

Application of Signal Classifiers in Auditory Evoked Potentials for the Detection of Pathological Patients

M. G. BALDIVIEZO¹, J. L. BARBERÍA², C. B. BONTEMPO³, Y. CORSARO⁴,
F. FERNANDEZ BIANCARDI⁵, M. R. HERNANDO⁶, L. LICATA CARUSO^{7*},
A. PAGLIA⁸, M. R. RODRIGUEZ⁹ and W. E. LEGNANI¹⁰

Received on October 6, 2021 / Accepted on July 10, 2022

ABSTRACT. The auditory brainstem response (ABR) by evoked potentials is a widespread auditory pathway assessment technique. This is largely applied due to its cost-effectiveness, practicality and ease of use. In contrast, it requires a trained professional to carry out the analysis of the results. This motivates several research efforts to increase the independence of the diagnostician. To this end, the present work shows the ability of three signal classification tools to differentiate ABR studies of normal hearing subjects from those who may have some pathology. As a starting point, the PhysioNet short term auditory evoked potentials databases are used to calculate the features later applied to construct the dataset. The features used are diverse classes of permutation entropy, fractal dimension, the Lyapunov exponent and the zero crossing rate. To ensure more accurate results, a Montecarlo simulation of one thousand trials is employed to train the

*Corresponding author: Lorenzo Licata Caruso – E-mail: lorenzolicata@frba.utn.edu.ar

¹Signal and Image Processing Center, Facultad Regional Buenos Aires, Av. Medrano 951, 1179, Ciudad Autonoma de Buenos Aires, Argentina - Email: mbaldiviezo@frba.utn.edu.ar <https://orcid.org/0000-0003-1271-9373>

²Signal and Image Processing Center, Facultad Regional Buenos Aires, Av. Medrano 951, 1179, Ciudad Autonoma de Buenos Aires, Argentina - Email: jbarberia@frba.utn.edu.ar <https://orcid.org/0000-0001-7632-6830>

³Signal and Image Processing Center, Facultad Regional Buenos Aires, Av. Medrano 951, 1179, Ciudad Autonoma de Buenos Aires, Argentina - Email: cbontempo@frba.utn.edu.ar <https://orcid.org/0000-0002-5867-162X>

⁴Signal and Image Processing Center, Facultad Regional Buenos Aires, Av. Medrano 951, 1179, Ciudad Autonoma de Buenos Aires, Argentina - Email: yaninacorsaro@hotmail.com <https://orcid.org/0000-0001-6227-8591>

⁵Signal and Image Processing Center, Facultad Regional Buenos Aires, Av. Medrano 951, 1179, Ciudad Autonoma de Buenos Aires, Argentina - Email: jfernandezbiancardi@frba.utn.edu.ar <https://orcid.org/0000-0003-2245-0064>

⁶Signal and Image Processing Center, Facultad Regional Buenos Aires, Av. Medrano 951, 1179, Ciudad Autonoma de Buenos Aires, Argentina - Email: rhernando@frba.utn.edu.ar <https://orcid.org/0000-0003-2471-7201>

⁷Signal and Image Processing Center, Facultad Regional Buenos Aires, Av. Medrano 951, 1179, Ciudad Autonoma de Buenos Aires, Argentina - Email: lorenzolicata@frba.utn.edu.ar <https://orcid.org/0000-0001-6652-5318>

⁸Signal and Image Processing Center, Facultad Regional Buenos Aires, Av. Medrano 951, 1179, Ciudad Autonoma de Buenos Aires, Argentina - Email: adrianepaglia@gmail.com <https://orcid.org/0000-0002-8906-7054>

⁹Signal and Image Processing Center, Facultad Regional Buenos Aires, Av. Medrano 951, 1179, Ciudad Autonoma de Buenos Aires, Argentina - Email: matias.rodriguez.97.13@gmail.com <https://orcid.org/0000-0002-1784-2613>

¹⁰Signal and Image Processing Center, Facultad Regional Buenos Aires, Av. Medrano 951, 1179, Ciudad Autonoma de Buenos Aires, Argentina - Email: walter@frba.utn.edu.ar <https://orcid.org/0000-0002-6949-0728>

classifiers and test the results. The goodness of the classification is performed upon the basis of the computation of quality parameters, such as the area under the receiver operating characteristic curve (ROC), the F1 score coefficient and the accuracy. In all the cases the methodology proposed gives high performance results that encourage the line of research of the present work.

Keywords: auditory evoked potentials, permutation entropy, signal classification, fractal dimension, Lyapunov exponent, zero crossing rate, support vector machines, random forest, K-nearest neighbours.

1 INTRODUCTION

When a sound stimulus reaches the ear, it produces several responses of the auditory structures, originating a rise of the evoked auditory potential. Mathematically, this response is made up of continuous waves, and can be registered with the appropriate systems.

The technique under analysis is a non-invasive procedure, which allows the assessment of sensory and neural function of the auditory nervous system and also represents an indirect but important resource for verifying and localizing neurological dysfunction within non-sensory systems, as the sensory pathways run parallel to the motor pathways and pass close to the encephalic areas linked to the vegetative processes of consciousness and cognition. This allows the clinic to reveal nervous system dysfunction that is not otherwise apparent (*cf* [11, 16]). It is currently a very useful study, but it has become very dependent on the experience of the audiologist or neurologist who analyzes the study. In this sense, the aim of this proposal is to develop an alternative methodology using objective parameters that contribute to the diagnosis and characterisation of the auditory response [28].

The ABR have become an integral part of current otological and audiological study methods and spectral methods have been successfully applied, as for example in [8]. The relevance in the use of this kind of electro-physiological procedures in the clinic can be understood by some example applications, in this sense can be mentioned: assessment of auditory pathway function and pathology, evaluation of brainstem development and pathologies, auditory threshold testing, selection of hearing aids, intra-surgical monitoring and analysis of brain activity for the determination of the so-called "brain death", among other applications.

To accomplish the objective of this work several magnitudes are calculated from the ABR records, that can be divided in three sets: one characteristic of the information measures, other corresponding to nonlinear system and signal analysis, and the third an optimization of the permutation entropy from the best choice of embedding time delay. Due to the noisy content of the ABR signals, taken from the Physionet databases, it is necessary to apply a prefiltering in order to avoid the contamination of the results with spurious components. Joining the entropic measures with the nonlinear parameters a wide dataset is constructed with the data of pathological and normal ABRs. Added these magnitudes to the database the dataset is composed by different combinations of them to analyze in the feature space. The entropy measures calculated from the signals are the Bandt and Pompe permutation entropy, Rényi Entropy, weighted permutation entropy, an improvement of the Bandt and Pompe permutation entropy to manage the missing

pattern problem of the records, and a Bandt and Pompe entropy calculated with the optimization of the parameter mentioned above. The nonlinear parameters calculated to be integrated in the dataset are the Lyapunov exponent and the fractal dimension. Lastly, the Zero Crossing Rate is also used as a feature in this study.

To formulate a model using the dataset constructed three conceptually different methodologies are analyzed: support vector machines, random forest and k-nearest neighbours. The aim of this stage of the work is to test different classification algorithms to evaluate the performance in differentiating between normal and pathological ABR records. Then, a Montecarlo simulation of one thousand trials is performed to improve the results of the classification algorithms applied. In every Montecarlo trial a seventy percent of the whole dataset is assigned to train the classifier and the remaining is used to classify. The quality of the classification is measured by the calculation of the area under the receiver operating characteristic curve (ROC), the F1 score coefficient and the accuracy. In all cases, with the datasets constructed, the selected classifiers gave very satisfactory results.

The present research is structured in the following sequence: section one dedicated to the introduction, section two where the database and prefiltering are described, section three presents the generalities of the classifiers selected, section four describe the dataset features, section five shows the results (including the features of the dataset employed), and section six contains the discussion and analysis, the conclusions are developed in section seven, the final section of the work presents the corresponding bibliography.

2 DATABASE AND PREFILTERING

2.1 Database

Two databases are used in this study, both located at the PhysioNet website: Evoked Auditory Responses in Hearing Impaired (EARH) [34] and Evoked Auditory Responses in Normals (EARNDB) [33]. They were provided under the same conditions, and consisted of Auditory Brainstem Response (ABR) and Otoacoustic Emission (OAE) signals, although in this work only the ABR ones are used.

The EARH database holds the signals corresponding to eight listeners with a known history of hearing impairment confirmed through clinical tests, ages ranging from 40 to 85 years old, four females and four males. The records were acquired on the ear with the best threshold: three subjects with their left ear and five with their right ear. The EARNDB provides the signals of eight subjects with normal hearing, again four females and four males, whose ages go from 19 to 31 years old. This time all of them were arbitrarily tested on their right ear.

For both sets the ABRs were recorded simultaneously with the OAEs in a sound-attenuating, electrically shielded booth. Listeners had three electrodes affixed as follows: the non-inverting electrode was positioned on the forehead, the inverting electrode was positioned on the ipsilateral mastoid (behind the ear), and the ground electrode positioned on the contralateral mastoid.

The tests used tone burst stimuli, specifically of 1 kHz tones with 4 ms duration and 4 kHz tones with 1 ms duration, each stimulus consisted of 12 concatenated 41.7 ms intervals in order to generate a train of tone bursts that lasted approximately 0.5 seconds. All signals were converted from analog to digital sampling at 48 kHz, resulting in a signal length of 10^6 points. The stimulus level ranged from the listeners threshold to 100 dB in steps of 5 dB, and due to this threshold some of the hearing impaired listeners could not be tested at 4 kHz.

From a previous study [4] with these signals, they suggested two things: a denoising filter was needed and the signals had to be trimmed to 5000 points. Besides, less than 20% of the EARNDDB signals presented some issue with the recordings so those were withdrawn from this study.

After removing the questionable signals, trimming all of the rest down to 5000 points and filtering them, a unique database with 2039 signals corresponding to normal hearing subjects and 565 signals for the hearing impaired patients was generated.

2.2 Prefiltering

The signals used to construct the dataset are taken from the PhysioNet database that are available in raw format, which means they require a prefiltering to remove the noise with which they were contaminated. This signals pre-processing accomplishes to minimize the effect of noise on the calculations of the features to be used with the classifiers.

To avoid unwanted attenuation and phase shifting of the signals a wavelet filter is selected [7]. In such a way a prefiltering based on denoising by wavelets is implemented using the Symlet 5 mother wavelet, in a filter bank with 6 levels of analysis and detail [36], [21], which allows the signal to be reconstructed with the contribution of the most significant levels of analysis, thus discarding the noisy component of the signal associated with the detail coefficients corresponding to the higher frequencies. The computational implementation of the denoising is done in Python using the library PyWavelets [20].

3 CLASSIFIERS

There is some background about the application of classifications methodologies using evoked potential records, see for example [19] [22] and [2]. Evaluation of the progression of the Multiple Sclerosis Disease has been studied in [39] and distinguishing age from infants can be seen in the work presented in [29]. Classifications algorithms in the use in clinical rehabilitation have been used in [35]. In the pathway to increase the objectivity of the ABR analysis, the work of [23] is an interesting study. To distinguish different hearing perception levels using machine learning, the work of [25] constitutes a good source of information. In [1] a study about ABR using support vector machines (SVM) is presented. Finally, to increase the impact of the machine learning methodologies in the clinical research the results of [30] are profusely illustrative.

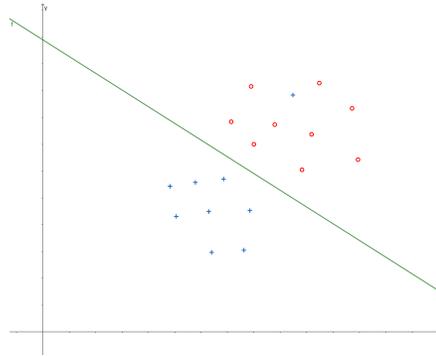


Figure 1: Outliers.

3.1 Support Vector Machines

Support Vector Machines (SVMs) are non probabilistic linear binary classifiers, with several applications in Biology [24]. This methodology is based upon finding a hyperplane that maximizes the gap separating two categories in the multidimensional feature space.

One of the advantages of SVMs is its capability to detect outliers. This allows a limited part of the dataset to stay on the wrong side of the hyperplane, as can be seen in figure 1 without affecting the final result, taking into account the amount of outliers and the distance of these to the hyperplane.

The SVM used in this work is set up with a third degree polynomial kernel type. To fit this kernel it uses the features set selected according to the capability of the library SKLearn of Python [26] and null constant kernel coefficient.

3.2 Random forest

The random forest algorithm is an improvement of the single decision tree method, generating a forest of several trees which splits with randomly generated vectors, or random subsets of training data, and computes the score as a function of these different components [8] as shown in Figure 2.

This algorithm presents a series of advantages: it is more precise than a single decision tree, it reduces the over fitting problems, and a random subset of features is selected at each node in every tree of the forest.

Among the disadvantages that can be mentioned: it is computationally more expensive than a single decision tree algorithm and it is harder to program and execute.

The Random Forest used in this study is configured with 100 decision trees. The criterion used to decide the branch selected is based upon the computation of the Shannon entropy. Due to the

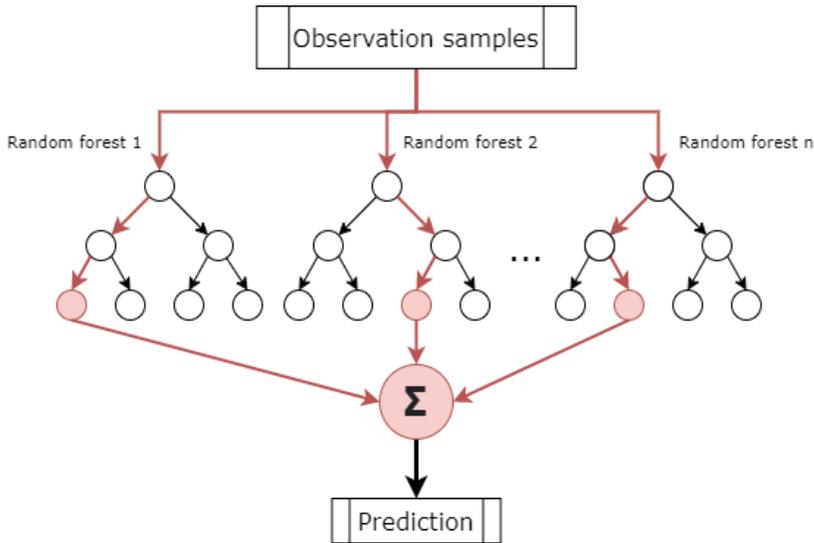


Figure 2: Exemplification of the random forest algorithm.

dichotomy of the diagnostic task, every node is split in 2 branches. Also, every decision tree uses 3 variables of the feature set. The library used is the same that indicated in the previous case.

3.3 *K-Nearest Neighbors*

The K-Nearest Neighbors (KNN) algorithm is one of the most basic and simplest classification methods. It is a good choice when there is little knowledge about the distribution of the data [27]. This algorithm uses as classification criteria the distance among the neighbors, as can be seen in Figure 3.

The KNN algorithm used in the present research is set up with 100 neighbors weighted by the euclidean distance in which the near neighbors have more influence in the construction of the classes. This classification is implemented by the same library above mentioned.

4 FEATURES

4.1 Higuchi's Fractal Dimension

The fractal dimension is an exponent that allows us to quantify how much a fractal seems to fill a geometric space as we get closer to smaller scales. It can be used to determine the degree of irregularity of a signal as well as, for example, Fourier spectral analysis, which can be used more precisely through the power spectrum that exists on every frequency (denoted by ν). As the power spectrum of a signal follows a law of the type $P(\nu) \propto \nu^{-\alpha}$, the exponent can be interpreted as an index of the irregularity of the signal. However, the algorithms normally used to calculate this spectrum, such as the fast Fourier transform, provide noisy fluctuations of these indexes,

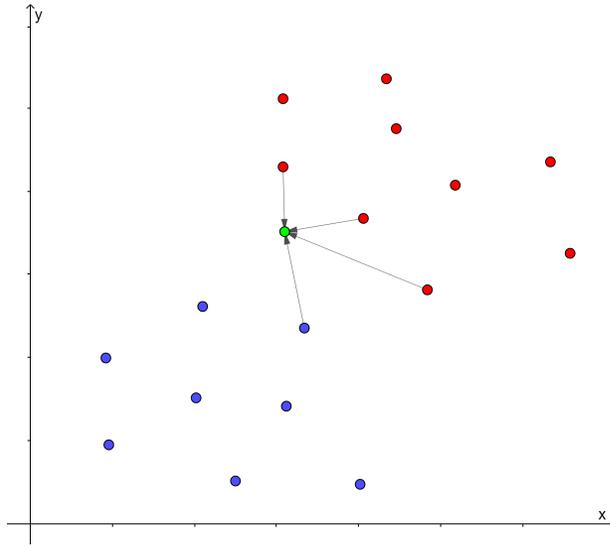


Figure 3: Example with four neighbors, the green dot would be classified in the same group as the red dots as the majority of the neighbors are red.

which is why an average of the power spectrum should be taken over a long interval to obtain a stable index. As the statistical nature of the signal generally varies over small time intervals, it would not be appropriate to use this technique. As an alternative to solve this problem, the fractal dimension can be used and in this work the algorithm introduced by Higuchi in 1988 [15] is used.

In order to calculate the fractal dimension of a time series $X = \{X_1, X_2, \dots, X_N\}$ which is reconstructed in k subseries as follows:

$$X_k^m = \{X_m, X_{m+k}, \dots, X_{m+\lceil \frac{N-m}{k} \rceil k}\} \forall m \in 1, 2, \dots, k.$$

Upon every one of these subseries the length $L_m(k)$ is calculated:

$$L_m(k) = \left(\frac{N-1}{\lceil \frac{N-m}{k} \rceil k} \right) \cdot \sum_{i=1}^{\lceil \frac{N-m}{k} \rceil} |X_{m+ik} - X_{m+(i-1) \cdot k}|,$$

where $\frac{N-1}{\lceil \frac{N-m}{k} \rceil k}$ is a normalization factor of the length of the signal.

Finally, to calculate the fractal dimension D a regression must be performed on the averages of the lengths of the subseries because these will follow a law of the type $L(k) \propto k^{-D}$.

The fractal dimension indicates the "roughness" degree of the signal shape. For continuous time series of the type $f(X) : X \in \mathfrak{R}^1 \rightarrow \mathfrak{R}^1$, if the signal tends to be linear, the index will be close to 1, while if the roughness increases, the signal will seem to be filling more space and the index will be close to 2.

4.2 Lyapunov Exponents

The Lyapunov Exponents [38] [32] of a dynamic system define the divergence or convergence degree of two nearby orbits, meaning, they show the evolutionary feature of the system.

In the case of a n -dimensional system, there are n Lyapunov exponents. These n exponents can be ordered from largest to lowest, obtaining the Lyapunov spectrum of exponents for a particular system. The sign of every exponent gives a notion of where the system tends to evolve (attractors). If they are all negative, it indicates the presence of a fixed point. When they are of the same sign, but with the presence of null values, it indicates a limit cycle. Finally, when the exponents present different signs it indicates chaos in the system.

In particular, the greatest exponent is the most important. From it we can determine whether the system is chaotic or not. In case it is greater than 0, two initially nearby orbits of the attractor diverge rapidly as time progresses, exponentially, showing extreme sensitivity and a large change with respect to the initial conditions, that is, the presence of chaos.

The algorithm selected to compute the Lyapunov exponent in this work is proposed by Kantz in 1994 [17] which provides the maximal LE and requires only the time between samples as input parameter.

4.3 Permutation Entropy

This simple complexity measure of a signal waveform was introduced by Bandt and Pompe in 2002 [6] and has been used in multiple areas of research because of its robust and fast method of calculation and its signal discriminating power. Across the years different additions or changes were made to this original measure giving place to new entropies, some of which are used in this work. The algorithm is as follows:

Given a time series $X = \{X_1, X_2, \dots, X_N\}$, an embedding dimension m and a time delay τ , the Permutation Entropy (PE) of this time series can be calculated. The time series X is sectioned into small sub-sequences called *embedding vectors* of length m .

$$X_k^m = \{X_k, X_{k+\tau}, \dots, X_{k+(m-1)\cdot\tau}\},$$

where $k = 1, 2, \dots, N - (m - 1) \cdot \tau$. The embedding vectors are then assigned a set of the same length consisting of indices in ascending order from 0 to $m - 1$.

$$\pi^m = \{0, 1, \dots, m - 1\}.$$

Now the embedding vectors are sorted in ascending order and this change is reflected in their assigned set of indices. These sets are now the corresponding *patterns* of every embedding vector. There will be $m!$ different *ordinal patterns* in every time series. The next step is to calculate their probability distribution p_i and finally use it to get the PE of the signal according to:

$$PE(X, m, \tau) = \sum_{i=1}^{m!} p_i \cdot \log p_i.$$

When the logarithm base is 2, the PE is measured in bits. the number of publications that use any version of the permutation entropy is grow during the recent times, among a lot of examples, it can be advise to see [3], [5], [18], [12], [37].

4.3.1 Modified Permutation Entropy

One of the variations of the original PE algorithm that is used in this study is the Modified Permutation Entropy (MPE) [10]. The difference lays in the way of calculating the probability distribution. Instead of using:

$$p_i = \frac{c_i}{N - (m - 1)},$$

where c_i stands for the amount of times the pattern π_i^m was found in the time series, for the MPE now that equation changes to:

$$p_i = \frac{c_i}{T},$$

where T is the amount of different patterns found. Then the MPE is calculated in the same way that PE.

This change is based on the fact that missing patterns, called *forbidden patterns*, could increase the obtainable information about time series. This small modification can be applied not only to the PE but also to any other kind of entropy and, possibly, increase their signal discrimination power.

4.3.2 Weighted Permutation Entropy

Another variant of the PE is the *Weighted Permutation Entropy* (WPE) [13]. First the average value X_k^m of every embedding vector is calculated and assigned an object from a previously defined alphabet Π containing $m!$ elements and upon this average value is defined a *weight* w_k .

$$X_k^m = \frac{1}{m} \sum_{k=1}^m X_{k+(m+1) \cdot \tau},$$

$$w_k = \frac{1}{m} \sum_{k=1}^m [X_{k+(k-1) \cdot \tau} - X_k^m]^2.$$

This of course affects the probability distribution, which now is :

$$p_i = \frac{\sum_{k \leq N} 1_{u:\text{type}(u)=\pi_i}(X_k^{m,\tau}) \cdot w_k}{\sum_{k \leq N} 1_{u:\text{type}(u)=\Pi}(X_k^{m,\tau}) \cdot w_k},$$

where the indicator function $1_A(u)$ is defined as $1_A(u) = 1$ in the case of that the embedding vector u belongs to a set A of the alphabet and $1_A(u)=0$ otherwise.

The WPE is computed exactly like the PE. This addition is motivated by the fact that PE overlooks the amplitude between points in a signal. Two subsections of a signal could have different amplitude variation between their points and yet have the same ordinal pattern associated to them. WPE offers with the addition of this new *weight* w_k a possible solution to this problem.

4.3.3 Rényi Entropy

There is another entropy used in this study and that is the *Rényi Entropy* denoted as (RE), which is actually a generalization of the original one. The ordinal patterns and their probabilities are calculated the same way than PE, also a new parameter α is incorporated but the final way of computing this entropy is:

$$RE(X, m, \tau, \alpha) = \frac{1}{1 - \alpha} \log \left(\sum_{i=1}^{m!} p_i^\alpha \right),$$

with $\alpha > 0$ and $\alpha \neq 1$. If $\alpha \rightarrow 1$ then RE approaches PE.

$$\lim_{\alpha \rightarrow 1} RE(X, m, \tau, \alpha) = PE(X, m, \tau)$$

In this work is implemented the Cuesta Frau's [10] modification to compute the Rényi Entropy thus getting a *Modified Rényi Entropy* (MRE).

4.3.4 Entropy with optimal embedding time delay

An optimal time delay is used for every signal as the last feature for the time series on this study. We follow the steps in the work of [14] to find an optimal embedding time delay based on auto-correlation and mutual information of each signal. Next is described how to determine this parameter.

Given a time series there are several methods to determine the optimal embedding time delay τ , some of them use the mutual information between the signal at time t and $t + \tau$, searching in this case the minimum value of this parameter. To achieve this objective, the assignment $[S(t), Q(t)] = [X(t), X(t + \tau)]$ can be done and then calculate the mutual information of this pair of signals as follow:

$$I(Q, S) = H(Q) - H(Q|S) = H(Q) + H(S) - H(S, Q)$$

where:

$$H(S, Q) = - \sum_{i,j} P_{s,q}(S_i, Q_j) \log(P_{s,q}(S_i, Q_j))$$

In this work the time delay τ is constantly increased. The first τ that generates a minimum in the mutual information function is selected. Finally with this τ and an embedding dimension of $m = 6$ the permutation entropy and the weighted permutation entropy are calculated.

4.4 Zero Crossing Rate

The zero crossing rate (ZCR) of a signal is the rate at which the signal changes its sign. That is, the amount of times the values go from positive to negative and viceversa in a given time interval or frame. In this study, the ZCR is measured all across the length of the signal.

ZCR is largely used in voice signal processing and other audio related applications.

Given a series $X = \{X_1, X_2, \dots, X_N\}$, ZCR is calculated as follows:

$$ZCR = \frac{\sum R}{N},$$

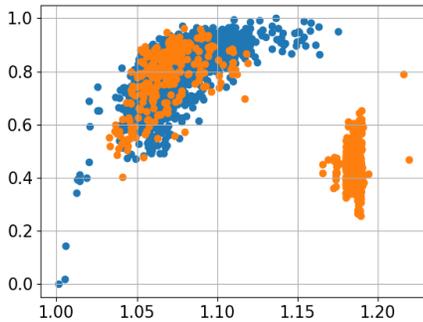
where R stands for the amount of zero crossings in the time series. the application of ZCR in machine learning algorithms is increased in the last twenty years specially in classification of speech signals (see i.e. [9] and [31]).

5 RESULTS

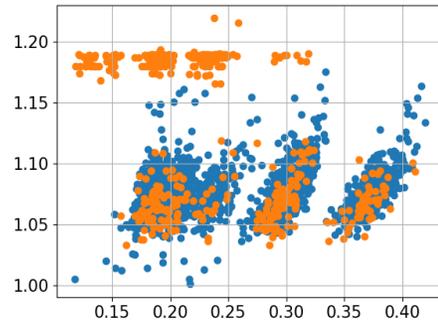
As described in the previous sections, the feature space of the dataset is calculated on the basis of every of the ABR records for both classes, such as for normal hearing subjects and for those with a diagnosed pathology. By varying the feature parameters in precise ranges, the values of these are set to produce the best results in order to achieve the best classification of pathological recordings.

When calculating the Fractal dimension, Higuchi's k values are ranging from 50 to 100. For all entropies, embedding dimension m is selected in a range from 3 to 6 and the embedding time delay τ in the interval between from 1 to 20. For the Rényi entropy, α values go from 0.01 to 0.5 and from 2 to 5.

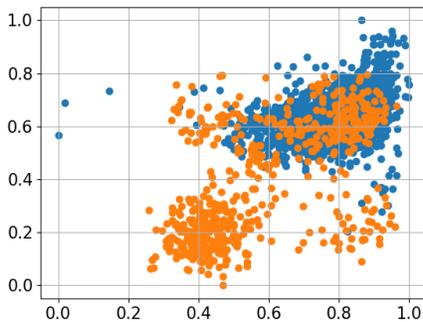
Out of all the calculated features, only eight are selected for the final classification results, which are the fractal dimension with $k = 50$, the Lyapunov Exponent, the ZCR, the modified permutation entropy with $m = 3$ and $\tau = 20$, the modified Rényi entropy with $m = 3$, $\tau = 20$ and $\alpha = 5$, the permutation entropy with $m = 6$ and optimal time delay and the weighted permutation entropy with $m = 4$, $\tau = 20$ and $m = 6$ and an optimal time delay. This selection is based on their scatter plots and histograms, shown in Figure 5 and 4 respectively. Both these plots show some kind of separations between the two groups, which could justify their presence. Moreover, when added to it, their contribution to the classification model improves the quality parameters that are evaluated in this work.



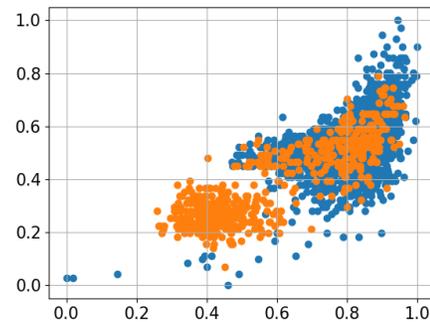
Fractal Dimension vs Permutation Entropy



Lyapunov Exponent vs Fractal Dimension



Modified Entropy vs Rényi Entropy



Permutation Entropy vs Rényi Entropy

Figure 4: Scatter plot of features, (blue: normal, orange: pathological).

In figure 6 the boxplots of the corresponding optimal τ are shown. This embedding time delay is calculated through the methodology proposed in 4.3.4. The mean for normal hearing subjects is $\tau = 47.2$ and for pathological ones is $\tau = 32.6$.

With the previously described dataset, in order to do a comparative study of the classification algorithms, 1000 Montecarlo trials of every type of classifier are performed using a 70% of the complete database randomly selected to train the model and the other 30% to classify it. From the results of this stage of the work, the area under the ROC curve, the F-1 score and the accuracy are calculated to compare the performance of every classification algorithm as can be seen in Figure 7.

Every classifier has a different accuracy. Although all the classifiers have achieved high level performance, the SVM and the Random Forest classifier have the best accuracy with a mean value above 91.8% and the Random Forest classifier has the greater area under the R.O.C. curve with 87.5%. The F1-score indicates the weighted average of the precision and the recall. In the Montecarlo trials results the SVM and the Random Forest classifiers get the best results with

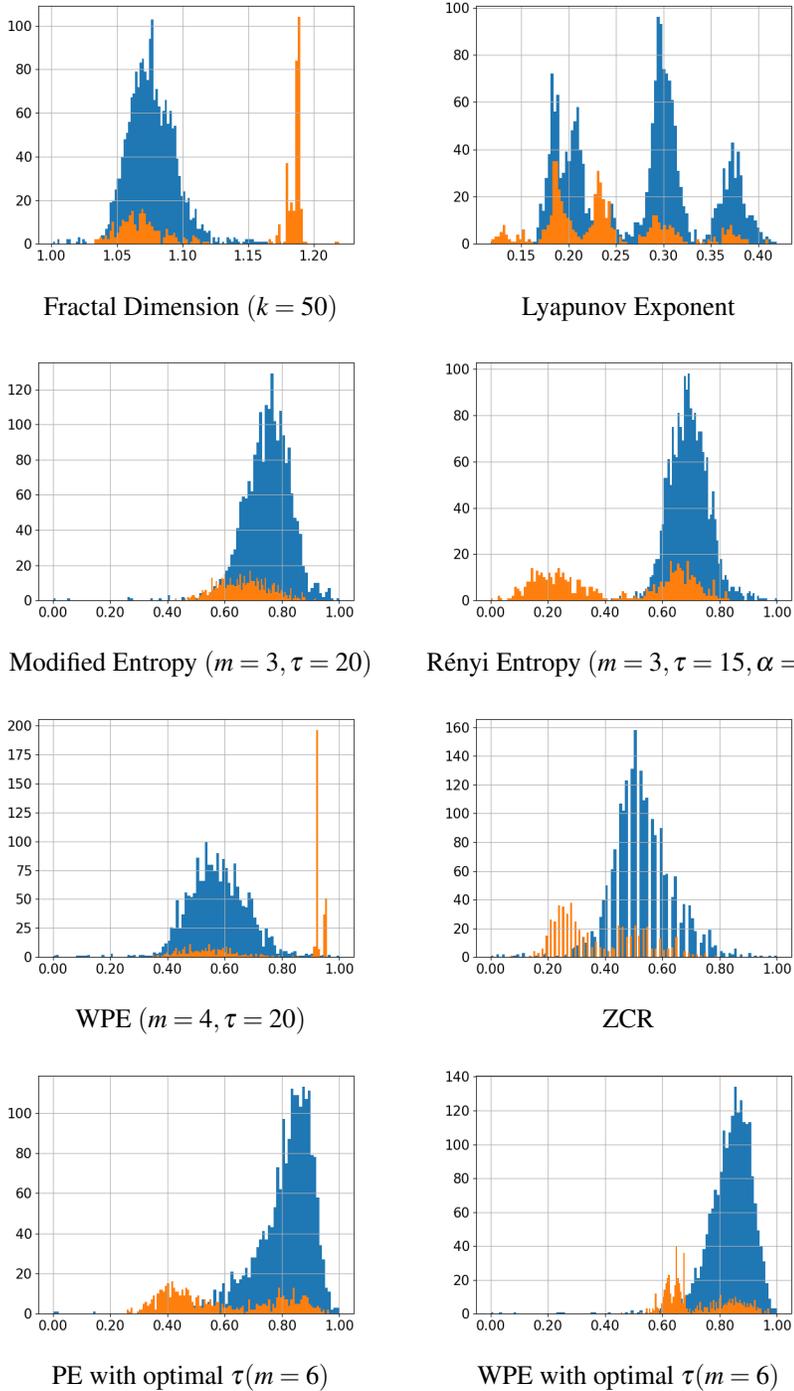


Figure 5: Histograms of features, (blue: normal, orange: pathological).

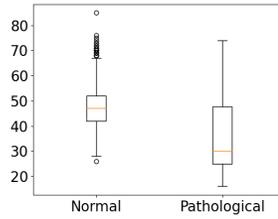


Figure 6: Optimal embedding time delay τ

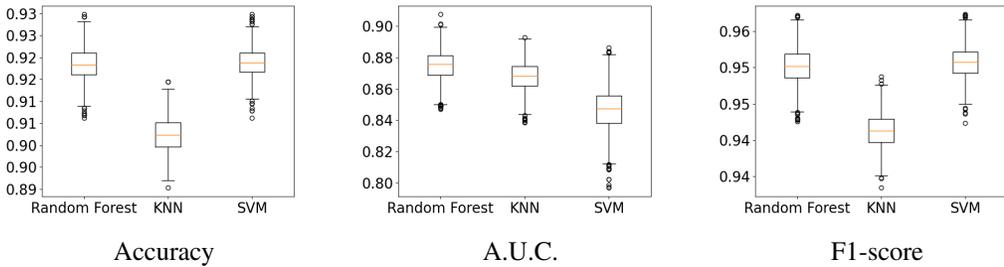


Figure 7: Results after 1000 Montecarlo trials for every algorithm.

overall mean values above 95%. However, the KNN classifier had an equally acceptable result with an f1 score of 94%.

The confusion matrices of one iteration were made and they are shown in the tables 1, 2 and 3 for the Random Forest, KNN and SVM classifier respectively. A low amount of false positives can be observed in these matrices.

A more visual way to see these results is through the ROC curves. Figure 8 shows one of the iterations carried out to calculate the performance of the models.

6 DISCUSSION AND ANALYSIS

From the box plots in Figure 7 it can be seen that SVM and Random Forest have the best results regarding the accuracy and F-1 Score while for the area under the ROC curve the best classifier is Random Forest. However, the three classifiers have satisfactory results in every classification

Table 1: Confusion matrix of one Montecarlo Trial for Random Forest classifier.

Random Forest	Pred. Norm.	Pred. Pat.
True Norm.	604	2
True Pat.	63	113

Table 2: Confusion matrix of one Montecarlo Trial for KNN classifier.

KNN	Pred. Norm.	Pred. Pat.
True Norm.	606	0
True Pat.	81	95

Table 3: Confusion matrix of one Montecarlo Trial for SVM classifier.

SVM	Pred. Norm.	Pred. Pat.
True Norm.	606	0
True Pat.	67	109

quality parameters. They all have an accuracy between 90% and 92%, an area under the ROC curve between 84% and 88% and an F-1 Score between 94% and 95.5%.

Analyzing the scatter plots in Figure 4, it is clear that healthy patients are placed in a concentrated point cloud, whereas pathological records may appear inside this cloud or very far away. From this observation it can be assumed that some signals from pathological patients may seem like healthy but no signal from a healthy patient is outside the mentioned cloud. This means that if one auditory evoked potentials signal from a patient is predicted as pathological, that patient is for sure pathological.

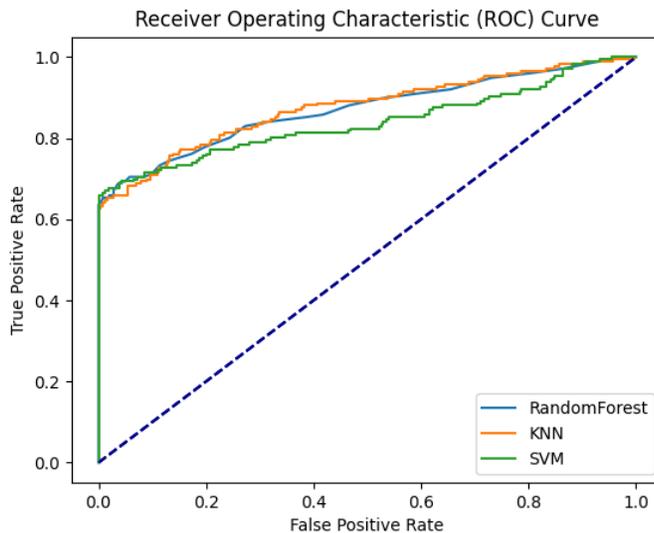


Figure 8: Example of the ROC curve for a Montecarlo trial for every classification algorithm.

The example shown in tables 2 and 3 coincides with the characteristics of the classifying methodologies of the K-Nearest Neighbours and Support Vector Machine, in respect to the prediction capability of the method. In the opposite corner is the Random Forest algorithm which has a low profile to prediction.

An interesting observation is how the ROC curve works in this case. Looking at Figure 8, it is clear how the curve is laying in the left axis, meaning the false positive rate is very low. The probability of misclassifying a normal hearing subject is minimal.

In the three proposed algorithms, the ROC curve obtained is very similar. In them, the cutoff points are very close to each other, which could not easily justify the use of one algorithm over another.

In both Figures 5 and 4 three point clouds can be seen for the Lyapunov exponent. This result is found in all the records analyzed, both normal hearing and pathological, and may be an indication of different behavioral regimes in the response of the auditory system. Not enough evidence has been found that these groupings are related to age since they manifest equally for the same patient.

The need of a classification algorithm emerges from the distribution functions, which many exhibit a remarkable overlapping as seen in Figure 5.

From Figure 6, the calculation of the optimal τ for both groups of signals give as a result lower values for the pathological subjects than the normal hearing subjects. As the embedding time delay is related with the auto correlation of a signal it is valid to express that the correlation of the pathological ABRs has a lower correlation than the normal ABRs. This result could indicate a characteristic of the auditory pathway pathologies that generates an evoked response with lower auto correlation.

7 CONCLUSIONS

Although three training methodologies with different operating principles were used, the results obtained with them were of the same level of quality in the classification.

Given the scatter plots and knowing that SVM classification algorithms work best when a curve that separates the point clouds is easily visualized, it is expected that SVM would have the best results in matter of accuracy and F-1 Score. Random Forest has almost as good results in these indicators and even has the greatest result in AUC whereas SVM has the lowest one. However, the fact that SVM has none false positives in the shown example of the confusion matrix and given the obtained scatter plots, it could be suggested that SVM is slightly better when classifying these signals than the other selected alternatives.

Although similar variables are calculated to be used as features in the dataset, such as the case of the WPE and PE with prefixed and optimal embedding time delay, and the case of the RE, it is clear from the corresponding histograms that the mean values of the distributions are different

as well as their relative locations, which makes them additional variables that are incorporated in the feature space contributing to improve the result of the classification stage.

The satisfactory results obtained in the present work can be interpreted as a first step in incorporating different algorithms to classifying signals for the detection of ABRs of subjects with possible pathologies and thus contribute to the analysis and subsequent objective diagnosis.

Acknowledgments

The authors wish to express their gratitude to the student scholarship programme of the Facultad Regional Buenos Aires Universidad Tecnológica Nacional and its Secretary of University Affairs and to the Secretary of Science and Technology and the Scholarship programme of the Secretary of Science, Technology and Graduate Studies of the Universidad Tecnológica Nacional. This work was developed within the framework of the projects PID UTN 4729 and PID UTN 8120.

REFERENCES

- [1] N. Acır, Ö. Özdamar & C. Güzeliş. Automatic classification of auditory brainstem responses using SVM-based feature selection algorithm for threshold detection. *Engineering Applications of Artificial Intelligence*, **19**(2) (2006), 209–218.
- [2] R. Al Osman & H. Al Osman. On the use of machine learning for classifying auditory brainstem responses: A Scoping Review. *IEEE Access*, (2021).
- [3] J. Amigó. “Permutation complexity in dynamical systems: ordinal patterns, permutation entropy and all that”. Springer Science & Business Media (2010).
- [4] M. Baldiviezo, C. Bontempo, J. Barbería, Y. Corsaro, F. Fernandez Biancardi, M. Hermando, A. Paglia, M. Rodriguez & W. Legnani. Application of entropic measures in the study of auditory evoked potentials for the detection of pathological patients. In “Proceedings of VIII MACI 2021, Volume 8”. Asociación Argentina de Matemática Aplicada, Computacional e Industrial (2021), p. 633–636.
- [5] C. Bandt. A new kind of permutation entropy used to classify sleep stages from invisible EEG microstructure. *Entropy*, **19**(5) (2017), 197.
- [6] C. Bandt & B. Pompe. Permutation entropy: a natural complexity measure for time series. *Physical review letters*, (2002).
- [7] B. Boashash. “Time-frequency signal analysis and processing: a comprehensive reference”. Academic press (2015).
- [8] L. Breiman. Random forests. *Machine learning*, **45**(1) (2001), 5–32.
- [9] N. Chauhan, T. Isshiki & D. Li. Speaker recognition using LPC, MFCC, ZCR features with ANN and SVM classifier for large input database. In “2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS)”. IEEE (2019), p. 130–133.

- [10] D. Cuesta-Frau. Using the Information Provided by Forbidden Ordinal Patterns in Permutation Entropy to Reinforce Time Series Discrimination Capabilities. *Entropy*, (2020).
- [11] J.R. Daube & D.I. Rubin. Needle electromyography. *Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine*, **39**(2) (2009), 244–270.
- [12] F.H.A. de Araujo, L. Bejan, O.A. Rosso & T. Stosic. Permutation entropy and statistical complexity analysis of Brazilian agricultural commodities. *Entropy*, **21**(12) (2019), 1220.
- [13] B. Fadlallah, B. Chen, K. A. & J. Príncipe. Weighted-permutation entropy: A complexity measure for time series incorporating amplitude information. *Physical Review E*, **87** (2013).
- [14] A.M. Fraser & H.L. Swinney. Independent coordinates for strange attractors from mutual information. *Physical review A*, **33**(2) (1986), 1134.
- [15] T. Higuchi. Approach to an irregular time series on the basis of the fractal theory. *Physica D: Nonlinear Phenomena*, **31**(2) (1988), 277–283.
- [16] D.L. Jewett & J.S. Williston. Auditory-evoked far fields averaged from the scalp of humans. *Brain*, **94**(4) (1971), 681–696.
- [17] H. Kantz. A robust method to estimate the maximal Lyapunov exponent of a time series. *Physics letters A*, **185**(1) (1994), 77–87.
- [18] K. Keller, T. Mangold, I. Stolz & J. Werner. Permutation entropy: New ideas and challenges. *Entropy*, **19**(3) (2017), 134.
- [19] N.S. Kwak, K.R. Müller & S.W. Lee. A convolutional neural network for steady state visual evoked potential classification under ambulatory environment. *PloS one*, **12**(2) (2017), e0172578.
- [20] G. Lee, R. Gommers, F. Waselewski, K. Wohlfahrt & A. O’Leary. PyWavelets: A Python package for wavelet analysis. *Journal of Open Source Software*, **4**(36) (2019), 1237.
- [21] S. Mallat. A wavelet tour of signal processing (1999).
- [22] D. Mao, H. Innes-Brown, M.A. Petoe, C.M. McKay & Y.T. Wong. Spectral features of cortical auditory evoked potentials inform hearing threshold and intensity percepts in acoustic and electric hearing. *Journal of neural engineering*, **18**(4) (2021), 046078.
- [23] R.M. McKearney & R.C. MacKinnon. Objective auditory brainstem response classification using machine learning. *International journal of audiology*, **58**(4) (2019), 224–230.
- [24] W.S. Noble. What is a support vector machine? *Nature biotechnology*, **24**(12) (2006), 1565–1567.
- [25] M. Paulraj, K. Subramaniam, S.B. Yaccob, A.H.B. Adom & C. Hema. A Machine learning approach for distinguishing hearing perception level using auditory evoked potentials. In “2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES)”. IEEE (2014), p. 991–996.
- [26] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot & E. Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, **12** (2011), 2825–2830.

- [27] L.E. Peterson. K-nearest neighbor. *Scholarpedia*, **4**(2) (2009), 1883.
- [28] R.A. Prayson & J.R. Goldblum. “Neuropathology”. Elsevier Churchill-Livingstone (2005).
- [29] M. Ravan, J.P. Reilly, L.J. Trainor & A. Khodayari-Rostamabad. A machine learning approach for distinguishing age of infants using auditory evoked potentials. *Clinical neurophysiology*, **122**(11) (2011), 2139–2150.
- [30] N.H. Shah, A. Milstein & S.C. Bagley. Making machine learning models clinically useful. *Jama*, **322**(14) (2019), 1351–1352.
- [31] G. Sharma, K. Umopathy & S. Krishnan. Trends in audio signal feature extraction methods. *Applied Acoustics*, **158** (2020), 107020.
- [32] B.K. Shivamoggi. “Nonlinear dynamics and chaotic phenomena: An introduction”, volume 103. Springer (2014).
- [33] I. Silva & M. Epstein. Estimating loudness growth from tone-burst evoked responses. *The Journal of the Acoustical Society of America*, **127**(6) (2010), 3629–3642. doi:10.1121/1.3397457.
- [34] I. Silva & M. Epstein. Objective estimation of loudness growth in hearing-impaired listeners. *The Journal of the Acoustical Society of America*, **131**(1) (2012), 353–362. doi:10.1121/1.3666024.
- [35] J. Sohn, I.Y. Jung, Y. Ku & Y. Kim. Machine-Learning-Based Rehabilitation Prognosis Prediction in Patients with Ischemic Stroke Using Brainstem Auditory Evoked Potential. *Diagnostics*, **11**(4) (2021), 673.
- [36] C. Torrence & G.P. Compo. A practical guide to wavelet analysis. *Bulletin of the American Meteorological society*, **79**(1) (1998), 61–78.
- [37] S.J. Watt & A. Politi. Permutation entropy revisited. *Chaos, Solitons & Fractals*, **120** (2019), 95–99.
- [38] A. Wolf, J.B. Swift, H.L. Swinney & J.A. Vastano. Determining Lyapunov exponents from a time series. *Physica D: nonlinear phenomena*, **16**(3) (1985), 285–317.
- [39] J. Yperman, T. Becker, D. Valkenburg, V. Popescu, N. Hellings, B. Van Wijmeersch & L.M. Peeters. Machine learning analysis of motor evoked potential time series to predict disability progression in multiple sclerosis. *BMC neurology*, **20**(1) (2020), 1–15.

