A voice analysis approach for recognizing Parkinson's disease patterns

Yu Chen Tai * **Bryan Paucar G.** * **Francis Loayza** **
Enrique Peláez *

* *Escuela Superior Politécnica del Litoral - ESPOL University, Electrical and Computer Engineering, Guayaquil, Ecuador (e-mail: yuchtai@espol.edu.ec, bryalpau@espol.edu.ec, epelaez@espol.edu.ec)*
** *Escuela Superior Politécnica del Litoral - ESPOL University, Mechanical and Production Sciences Engineering, Guayaquil, Ecuador (e-mail: floayza@espol.edu.ec)*

Abstract: Many of the patients diagnosed with Parkinson's disease (PD) do not know they have it until the most severe symptoms appear, sometimes they must wait months or even years to get the correct diagnosis, so detection in its early stage is important to improve the quality of life of patients and families. We propose the creation of a model based on supervised learning, to learn the patterns associated with the voice of PD patients. We used 1400 voice recordings of PD patients and controls which were preprocessed, further were obtained 70 features for each recording, and then we used a supervised learning algorithms such as a Multilayer Perceptron (MLP), Random Forest (RF), Logistic Regression (LR), and Support Vector Machines (SVM) to classify the data between patients and controls. From all machine learning models evaluated the SVM model showed the best performance, with an accuracy of 88%. This work presents the possibility to incorporate the voice analysis as digital biomarker to facilitate diagnosis in PD.

Keywords: Parkinson, Voice Analysis, Machine Learning.

1. INTRODUCTION

The Parkinson disease (PD) is a neurodegenerative disorder that affects predominantly dopamine-producing neurons located in the substantia nigra of the brain. According to WHO [1], about 5.4 million people worldwide were diagnosed with PD in 2013 and its cause still is not well understood. For diagnosing PD, clinical observations of cardinal signs of motor deterioration is done, such as distal resting, tremors, rigidity, bradykinesia in asymetrical onset and others. Physical examination includes postural instability, facial expression, micrographia and decreased olfaction; there is also a psychological evaluation. However, it is clear that in pre-clinical stages, 4 to 6 years before the diagnosis, PD is preceded by a prodromal stage that predates clinical diagnosis, and there are not established methods for detecting this stage. Despite advances in imaging analysis and radiologic testing, PD diagnosis at the early stages of the disease remains complex and in most cases a non-accurate task.

There have been some efforts to developed methods to diagnose the PD at an early or prodromal stage; such as the study performed by Rolheiser et al. (2011), which is based on diffusion tensor imaging of the olfactory tract combined with behavioral olfaction, as a biomarker to identify early PD, others focused in identify differences in white matter hyperintensities features Chancay et al. (2015); Viteri et al. (2021). Several studies have also described the effects of the disease on speech impairment; Ho et al. (1998), found that voice was the leading deficit, as compared to other effects in initial stages, and recent research is focusing on analyzing the speech impairment identified in PD patients. According to New et al. (2015), one of the first PD symptoms identified is the change in speech, as well as a language impairment. It has been estimated that more than 90% of PD patients and Parkinsonism, present high degree of speech disorders; such as, articulation (dysarthria), spoken language production (dysprosody) or voice (dysphonia), as reported by Sapir (2014); Mostafa et al. (2019) and Lirani-Silva et al. (2015). These speech disorders have led to the proposal of telemedicine systems, which are helping in the early detection of the disease as stated by Tsanas et al. (2009) and Peker et al. (2015), which use simple and low cost methods and technologies to support physical examinations and the workload of physicians. Recent work has been focused on exploring speech distortions, through the audio analysis, using protocols that allow to identify specific changes in speech of PD patients, as it has been proposed by J. Holmes et al. (2000), who defined a protocol for recording the patient's audios, e.g. asking a patient to speak and hold the vowel *'aaaa'* for about 10 seconds, and their subsequent analysis. Over time these types of protocols were replicated in other studies, as in Tsanas et al. (2012); Sakar and Kursun (2010) and Arora et al. (2018) in the differentiation between genetic and idiopathic Parkinson's disease, using Artificial Intelligence (AI) techniques for classifying speech signals.

[1] World Health Organization

In this work, we propose a methodology for preprocessing voice recordings, then identifying and extracting the relevant features related to the PD patterns within those signals, then we have evaluated 4 Machine Learning (ML) models to carry out a binary classification, which allowed us to classify the PD or not PD based on the voice analysis, these models are: Random Forest (RF), Support Vector Machines (SVM), Logistic Regression (LR), and a Multilayer Perceptron (MLP). Furthermore, we have compared their performance in classifying the patterns, as PD or not PD with a given probability.

2. THE DATASET AND METHODOLOGY

The dataset used in this work comes from the Sage Bionetworks mPower Project, called "Mobile Parkinson Disease Study", Bot et al. (2016), This dataset contains a collection of more than 65000 audios, which have been recorded using smart phones. The audios have the voice recordings of the vowel *'aaaa'*, with a duration of 10 seconds, sampled at a frequency of 44.1 kHz. and have been saved in `".m4a"` format. However, most of them are replications from the same subject or noise contaminated samples. Also, for the purpose of this work, a set of voice recordings from a control group of subjects was included. Therefore, the criteria used for filtering the audio files is described in Table 1:

Table 1. Selection criteria for selection of audio files

Patients
Who did the recording prior to taking Parkinson's medication.
Have a positive PD diagnosis from a professional.
Consume any medication for PD.
Range in age is from 50 to 75 years old.
Control group
Defined as healthy controls
Age range from 50 to 75 years old.

Using this filter selection criteria, we found 1201 recording audios from healthy controls, and 6405 audios of people with positive PD.

For preprocessing the signals, as well as for training the models, we used: Python ver. 3.0, and for feature extraction from the preprocessed audio set, we used the following libraries: Librosa and Parselmouth available for Python. Additionally, for model implementing we used Google Colab [2] , which provided us with 12 Gb of RAM and access to GPUs for free.

2.1 Audio preprocessing

Since the voice data was captured using smart phones, the recording environment was not controlled, affecting the audio's quality. Therefore, the cleaning of the background noise was performed in the preprocessing phase; noise, such as sounds coming from TV sets, animals, cars, people talking, among others; cleaning included intensity audio volume normalization.

According to Rueda and Krishnan (2018), audio signal alterations will be present depending on the recording decibel's level. If the signal has a value greater than 0 decibels (the saturation point), or less than -30 decibels, then some important features could be lost. Therefore, it is advisable to use audio signals within the frequency range [-20db to -3db]. Additionally, consider the time window selection where the signal is stable. In this work, after performing a time window analysis we found that, in average after the first second, there was about 3 seconds stability of the voice across all subjects.

In order to increase the number of healthy control samples, 36 new local recordings were acquired from 16 men and 20 women, which were analyzed using the same protocol as the mPower Project, as well as the same audio recording format.

After applying this criteria, 1400 qualified audios were selected, 700 samples corresponded to PD patients and the remaining 700, to healthy controls.

Table 2 shows a summary of the audio files belonging to each genre and the PD status.

Table 2. Summary of the final Dataset

Parkinson's disease	**Genre**	**Age**	**Quantity**	**Total**
Parkinson's disease	Male	56 - 71	442	700
	Female	56 - 71	258	
Healthy controls	Male	50 - 74	474	700
	Female	50 - 74	226	

2.2 Feature extraction

As recommended by J. Holmes et al. (2000), and once the audio signals have been preprocessed, we extracted 70 features as described in the Table 3, which allowed us to determine: The shimmer, jitter, maximum pitch, minimum pitch, minimum tone, harmonics-to-noise ratio (HNR), which calculates the amount of noise in the speech signal, number of pulses, and the fundamental frequency. Also, to analyze healthy voice signals and dysphonia, MFCCs (Mel Frequency Cepstral Coefficients) were extracted, which are the features used in speech recognition and allow to model the spectral energy distribution of a signal, based on a linear cosine transform of a log power spectrum on a nonlinear Mel scale frequency Majeed et al. (2015); using these functions, the following features were extracted: The MFCC, a delta MFCC and their means and standard deviations, as recommended by Rueda and Krishnan (2018). Table 3, shows the list of these features. Notice that row 71 was added, which contains the label (PD or HCs).

2.3 Dimensionality reduction

To simplify the number of features, as well as to reduce the possibilities of overfitting, and the model's training costs, two techniques were applied: a High Correlation Filter (HCF); which allowed us the determine the best correlations between the features and the expected output label; and, Principal Component Analysis (PCA), to determine those features that contribute with the most information from the preprocessed dataset.

High Correlation Filter - HCF. The HCF allowed us to find the correlation of each and all features, as

[2] www.colab.research.google.com

Table 3. List of features

Id	Features	Description
1	HNR	Harmonic-to-noise
2	apq11Shimmer	Shimmer (apq11)
3	apq3Shimmer	Shimmer(apq3)
4	apq5Shimmer	Shimmer (apq5)
5	ddaShimmer	Shimmer (dda)
6	ddpJitter	Jitter (ddp)
7-19	desvMFCC	13 MFCC - Standard Deviation
20-32	desvMFCCdelta	13 delta MFCC - Standard Deviation
33	localJitter	Jitter (local)
34	localShimmer	Shimmer(local)
35	localabsoluteJitter	Jitter (local, absolute)
36	localdbShimmer	Shimmer (localdb)
37	max_pitch	Maximum pitch
38	meanF0Hz	F0 - Mean
39-51	meanMFCC	13 MFCC - Mean
52-64	meanMFCCdelta	13 delta MFCC - Mean
65	min_pitch	Minimum pitch
66	n_periods	Number of periods
67	n_pulses	Number of pulses
68	ppq5Jitter	Jitter (ppq)
69	rapJitter	Jitter (rap)
70	stdevF0Hz	F0 - Standard Deviation
71	Status	Label (PD or HCs)

compared in pairs to the expected output label, and estimated by (1):

$$r = \frac{(XY)}{\sqrt{(XX)(YY)}} \tag{1}$$

Where X, the input feature matrix, explains the total variability with respect to the output vector Y. The correlation value r can be positive or negative in the range [-1, 0, +1]; where zero indicates the absence of a relationship; closer to 1, regardless of the sign, shows a strong association between X and Y.

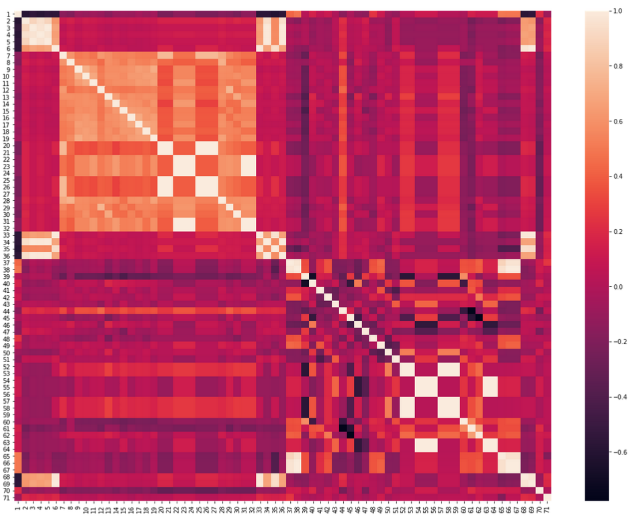


Fig. 1. Heat Map showing the total variability of each feature with respect to the output

To visualize the strength of these associations a *Heat Map* was created, as it can be seen in Figure 1, the brighter the color the stronger the positive correlation, and the darker the color, the stronger the negative correlation).

Principal Component Analysis - PCA. Through this technique, a matrix of Principal Components (PCs) was generated, from the set of selected features extracted from the Co-variance matrix and generated by (2); these PCs are linear combinations of the features or independent variables; as a result, the PC with the highest variance is the first component, the second with the highest variance is the second component, and so forth, as defined in Maćkiewicz and Ratajczak (1993).

$$Cov(A_iX, A_jX) = 0 \qquad for \quad i \neq j \tag{2}$$

Where X is a matrix, with columns representing the features' vectors of the signals, and labeled as $X_1, X_2, ...X_p$; each column vector represents a point in $p-dimensions$, as expressed in (3). A contains the orthonormal eigen-vectors of the variance-covariance matrix, derived from X.

$$X = \begin{bmatrix} x_{1_1} & x_{1_2} & \cdots & x_{1_n} \\ x_{2_1} & x_{2_2} & \cdots & x_{2_n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{p_1} & x_{p_2} & \cdots & x_{p_n} \end{bmatrix} \tag{3}$$

The selection of PCs is based on the number of components that maximizes the percentage of variance, in this case we have selected the first 27 PCs, which preserve around 97% of the information. Figure 2 shows the corresponding percentage of variance for each of the 27 PCs.

	var		
PC1	23.382955	PC16	1.319273
PC2	14.440023	PC17	1.071065
PC3	9.650829	PC18	0.923636
PC4	8.800972	PC19	0.819421
PC5	7.613228	PC20	0.792613
PC6	4.906104	PC21	0.675027
PC7	3.655843	PC22	0.632107
PC8	2.886277	PC23	0.525305
PC9	2.500326	PC24	0.521412
PC10	2.333172	PC25	0.456264
PC11	2.217188	PC26	0.439215
PC12	1.878570	PC27	0.399676
PC13	1.545714		
PC14	1.464331		
PC15	1.396541		

Fig. 2. Percentage of variance for the 27 Principal Components

2.4 Description of hyper-parameters

Multilayer Perceptron - MLP. After evaluating several configurations and models, an MLP architecture with 4 layers was designed. The input layer has 27 neurons, based on the number of input features; 2 hidden layers with 128 and 64 of neurons respectively; and, since this is a binary classification problem, the output layer has 1 neuron.

The activation function defined for the hidden layers was `tanh` mainly because, as it was seen before, the features extracted are in the [-1, +1] range; at the output layer the `Sigmoid` function was used since it is a binary classification problem. The function `binary-crossentropy` was chosen to evaluate the loss during the learning process.

For implementing the neural network we used the Scikit-learn and Keras libraries available in Python, and the optimizer selected for this model was `Adam`. Searching for the best parameters was performed using `Stochastic Gradient Descent (SGD)` and for updating the parameters we used `Back-propagation` during 42 epochs; that is, until we reached the maximum accuracy and the lowest loss. With this model an accuracy of 86% was achieved, using the PCA as a dimensional reduction method and applying a fine tuning technique, the following hyper-parameter configuration was obtained. (see Table 4)

Table 4. MLP configuration

Parameters	Value
Learning rate	0.1
momentum	0.8
kernel_initializer	he_normal
Dropout1	0.5
Dropout2	0.0
Dropout3	0.1
neurons1	128
neurons2	256
neurons3	64
neurons4	1
validation split	0.2
epochs	100
batch_size	40

Support Vector Machine (SVM). The SVM is an algorithm widely used in classification, which is based on statistical learning theory. This algorithm is a discriminative classifier defined on a separation hyperplane, as presented by Vapnik (1995), the linear classification equation is defined as follows:

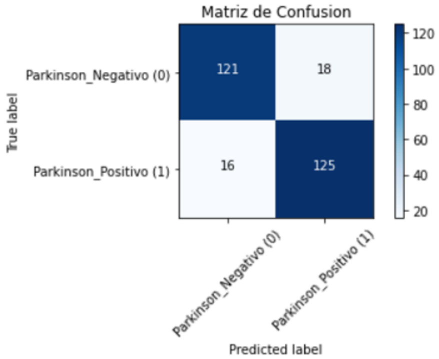
$$f(x) = sign(\sum_{i=1}^{l} y_i c_i K^*(z(x), z(x_i))) \quad (4)$$

Where K is the weight with which the dot product of the support vector *z(x)* and the input vector $z(x_i)$ is multiplied with the combination of y_i and c_i. With this model an accuracy of 88% was achieved, using the high correlation filter method and after fine tuning, the following hyper-parameter configuration was obtained. (see Table 5)

Table 5. SVM Configuration

Parameters	Value
C	5
gamma	0.00001
kernel	rbf

Using these hyper-parameters the percentage of predictions for class 0, or negative PD, was 87%; and, 89% for class 1, or positive PD. Looking at the confusion matrix shown in Figure 3, this model predicted 18 false negatives out of 139 audios in the test set for class 0, and 16 false positives out of 141 audios for class 1.



	precision	recall	f1-score	support
0	0.88	0.87	0.88	139
1	0.87	0.89	0.88	141
accuracy			0.88	280
macro avg	0.88	0.88	0.88	280
weighted avg	0.88	0.88	0.88	280

Fig. 3. Confusion Matrix for the SVM model

Logistic Regression. Logistic Regression is one of the simplest supervised autonomous learning algorithms used for classification, it basically contains one processing unit; it is a statistical method for predicting binary classes with dichotomous outputs; hence, used for binary classification problems. This algorithm is based on a `Sigmoid` function, which can be designed to estimate a probability of the expected prediction. The model is defined as follows:

$$f(l) : p(x) = \frac{1}{1 + \mathrm{e}^{(-f(l))}} \quad (5)$$

Where *f(l)* represents a linear model, however the `Sigmoid` function is applied to restrict the values in the range of (0,1). With this model and using the high correlation filter an 80% accuracy was achieved, after fine tuning, the following hyper-parameter configuration was obtained. (see Table 6).

Table 6. Logistic Regression Configurations

Parameters	Value
C	100
max_iter	500
penalty	L2

Random Forest. A Random Forest model is a supervised learning algorithm that can be used for regression and classification tasks. At start, each decision tree, as a weak classifier, results in a binary classification, which is subject to a series of binary tests at each node of the tree, these are called Splits, and the more trees there are the more robust the forest becomes, resulting in a strong classifier. This technique is useful to convert large problems into simple ones, finally to classify a sample, all trees are averaged, as expressed in (6).

$$f = \frac{1}{B}\sum_{b=1}^{B} f_b(x^{"}) \quad (6)$$

Where B is the number of random samples and fb is the classification or the regression tree function.

A Random Forest classifier is ideal for handling a large amount of data and multiple variables, since it identifies subsamples to elaborate each tree, it is considered as a dimensional reduction method, it also provides methods for estimating missing data. With this model, an accuracy of 82% was achieved, using the PCA method and after fine tuning the design parameters, the following hyper-parameter configuration was obtained, (see Table 7).

Table 7. Random Forest Configurations

Parameters	**Value**
n_estimators	50
min_samples_leaf	10
max_features	log2

Assessment metrics. To evaluate the effectiveness of the classifiers, different metrics were computed and evaluated: A confusion matrix for each model, such as the one shown in Figure 3; and the accuracy, sensitivity and specificity scores. With the confusion matrix we obtained information about the predicted and expected classification. It allows us to visualize the hits and misses in the prediction and through this, we were able to obtain metrics of error, precision, sensitivity and specificity, as represented in Table 8.

In this work the accuracy score was used to measure the model's prediction accuracy with respect to all instances and is calculated as follows:

$$Accuracy = \frac{TP + TN}{(TP + FP + TN + FN)} \quad (7)$$

Where TP, the true positives and FP, the false positives represent the correct predictions of the model. And, TN, true negatives and FN, the false negatives represent the predictions where the model was wrong. The specificity measures the true negative rate as the ratio of the number of true negative predictions compared to the total true negatives and false positives, as expressed by (8)

$$Specificity = \frac{TN}{(FP + TN)} \quad (8)$$

Table 8. Confusion matrix structure

	Positive	**Negative**
Positive	TP	FN
Negative	FP	TN

We used the Sensitivity score to measure the true positive rate, as the ratio of the number of true positive predictions compared to the total true positive actual values, as expressed by (9).

$$Sensitivity = \frac{TP}{(TP + FN)} \quad (9)$$

3. RESULTS AND DISCUSSION

Once the training, validation and evaluation of performance scores of all models have been extracted, we found that the SVM model obtained the best scores, either using the High Correlation Filter (88% accuracy) or the dimensional reduction model (87% accuracy). The results showed that using the HCF, with the classifiers based on the extraction of a decision border, such as the SVM, MLP and LR performed better than the model with dimensional reduction, with an accuracy between 80% and 87%, as compared to the Random Forest, with an accuracy of 75%. However, using PCA the RF model improves its prediction, with an accuracy of 82%, and at the same time affecting the performance of the LR model, with an accuracy of 71%.

As Table 9 shows, the performance of the MLP and the SVM models are comparable after using the two dimensional reduction techniques, with an accuracy above 85% in average; however, the SVM outperforms all models, with an accuracy of 88% in average. These results were also confirmed with the Receiver Operating Characteristic (ROC) curves analysis. Figure 4 shows the ROC curves for each model. These curves measure the balance between sensitivity and specificity. From Figure 4, the classification models that come closest to the upper part of the "Y" axis perform better, such models are identified as ideal discriminators, in our case these are the SVM and MLP classifiers. However, as we can see in this figure, the SVM classifier performed the best. On the other hand, the classifier that came closest to the 45-degree diagonal of the ROC space is the Logistic Regression, which makes it the least accurate in this evaluation.

Considering the number of features extracted from the voice signals, and to be able to classify them as a positive or negative class for PD, it was necessary to reduce their dimensions. Two dimensionality reduction techniques were evaluated, a High Correlation Filter (HCF) and Principal Component Analysis (PCA), which allowed us to improve the performance of the classifiers. In both cases, a decision has to be taken as to define the threshold of the correlated features, which allows to keep those features with strong correlation with the expected diagnosis, such as in the HCF technique; or, to decide the maximum variance to be preserved, which is defined with the number of PCs to be included in the analysis.

Table 9. Accuracy scores obtained using dimension reduction techniques

	High Correlation Filter	**Principal Component Analysis**
MLP	83%	86%
RF	75%	82%
SVM	88%	87%
LR	80%	71%

Table 9 presents a summary of the accuracy scores for each model, after applying the two dimensional reduction techniques; form this evaluation the Support Vector Machine (SVM) model, using a High Correlation Filter, was the model with the best performance.

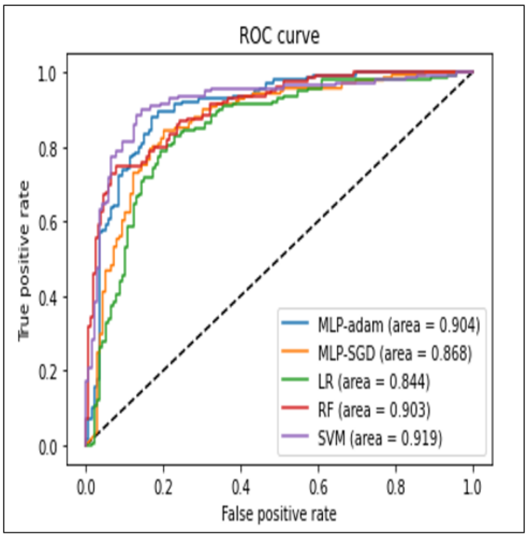


Fig. 4. ROC curve of supervised learning algorithms after using dimensionality reduction

4. CONCLUSIONS

The analysis of speech distortions, using preprocessing techniques and Machine Learning (ML) models, for identifying specific signals in speech, allowed us to extract relevant features related to the PD patterns of their voice recordings in a non controlled environment. From the 4 ML models evaluated, to perform a binary classification, the SVM model outperformed all models, with an average accuracy of 88%. This results demonstrate that the voice or speech analysis in a controlled environment could be an important digital biomarker which may be helpful for an early PD diagnosis, specially for those individuals in a prodromal phase.

REFERENCES

Arora, S., Visanji, N.P., Mestre, T.A., Tsanas, A., Al-Dakheel, A., Connolly, B.S., Gasca-Salas, C., Kern, D.S., Jain, J., Slow, E.J., et al. (2018). Investigating voice as a biomarker for leucine-rich repeat kinase 2-associated parkinson's disease. *Journal of Parkinson's disease*, 8(4), 503–510.

Bot, B.M., Suver, C., Neto, E.C., Kellen, M., Klein, A., Bare, C., Doerr, M., Pratap, A., Wilbanks, J., Dorsey, E.R., et al. (2016). The mpower study, parkinson disease mobile data collected using researchkit. *Scientific data*, 3(1), 1–9.

Chancay, O., Haro, T., Yapur, M., Alvarado, R., Pastor, M., and Loayza, F. (2015). Nuevo biomarcador en la enfermedad de parkinson mediante el análisis y cuantificación de lesiones cerebrales en secuencias flair obtenidas por resonancia magnética (acl-tool). *Revista Tecnológica-ESPOL*, 28(5).

Ho, A.K., Iansek, R., Marigliani, C., Bradshaw, J.L., and Gates, S. (1998). Speech impairment in a large sample of patients with parkinson's disease. *Behavioural neurology*, 11(3), 131–137.

J. Holmes, R., M. Oates, J., J. Phyland, D., and J. Hughes, A. (2000). Voice characteristics in the progression of parkinson's disease. *International Journal of Language & Communication Disorders*, 35(3), 407–418.

Lirani-Silva, C., Mourão, L.F., and Gobbi, L.T.B. (2015). Dysarthria and quality of life in neurologically healthy elderly and patients with parkinson's disease. In *CoDAS*, volume 27, 248–254. SciELO Brasil.

Maćkiewicz, A. and Ratajczak, W. (1993). Principal components analysis (pca). *Computers & Geosciences*, 19(3), 303–342.

Majeed, S.A., Husain, H., Samad, S.A., and Idbeaa, T.F. (2015). Mel frequency cepstral coefficients (mfcc) feature extraction enhancement in the application of speech recognition: a comparison study. *Journal of Theoretical and Applied Information Technology*, 79(1), 38.

Mostafa, S.A., Mustapha, A., Mohammed, M.A., Hamed, R.I., Arunkumar, N., Abd Ghani, M.K., Jaber, M.M., and Khaleefah, S.H. (2019). Examining multiple feature evaluation and classification methods for improving the diagnosis of parkinson's disease. *Cognitive Systems Research*, 54, 90–99. doi: https://doi.org/10.1016/j.cogsys.2018.12.004.

New, A.B., Robin, D.A., Parkinson, A.L., Eickhoff, C.R., Reetz, K., Hoffstaedter, F., Mathys, C., Sudmeyer, M., Michely, J., Caspers, J., et al. (2015). The intrinsic resting state voice network in parkinson's disease. *Human brain mapping*, 36(5), 1951–1962.

Peker, M., Sen, B., and Delen, D. (2015). Computer-aided diagnosis of parkinson's disease using complex-valued neural networks and mrmr feature selection algorithm. *Journal of healthcare engineering*, 6(3), 281–302.

Rolheiser, T.M., Fulton, H.G., Good, K.P., Fisk, J.D., McKelvey, J.R., Scherfler, C., Khan, N.M., Leslie, R.A., and Robertson, H.A. (2011). Diffusion tensor imaging and olfactory identification testing in early-stage parkinson's disease. *Journal of neurology*, 258(7), 1254–1260.

Rueda, A. and Krishnan, S. (2018). Clustering parkinson's and age-related voice impairment signal features for unsupervised learning. *Advances in Data Science and Adaptive Analysis*, 10(02), 1840007.

Sakar, C.O. and Kursun, O. (2010). Telediagnosis of parkinson's disease using measurements of dysphonia. *Journal of medical systems*, 34(4), 591–599.

Sapir, S. (2014). Multiple factors are involved in the dysarthria associated with parkinson's disease: a review with implications for clinical practice and research. *Journal of Speech, Language, and Hearing Research*, 57(4), 1330–1343.

Tsanas, A., Little, M., McSharry, P., and Ramig, L. (2009). Accurate telemonitoring of parkinson's disease progression by non-invasive speech tests. *Nature Precedings*, 1–1.

Tsanas, A., Little, M.A., McSharry, P.E., Spielman, J., and Ramig, L.O. (2012). Novel speech signal processing algorithms for high-accuracy classification of parkinson's disease. *IEEE transactions on biomedical engineering*, 59(5), 1264–1271.

Vapnik, V. (1995). The nature of statistical learning theory. *New York*.

Viteri, J.A., Loayza, F.R., Peláez, E., and Layedra, F. (2021). Automatic brain white matter hypertinsities segmentation using deep learning techniques. In *HEALTHINF*, 244–252.