. Tsanas, M. A. Little, P. McSharry, J. Spielman, L. O. Ramig, Novel speech signal processing algorithms for high accuracy classi_cation of Parkinson's disease, IEEE Transactions of Biomedical Engineering 59 (5) (2011) 1264-1271.

[11] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, L. O. Ramig, Suitability of dysphonia measurementes for telemonitoring of Parkinson's disease, IEEE Transactions of Biomedical Engineering 56 (4) (2009) 1015-1022.

[12] P.Schwab, W. Karlen. PhoneMD: Learning to Diagnose Parkinson's Disease from Smartphone Data. arXiv preprint arXiv: (2018). 1810.01485.