

Epileptic Seizure Prediction based on Ratio and Differential Linear Univariate Features

Jalil Rasekhi, Mohammad Reza Karami Mollaei, Mojtaba Bandarabadi¹, César A. Teixeira¹, António Dourado¹

Department of Biomedical Engineering, Faculty of Engineering, Babol Noshirvani University of Technology, Babol, Iran, ¹Department of Informatics Engineering, CISUC/DEI, Center for Informatics and Systems of the University of Coimbra, Polo II 3030-290, Coimbra, Portugal

Submission: 28-08-2013

Accepted: 02-01-2015

ABSTRACT

Bivariate features, obtained from multichannel electroencephalogram recordings, quantify the relation between different brain regions. Studies based on bivariate features have shown optimistic results for tackling epileptic seizure prediction problem in patients suffering from refractory epilepsy. A new bivariate approach using univariate features is proposed here. Differences and ratios of 22 linear univariate features were calculated using pairwise combination of 6 electroencephalograms channels, to create 330 differential, and 330 relative features. The feature subsets were classified using support vector machines separately, as one of the two classes of preictal and nonpreictal. Furthermore, minimum Redundancy Maximum Relevance feature reduction method is employed to improve the predictions and reduce the number of false alarms. The studies were carried out on features obtained from 10 patients. For reduced subset of 30 features and using differential approach, the seizures were on average predicted in 60.9% of the cases (28 out of 46 in 737.9 h of test data), with a low false prediction rate of 0.11 h⁻¹. Results of bivariate approaches were compared with those achieved from original linear univariate features, extracted from 6 channels. The advantage of proposed bivariate features is the smaller number of false predictions in comparison to the original 22 univariate features. In addition, reduction in feature dimension could provide a less complex and the more cost-effective algorithm. Results indicate that applying machine learning methods on a multidimensional feature space resulting from relative/differential pairwise combination of 22 univariate features could predict seizure onsets with high performance.

Key words: Classification, epilepsy, epileptic seizure prediction, features selection, support vector machine

INTRODUCTION

Epilepsy is described as the seizures occurring randomly and triggered by the intensive, unexplained, and coherent neuronal activities in the brain.^[1] About 60 million people worldwide are currently suffering from a known form of epilepsy. Epileptic seizures affect seriously the living conditions of these patients namely by exposing them to danger and social exclusion. Medication is usually employed to improve the living conditions of the patients by lowering the brain activity levels. However, around 30% of these patients should live with refractile epilepsy resistant to any drug.^[2] Surgical treatment is also suggested for a limited number of such cases, but bears high risks. Most patients would feel quite satisfied by receiving warnings on pending seizures, enabling them to prepare for the onset and the resulting unexpected harmful situation. Epileptic seizure prediction tries to fulfill this demand. For more than 50% of the epileptic cases, a certain cause cannot be identified,^[3] but anything damaging normal brain tissue can develop

seizures. Young children are easily affected with epilepsy, by birth-related issues, inheritance, usual infections such as meningitis and sometimes by uncontrolled fevers. In grownups, the accidents may cause seizures by harming brain tissue. Heart attacks and traumas are main triggers of the seizures in old people.^[3]

Although the problem of epileptic seizure prediction has attracted the minds of many researchers over the years,^[4-15] promising results of most groups have been challenged and questioned by other researchers.^[16-19]

In the literature, linear/nonlinear features have been investigated in univariate (extracted from one channel), bivariate (extracted from two channels simultaneously), and multivariate (extracted from three or more channels) fashions. Most seizure prediction methods extract some features from a time moving window of electroencephalogram (EEG) signals and study their behavior during the preictal time compared with the other times. The

Address for correspondence:

Jalil Rasekhi, Department of Biomedical Engineering, Faculty of Engineering, Babol Noshirvani University of Technology, Babol, Iran.
E-mail: jraskhi@gmail.com

linear univariate features of statistical moments,^[19] spectral power,^[20-22] Hjorth parameters of mobility and complexity,^[19] decorrelation time,^[19] wavelet coefficients^[23,24] have been investigated in seizure prediction studies.

Machine learning (ML) algorithms using a set of features could improve the results compared to the studies based on only a single feature.^[6,25,26] In our previous work^[26] we employed a combination of 22 linear univariate features extracted from 6 recording channels and classified the resulting feature space using support vector machine (SVM). The results were somehow promising.

In order to treat the seizure prediction as a classification problem, four distinct states are defined for an epileptogenic brain: Interictal, preictal, ictal, and postictal. Preictal state is the period during which the brain is evolving toward a new seizure, and the one possessing high processing values for seizure prediction. By detecting the preictal state, the solution for seizure prediction problem will be easily accomplished. Thus, the seizure prediction can be considered as a two class preictal/nonpreictal problem. Preictal time length is the prerequisite parameter to prepare the initial two classes data pool. The proper choice of optimal preictal time is very essential for achieving good prediction results. Four preictal times of 10 min, 20 min, 30 min, and 40 min are used in this study to address this issue. The results, in terms of prediction sensitivity (SS) and false prediction rate (FPR), are compared with those achieved from the original univariate features to indicate the performance of the proposed bivariate approach. The proposed method also outperforms its equivalent analytical random predictor model.

MATERIALS AND METHODS

Subjects

Long-term continuous multichannel EEG recordings of 10 patients with partial refractory epilepsy were investigated

and prepared to test the proposed methods. The patients' ages ranged from 15 to 57 years and were selected from European database on epilepsy.^[16,27] The data were supplied by two epilepsy centers from France (Pitié-Salpêtrière Hospital of Paris) and Portugal (University Hospitals of Coimbra), and included eight Scalp and two intracranial recordings. The scalp recordings were made using a 10–20 standard system. For more details refer to Table 1.

Seizure Prediction

The proposed method is built upon five blocks: (1) a univariate feature extraction stage from 6 EEG channels using a nonoverlapping 5-s moving window, (2) a channel fusion block in order to prepare differential or relative features, (3) a feature preprocessing block to both normalize and smooth the extracted features (4) a SVM classifier to discriminate different classes of the incoming EEG signal, and (5) a regularization algorithm in order to make alarms and to reduce the number of false predictions. The block diagram of the seizure predictor is depicted in Figure 1.

Feature Extraction

The EEG data of 6 selected channels, three located on the focal area and three far from the focus, were first segmented to nonoverlapping 5 s windows, and then processed by an infinite impulse response forward-backward Butterworth 50 Hz notch filter to eliminate strong 50 Hz distortions caused by AC power supply. By considering a linear model for short-term brain activities, 22 features were extracted, encapsulating phase/frequency and amplitude information of the EEG signal, so-called linear features. In order to obtain differential and relative feature sets, all possible pairwise combinations of these 6 channels were utilized. Employed univariate features include time domain, frequency domain, and time/frequency domain features. Regarding theoretical limitations on segmentation windows size, 5-s is neither much long nor much short and has been employed in various

Table 1: Information for the 10 studied patients

ID	Sex	Age	Onset age	Rec. type	Rec. time (h)	Number of seizure	Seizure type	Localize	Samp. rate (Hz)	Hospital
1	Male	57	6	Scalp	110.2	8	CP (7), UC (1)	Right T	512	Coimbra
2	Male	20	-	Scalp	89.8	13	CP (12), UC (1)	Left T	512	Coimbra
3	Male	15	3	Scalp	113.5	7	CP (6), UC (1)	Right T	512	Coimbra
4	Male	16	12	Scalp	91.9	7	SG (1), UC (6)	Left T	512	Coimbra
5	Male	35	1	Scalp	141.5	4	CP (3), UC (1)	Right T	512	Paris
6	Male	41	19	Scalp	140.6	6	CP (4), SP (1), UC (1)	Right T	400	Paris
7	Female	21	1	Scalp	252.6	9	UC (9)	Right T	400	Paris
8	Male	46	19	Scalp	91.7	7	UC (6), CP (1)	Right T	400	Paris
9	Female	23	15	Invasive	145.1	13	CP (13)	Right T	400	Paris
10	Male	36	1	Invasive	211.7	12	UC (12)	Right F	400	Paris
A/T	-	31	8.5	-	1388	86	-	-	-	-

Localize – Localization of seizures; T – Temporal lobe; F – Frontal lobe; P – Parietal lobe (between frontal and occipital); Seizure type – Type of the clinical seizures; SP: Simple partial; CP – Complex partial; SG – Secondarily generalized; UC – Unclassified

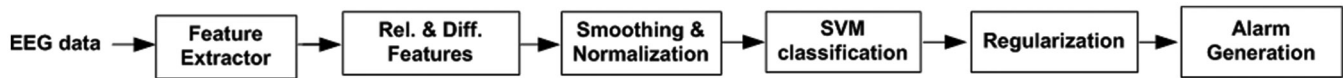


Figure 1: Block diagram for employed support vector machine seizure predictor, from EEG acquisition to alarm generation

research works successfully. Moreover, 5-s segmentation has been established as a basic part of EPILEPSIAE project. Similar lengths of segmentation windows sizes can also be frequently noticed in many seizure prediction reports.^[22,26,28]

Spectral power features

Power spectral density (PSD) represents the distribution of strength of signal components across different frequencies. EEG signal processing specialists commonly divide the EEG spectrum into five frequency bands δ (≤ 4 Hz), θ (4–8 Hz), α (8–12 Hz), β (13–30 Hz), and γ (≥ 30 Hz), also being adopted here. During particular brain tasks, the EEG activities increase or decrease in some particular frequency bands. Mormann *et al.*^[19] found that the spectral power of Delta band decreases during the preictal period while the spectral power of other bands increase. Using Welch method,^[29] the PSD of EEG signal was first calculated, and then by summing up the frequency components inside the desired frequency bands, the spectral power of all sub-bands was calculated. The spectral power of sub-band features was normalized by the total spectral power, in order to generate a more robust measure to the variations in the patient's daily life, for revealing the preictal state.^[30] As normalized power features are less dependent on the total power, they can be considered better measures for comparison.

Statistical moments

Statistical moments provide information about the amplitude distribution of time series. Four well-known statistical moments are “mean,” “variance,” “skewness,” and “kurtosis.” The location and span of the amplitude distribution of every series are related to the mean and variance of that series respectively, whereas the shape information is represented by skewness and kurtosis.^[17] The significant changes of these statistical moments during preictal state have been reported in several studies.^[7,19] Reports by Mormann *et al.*^[19] revealed a decrease in variance coinciding with an increase in kurtosis, during the preictal period.

Hjorth parameter

Activity, mobility, and complexity are 3 time domain parameters introduced by Hjorth^[31] to describe EEG signals quantitatively. Employing Hjorth parameters, Mormann *et al.*^[19] have discovered significant increases during preictal in mobility and complexity. Mobility is defined as the root mean square (RMS) of slopes of the EEG signal inside a moving window, divided by the RMS of amplitudes in the same window. It can provide an estimate of the mean frequency value. Complexity gives a measure

of the RMS of the rate of slope changes with reference to an ideal possible curve and provides an estimate for signal's bandwidth.

Long-term energy

The feature also known as accumulate energy has been used in seizure prediction studies. A possible theory for explaining the cause of temporal lobe epilepsy is that the seizures are developed by hours-long chains of events in the brain. This shall provide a justification for long-term monitoring of energy to predict epileptic seizures. Litt *et al.*^[32] has reported that the bursts of long-term energy increase approaching the seizure onsets. The feature is evaluated in three steps. First, the instantaneous energy of EEG samples is found by a simple squaring of the EEG sequence:

$$E_i(n) = x(n)^2 \quad (1)$$

a length-M averaging window is then applied on the resulting $E_i(n)$ sequence:

$$E_k = \frac{1}{M} \sum_{n=1}^M E_i(n) = \frac{1}{M} \sum_{n=1}^M x(n)^2 \quad (2)$$

finally, the accumulated energy is obtained by averaging of the E_k sequence using another length-N moving window:

$$AE_m = \frac{1}{N} \sum_{k=1}^N E_k + AE_{m-1} \quad (3)$$

where AE_m is the m -th value of the accumulated energy. The whole process is analogous to the integration over the energy values of original EEG signal.^[32]

Autoregressive error

Autoregressive model (AR) is a linear method used to predict the values of a linear time series at a given time, by employing weighted sum of a number of previous values plus noise. Apparently, the signal has to be stationary. However, for nonstationary signals like the EEG data, this technique can also be applied by taking very short periods of signal demonstrating quasi-stationary properties. Chisci *et al.*^[33] reported that AR-modeling of epileptogenic EEG data could locate the preictal changes. We estimated the 10th-order AR model of EEG time series using Burg's method.^[34] The mean square error (MSE) value between the output of the resulting AR model and the original EEG series is then used as one of the features for the prediction of seizure onsets.

Decorrelation time

“Decorrelation time” is defined as the time at which the first zero-crossing of the autocorrelation function occurs.^[35]

It is a useful measure of detecting stationarity in time series. Extracting decorrelation time from epileptogenic EEG signals, Mormann *et al.*^[19] could discriminate preictal and interictal periods from each other, by looking for a decrease in the measure.

Spectral edge frequency/power

Spectral power of EEG signals is mainly confined to the frequencies below 40 Hz. This characteristic can be quantified by means of two parameters: Spectral edge frequency and spectral edge power. Spectral edge frequency is the frequency below which $x\%$ of the whole energy of the signal is contained. x value of 50 was considered here, describing the minimum frequency up to which 50% of the overall power of the 0–40 Hz band is contained. Spectral edge power is the power covered under the spectral edge frequency.^[36]

Wavelet coefficients

Thanks to their multi-resolution nature, wavelet transforms play an important role in the field of processing, specifically for the nonstationary signals like EEG. Wavelet analysis decomposes the original signal into a sum of scaled and shifted versions of a specific signal known as a mother wavelet. Mother wavelet is a function specifically synthesized to bear particular mathematical properties. Bandarabadi *et al.*^[24] recently developed Seizure-specific wavelets for seizure prediction with promising results. However, designing patient-specific wavelets is a very time-consuming task, and beyond the objectives intended for this work. Direito *et al.*^[23] by studying 22 univariate features, and employing feature reduction techniques, emphasized the performance of wavelet coefficients in comparison to other features. The Daubechies-4 (db4) mother wavelet possesses good localization properties for EEG signals both in time and frequency domains.^[37] A five-level decomposition using db4 mother wavelet was applied on the segmented EEG signals and energy of the resulting wavelet coefficients were used as features.

Ratio and Differential Features

Quantification of generalized and localized perspectives of the brain activities is very important for discovering the abnormalities in the state of the brain which may possess the possibility of leading to seizures in patients suffering from epilepsy. The univariate features can only quantify the state of a localized part of the brain where the corresponding electrode is situated to record the electrical signals. Contrarily, the comparison of the same feature from different channels, will quantify the inter-relationships between two or more regions, thus providing a generalized view toward brain activities.

Relative and differential approaches were applied here, to quantify the inter-relationship between different channels.

The relative and differential features are obtained by dividing and subtracting the features of one channel to/from features of another channel, respectively. Considering 22 features in the feature vector F for each of six recording channels, relative and differential features can, therefore, be written as (4) and (5),

$$f_dif_k(i, j) = f_k(i) - f_k(j), k = 1, 2, \dots, 22; i, j = 1, 2, \dots, 6; i \neq j \quad (4)$$

$$f_rel_k(i, j) = \frac{f_k(i)}{f_k(j)}, k = 1, 2, \dots, 22; i, j = 1, 2, \dots, 6; i \neq j \quad (5)$$

where $f_k(i)$ stands for the k -th feature from the i -th channel.

Feature Preprocessing

The MA methods such as classification or clustering have been widely used for the combination of features. These methods are generally sensitive to the range, and span of the features,^[38] and feature processing has, therefore, become a nonseparable part of MA methods. In an earlier study,^[26] it was shown that in the context of seizure prediction problem, a combination of smoothing by a moving average window and normalization by the maximum feature value could provide better results. Smoothing is obtained by feeding the features into a 1-min length averaging window which covers twelve 5-s epochs:

$$\bar{a}_k = \frac{1}{m+1} \sum_{j=0}^m a_{k-j}, m = 11 \quad (6)$$

where \bar{a}_k and stand for the smoothed and unsmoothed feature, respectively. The normalization is also achieved by dividing by maximum value of that feature:

$$x_{norm} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (7)$$

x representing the feature sequence being normalized and x_{norm} indicating its normalized version. It should be noted that since the EEG patterns change slowly, for example 30 s, the choice of 1-min length preprocessing window (12 consecutive feature samples) should be reasonable enough to cover these low frequency trends as well.

Feature Labeling

As the proposed method relies on a supervised ML algorithm, all features are labeled and also divided into the training and test sets. Neurologists categorize the epileptogenic EEG recordings into four distinct classes: Interictal, preictal, ictal, and postictal, having different patterns [Figure 2]. The interictal state is the interval beginning right after one postictal state of a seizure and ending before the preictal state of the next seizure. The preictal state precedes the

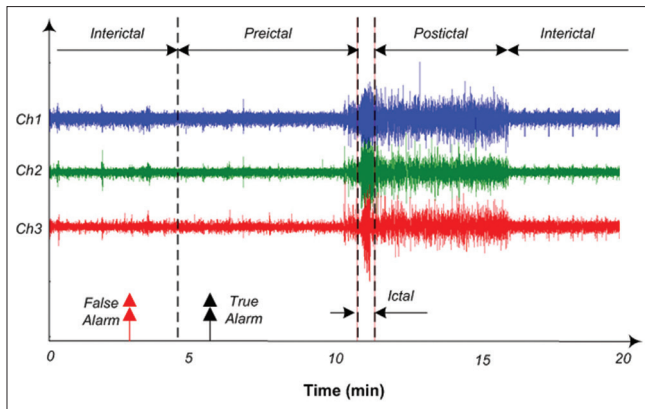


Figure 2: Epileptic brain states; interictal, preictal, ictal, and postictal

seizure onset and can be something from several seconds up to several hours. Ictal is the time period in which seizure happens, which lasts from few seconds to 2-3 min, and is initially marked by experts. In fact, the seizure prediction algorithms are mainly focused on distinguishing the preictal period-which precedes the seizure-as much as possible from the three other states. Thus, the seizure prediction was considered here as a two-class problem: Distinguishing between preictal and nonpreictal states.

Since the selection of the training samples is very important in supervised learning methods, the preictal period should be selected carefully, so as to cover the information and patterns best reflecting the preictal activities. The optimal choice of preictal time is, therefore, an essential part of any seizure prediction algorithm. To tackle this issue, we have used four different preictal times of 10 min, 20 min, 30 min, and 40 min. The resulting performances achieved using these four preictal times indicate which preictal time best suits a specific patient.

After labeling all samples, they were partitioned into the training and test sets. The samples belonging to the first three seizures and their corresponding preictal and nonpreictal recordings were used for training, and the rest of data were used for test. The features were subsequently fed to the SVM classifier so as to discriminate the resulting sequence into predefined preictal/nonpreictal states.

Feature Reduction

The extracted bivariate features produce a huge feature space with dimension of 330. Classifying such a high dimensional feature space would be quite challenging and can bear a high computational cost. Feature reduction is the act of selecting a subset of relevant features, which have more discriminative information for separating the classes. The well-known minimum redundancy maximum relevance (mRMR) feature selection method^[39] was applied on the labeled features in order to rank them. mRMR method uses the mutual information to find the relevance

and redundancy of features and to rank the features based on these two parameters.^[39]

Classification

Support vector machine classifiers have shown good performance in classifying data in nonlinearly separable feature spaces such as in seizure prediction problem.^[21,40] The patient-specific SVM classifier with Gaussian kernel is widely exercised in the literature as a powerful choice and is also used here:

$$K(x, y) = \exp\left(-\frac{|x - y|^2}{2\sigma^2}\right) \quad (8)$$

where x and y are feature vectors in the SVM input space, and σ is the scale parameter which controls the spread of the kernel. Gaussian-based SVM classification uses two tunable parameters of sigma and soft margin (C) that need to be optimized. To optimize these parameters, a grid search was used.

Regularization

A known issue with signal processing problems using hard classification is the lack of confidence usually existing about the proper choice of the threshold value and thus the validity of classification. This becomes even worst with SVM-based seizure prediction using sample feature lengths in the range of seconds, because of the highly changing fluctuations in the sequence of extracted feature samples from short windowed EEGs. In such cases, false alarms can be raised and lowered alternatively in a short period, and just by chance. Regularization methods can be used to improve the classification performance by monitoring several classification outputs in a row. The regularization method employed here takes a sliding window having a size related to the preictal time value and then computes a measure called firing power.^[41,42] The measure is defined as:

$$fp[n] = \frac{\sum_{k=n-\tau}^n o[k]}{\tau} \quad (9)$$

$fp[n]$ has a value between 0 and 1, τ is the number of samples inside a time-length equal to the preictal time, and $o[k]$ is the output of the classifier being 1 and 0 for preictal and nonpreictal states, respectively. For example, suppose the features are extracted using epochs of length 5-s, and preictal time is selected as 10 min. Thus, a time length of 10 min involves $10 \times 60/5$ sample features, and the firing power at every instance is calculated using the past $\tau = 120$ samples. To generate an alarm, the firing power is compared with a threshold value such as 0.5 under two restrictions: (1) after the first positive threshold pass, an alarm is raised, but further alarm generation is blocked for τ epochs: Alarms should not be raised during an ongoing

seizure warning. (2) As soon as the first restriction is eased, alarm generation remains blocked until the firing power returns back below the threshold value: Letting some time for the residual effects of the previous seizure in Eq. (9) to die out.

The final stage of alarms generation using firing power method is illustrated in Figure 3. As shown in this figure, by upward passing of regularization output (blue line) across the threshold value (0.5), while fulfilling two above-mentioned conditions, an alarm is generated. The preictal period is highlighted for reference, indicating that alarms outside this window will be considered as a false alarm.

Statistical Validation

The results obtained from the proposed method should simply exceed the prediction performance of a random predictor having no clue about the input EEG data. Here, we employ the analytical random predictor introduced by Schelter *et al.*^[43] According to their model, the probability of randomly predicting at least n out of N seizures by at least

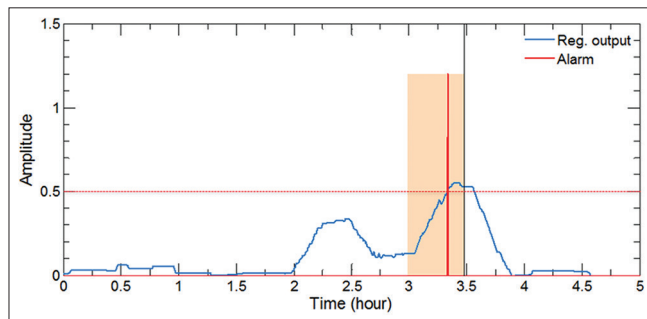


Figure 3: Generation of alarms using firing power method. Highlighted area indicates the preictal period, while blue and red lines represent the regularization output and the generated alarm, respectively. To obtain correct predictions, alarms should fall inside the preictal period of the pending seizure

one of the d different combinations of channels can be formulated as:

$$P_d \{n, N; P\} = 1 - \left[1 - \sum_{j \geq n} \binom{N}{j} P^j (1-P)^{N-j} \right]^d \quad (10)$$

And the critical sensitivity of the analytical random predictor for a significance level of α is written as (11),

$$S_{RP} = \operatorname{argmax}_n \{P_d \{n, N; P\} > \alpha\} / N \times 100\% \quad (11)$$

In this regard, any sensitivity above S_{RP} can be regarded as statistically significant.

RESULTS

Prediction performance of the proposed methods was evaluated on the test recordings of 10 patients with partial epilepsy. The patient-specific SVM classifiers were trained using the training set. The performance of the methods was evaluated using sensitivity (SS) and false positive rate (FPR) of the raised alarms. SS is defined as the ratio of correctly predicted seizures on the actual number of seizures, and FPR is the number of false alarms generated per hour. Table 2 presents the results of SS% and FPR for 10 subjects obtained using all three approaches: Relative (bivariate), differential (bivariate), and the original univariate features. Furthermore, the results of the reduced feature space are presented in Table 3.

By considering a trade-off between SS and FPR values, the reduced feature space of differential features could provide the best results, with an SS of 60.9%, and an FPR of 0.11 h^{-1} . Meanwhile, although the highest average SS = 63% was achieved from the ratio method with reduced feature set, but it provides twice the number of false alarms (FPR = 0.22 h^{-1}) on average, with respect to the reduced/differential method. Overall, both reduced/ratio

Table 2: Results of 10 subjects using all 22 linear features, as compared to the results from analytical random predictor ($d=15$, $\alpha=0.05$)

Patient ID	Test rec. time (h)	Number of test seizures	Univariate (all features)				Differential (all features)				Ratio (all features)			
			SS %	FPR (h^{-1})	PP (min)	SS _{RP}	SS %	FPR (h^{-1})	PP (min)	SS _{RP}	SS %	FPR (h^{-1})	PP (min)	SS _{RP}
1	64.5	5	40	0.09	20	40	40	0.16	40	60	40	0.05	20	20
2	48.5	6	50	0.02	40	16.67	33.3	0.00	40	0	50	0.02	40	16.67
3	24	4	50	0.16	40	75	50	0.04	30	25	75	0.33	40	75
4	83.6	4	75	0.26	10	50	75	0.23	30	75	75	0.29	30	75
5	65.1	1	100	0.08	20	100	100	0.06	10	100	100	0.08	20	100
6	40.8	2	100	0.46	40	100	50	0.15	10	50	100	0.46	40	100
7	139.6	6	100	0.12	20	33.33	100	0.14	40	50	100	0.12	20	33.33
8	58.8	2	50	0.03	40	50	100	0.13	30	100	50	0.03	40	50
Av./T	524.9	30	66.7	0.14	28.7	46.66	63.3	0.13	28.7	46.66	70	0.15	31.2	46.66
9	52.7	7	57.1	0.20	20	42.86	57.1	0.14	20	28.57	57.1	0.20	20	42.86
10	160.3	9	44.4	0.28	40	55.56	33.3	0.05	20	22.22	33.3	0.12	10	22.22
Av./T	213	16	50	0.26	30	50.01	43.7	0.07	20	24.99	43.7	0.14	15	31.25
Av./T	737.9	46	60.9	0.18	29	47.83	56.5	0.11	27	39.13	60.9	0.14	28	41.31

FPR – False prediction rate; RP – Random predictor; PP – Preictal period; SS – Sensitivity

Table 3: Results of 10 subjects using a reduced set of 22 linear features, as compared to the results from analytical random predictor ($d=15$, $\alpha=0.05$)

Patient ID	Test rec. time (h)	Number of test seizures	Differential (reduced 30 features)				Ratio (reduced 30 features)			
			SS %	FPR (h^{-1})	PP (min)	SS_RP	SS %	FPR (h^{-1})	PP (min)	SS_RP
1	64.5	5	40	0.05	10	20	40	0.16	40	60
2	48.5	6	50	0.06	30	33.33	50	0.02	30	16.67
3	24	4	50	0.08	40	50	75	0.37	30	75
4	83.6	4	75	0.22	30	75	75	0.29	20	50
5	65.1	1	100	0.02	40	100	100	0.14	30	100
6	40.8	2	100	0.23	40	100	100	0.33	40	100
7	139.6	6	100	0.17	40	50	100	0.10	10	33.33
8	58.8	2	50	0.02	20	50	50	0.00	40	0
Av./T	524.9	30	66.7	0.12	31.2	49.99	70	0.15	30	46.66
9	52.7	7	57.1	0.16	20	42.86	57.1	0.22	40	57.14
10	160.3	9	44.4	0.05	40	22.22	44.4	0.46	10	33.33
Av./T	213	16	50	0.08	30	31.25	50	0.4	25	43.75
Av./T	737.9	46	60.9	0.11	31	43.48	63.0	0.22	29	45.65

FPR – False prediction rate; RP – Random predictor; PP – Preictal period; SS – Sensitivity

and reduced/differential methods outperform prediction results of other three methods, reminding us of the significance of using feature reduction methods.

Statistical Validation

To verify the performance of the proposed predictor, upper critical SS values obtained from random predictor (RP) model for the significance level of $\alpha = 0.05$ are also presented in Tables 2 and 3. Upper SS values for RP model (SS_RP) were obtained using FPR values from the same tables and considering $d = 15$ (completely uncorrelated recording channels) corresponding to the 15 possible combinations of 2 out of 6 selected channels. As seen from both tables, most of the sensitivities and all of the average sensitivities produced by our method are higher than their respective upper critical SS values achievable from the analytical random predictor. However, as the EEG signals recorded from the neighboring electrodes may be correlated to some extent, the effective value d_{eff} of d might be even smaller than 15,^[43] thus providing a more concrete proof for the performance of the proposed method.

DISCUSSION

In this study, we have tried to uncover the relationship between the features of different channels during both preictal and nonpreictal states. Two bivariate approaches employing relative and differential features are studied, and the results are compared with the original univariate features.

Relative and Differential Features

The combination of 22 univariate features extracted from 6 channels in one feature space, and been classified using SVM, has shown significant improvements in contrast to the predictions using one single univariate features only. The

dimension of the feature space achieved from univariate features was 132, whereas the proposed bivariate relative and differential features produced a feature space with a dimension of 330. The results of univariate features and two bivariate relative and bivariate differential approaches, presented in Tables 2 and 3, indicate that no significant improvement is achieved using bivariate approaches without feature reduction compared to the univariate features. The reason behind this could be the nature of SVM classifier, which implicitly considers the best possible linear combination of the features, and therefore, the proposed linear combination approaches could not help improve the results further. However, as seen from the same tables, applying feature reduction on both differential and relative approaches has improved the results. More specifically FPR values are decreased for all patients (except patients 2 and 7) using a differential method and by employing reduced features. In fact, nondiscriminative features increase the difficulty of finding the right support vectors by SVM classifier, hence increasing the chance of wrong classification and false prediction.

Discriminative Features

Table 4 presents the highest ranked five features resulting from mRMR feature selection method. In this Table, the value under each focal channel indicates the ratio of observed seizure initiations on that channel. For example, if a patient experiences 5 seizures, two of which originating in the neighborhood of channel F3, the corresponding value for F3 will be 0.4.

As seen from Table 4, the channel combinations of top five selected features for the differential approach are 80% from focal-focal channels and 20% from focal-nonfocal channels, and none is non-focal-nonfocal. Furthermore, the channel combinations of top five selected features

Table 4: The five high ranked features for each of the 10 patients (differential and ratio)

ID	Patient info.			Channel comb. scheme	Selected features				
	Focal channels				Feat. 1	Feat. 2	Feat. 3	Feat. 4	Feat. 5
	Ch. 1	Ch. 2	Ch. 3						
1	F10	T8	P8	Reduced diff.	θ (T8, P8)	HjorthC (P8, T7)	β (F10, T7)	β (T8, P7)	θ (T8, P7)
	I	I	I	Reduced rel.	ARCoeff (F10, T8)	HjorthC (F10, P8)	Wavelet Band1 (F10, T8)	Wavelet Band2 (F10, T8)	Wavelet Band3 (F10, T8)
2	FP1	AF7	F3	Reduced diff.	ARCoeff (FP1, AF7)	HjorthC (FP1, AF7)	θ (FP1, F3)	α (FP1, F3)	Variance (FP1, AF7)
	0.86	0.93	0.79	Reduced rel.	ARCoeff (FP1, AF7)	δ (FP1, F3)	LTEnergy (FP1, AF7)	HjorthM (FP1, AF7)	HjorthC (FP1, AF7)
3	T8	P8	TP8	Reduced diff.	ARCoeff (T8, P8)	Wavelet Band5 (T8, P8)	Wavelet Band2 (T8, P8)	Wavelet Band3 (T8, P8)	DecorrTime (T8, T10)
	I	0.86	0.86	Reduced rel.	ARCoeff (T8, P8)	Skewness (T8, P8)	LTEnergy (T8, P8)	HjorthM (T8, P8)	HjorthC (T8, P8)
4	AF7	FP1	F3	Reduced diff.	ARCoeff (FP1, AF7)	θ (FP1, FP2)	α (FP1, FP2)	DecorrTime (FP1, FP2)	HjorthM (FP1, AF7)
	0.86	0.86	0.86	Reduced rel.	ARCoeff (FP1, AF7)	HjorthC (FP1, AF7)	ARCoeff (FP1, FP2)	DecorrTime (FP1, FP2)	LTEnergy (FP1, FP2)
5	FT10	T10	T4	Reduced diff.	ARCoeff (T3, T4)	Wavelet Band2 (T3, T4)	α (T3, T4)	θ (T3, T4)	γ (T3, T4)
	I	I	I	Reduced rel.	ARCoeff (T3, T4)	LTEnergy (T3, T4)	ARCoeff (T3, FT9)	DecorrTime (T3, FT9)	LTEnergy (T3, FT9)
6	FT10	T10	TP10	Reduced diff.	ARCoeff (FT10, T10)	α (FT10, TP10)	α (FT10, T10)	β (FT10, T10)	γ (FT10, T10)
	0.83	0.83	0.83	Reduced rel.	ARCoeff (FT10, T10)	δ (FT10, T10)	α (FT10, TP10)	β (FT10, T10)	γ (FT10, T10)
7	F4	FP2	C4	Reduced diff.	ARCoeff (F4, C4)	α (F4, FP2)	DecorrTime (F4, FP2)	LTEnergy (F4, FP2)	HjorthM (F4, FP2)
	I	0.89	0.78	Reduced rel.	Variance (F4, F3)	Wavelet Band6 (F4, FP2)	γ (FP2, C3)	ARCoeff (C4, C3)	Wavelet Band5 (C4, C3)
8	CZ	FZ	F7	Reduced diff.	Wavelet Band6 (CZ, FZ)	δ (CZ, T6)	θ (CZ, T6)	α (CZ, T6)	Wavelet Band4 (CZ, FZ)
	0.43	0.43	0.43	Reduced rel.	ARCoeff (CZ, FZ)	HjorthC (CZ, T6)	δ (CZ, T6)	θ (CZ, T6)	α (CZ, T6)
9	TEMOB3	TEMOB4	TEMBP2	Reduced diff.	β (TEMBP2, TEMOB3)	HjorthM (TEMBP2, TEMOB3)	LTEnergy (TEMBP2, TEMOB3)	α (TEMBP2, TEMOB3)	DecorrTime (TEMBP2, TEMOB4)
	0.62	0.77	0.23	Reduced diff.	ARCoeff (TEMBP2, TEMOB3)	Wavelet Band5 (TEMBP2, TEMOB3)	HjorthM (TEMBP2, TEMOB4)	Variance (TEMBP2, TEMOB3)	Skewness (TEMBP2, TEMOB3)
10	F2IM1	F2IM2	F2M3	Reduced diff.	ARCoeff (F2IM2, F2IM1)	HjorthM (F2IM2, F2IM3)	(F2IM2, F2IM1)	Variance (F2IM2, F2IM1)	θ (F2IM2, F2IM1)
	0.83	I	0.66	Reduced rel.	DecorrTime (F2M3, F2IA1)	DecorrTime (F2M3, F2PI)	(F2IM2, F2IM1)	Skewness (F2IM2, F2IA1)	α (F2PI, F2IA1)

for ratio approach include 64% from focal-focal channels, 28% from focal-nonfocal channels, whereas only 8% are nonfocal-nonfocal. Based on these findings, it could be concluded that: (a) gradual preictal changes mostly correlate with focal signals, features of which being the most discriminative. (b) Nonfocal signals bear the least significant data for seizure prediction. (c) Focal-non-focal combinations possess valuable information for predicting seizures and should never be excluded from the beginning from possible channel combinations.

Also by further analyzing Table 4, it is found that for the differential approach three features of spectral power of Alpha band (9 times), autoregressive error (7 times), and spectral power of Theta band (6 times) appear more than other features, demonstrating highest discrepancies for differential channel combinations. Regarding the ratio approach however, the three features of autoregressive error (11 times), Hjorth complexity (5 times), and long-term energy (5 times) appear more frequently,

hence providing the highest discrepancies to the ratio combination scheme.

Channel Selection

The features extracted from different channels were compared using the proposed methods in both relative and differential ways. The use of relative or differential features provides a measure of similarity between channels from different aspects (phase/frequency, amplitude). Selection of three channels on the focal area was carried out, bearing in mind that during preictal state the focal channels will grow more similar patterns, hence differential and relative features will approach towards zero and one, respectively. The other three channels were selected far from the focal area mainly to provide us with information about the general state of the brain, and will perhaps also record epileptogenic changes during the preictal state. The nonfocal channels can also act as references for focal channels, when using bivariate features.

Higher number of recording channels is avoided. Extra channels are not desirable as they will become both harder to be mounted and fixed on the skull, more energy demanding, and more uncomfortable in future transportable warning devices. Consequently, to keep the complexity to the lowest possible, while profiting from maximum spatial information, the number of six channels is considered reasonable.

Comparison with Our Previous Study

The results presented in our previous study^[26] were achieved for the best preprocessing combination for each patient, whereas here, we have used an identical (non-best-for-all) preprocessing combination for all patients. For a reasonable comparison of the results, we have selected only those results obtained in previous work using the same combination of preprocessing methods employed in the current study.

Compared with our previous work, the method could provide slightly better results using a reduced set of features, which can hugely decrease the complexity of prediction algorithm. The basic goal for this study has been to obtain a less-complex algorithm making it suitable for portable devices necessitating low-power consumption criteria, meanwhile to improve the performance. As a result, a reduced set of bivariate features (30 features) compared to the 132 univariate features in our previous work, have been able to better the FPR from 4.32 false alarms per day to 2.64 false alarms per day, without losing the same level of SS.

CONCLUSIONS

Applying ML method on the relative and differential features could predict the epileptic seizures with slightly high SS and considerable higher specificity (lower FPR). The performance of the bivariate features was approved by comparing their results with those obtained from original univariate features. Since the number of bivariate features is very high, mRMR feature selection method was applied on the features to reduce the computational cost of the algorithm. The reduced feature space could provide better results in comparison to the original high dimensional feature space.

Future work can use the nonlinear features in conjunction with linear ones, or separately. This can check for the superiority of nonlinear over linear features as argued in some literature.

ACKNOWLEDGMENTS

Part of this work has been done in the Center for Informatics and Systems of the University of Coimbra (CISUC), Portugal, during a research stay of Jalil Rasekhi. JR wishes to thank for all the

supports received from the highly respected professors at CISUC. Mojtaba Bandarabadi would particularly like to acknowledge the Portuguese Foundation for Science and Technology (FCT - SFRH/BD/71497/2010).

REFERENCES

1. Fisher RS, van Emde Boas W, Blume W, Elger C, Genton P, Lee P, et al. Epileptic seizures and epilepsy: Definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia* 2005;46:470-2.
2. Carney PR, Myers S, Geyer JD. Seizure prediction: Methods. *Epilepsy Behav* 2011;22 Suppl 1:S94-101.
3. Sun DA, Sombati S, DeLorenzo RJ. Glutamate injury-induced epileptogenesis in hippocampal neurons: An *in vitro* model of stroke-induced "epilepsy". *Stroke* 2001;32:2344-50.
4. D'Alessandro M, Esteller R, Vachtsevanos G, Hinson A, Echaz J, Litt B. Epileptic seizure prediction using hybrid feature selection over multiple intracranial EEG electrode contacts: A report of four patients. *IEEE Trans Biomed Eng* 2003;50:603-15.
5. Le Van Quyen M, Martinerie J, Navarro V, Boon P, D'Havé M, Adam C, et al. Anticipation of epileptic seizures from standard EEG recordings. *Lancet* 2001;357:183-8.
6. Mirowski P, Madhavan D, Lecun Y, Kuzniecky R. Classification of patterns of EEG synchronization for seizure prediction. *Clin Neurophysiol* 2009;120:1927-40.
7. Rogowski Z, Gath I, Bental E. On the prediction of epileptic seizures. *Biol Cybern* 1981;42:9-15.
8. Viglione SS, Walsh GO. Proceedings: Epileptic seizure prediction. *Electroencephalogr Clin Neurophysiol* 1975;39:435-6.
9. Mormann F, Lehnertz K, David P, Elger CE. Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients. *Physica D Nonlinear Phenom* 2000;144:358-69.
10. Litt B, Echaz J. Prediction of epileptic seizures. *Lancet Neurol* 2002;1:22-30.
11. van Drongelen W, Nayak S, Frim DM, Kohrman MH, Towle VL, Lee HC, et al. Seizure anticipation in pediatric epilepsy: Use of Kolmogorov entropy. *Pediatr Neurol* 2003;29:207-13.
12. Elger CE, Lehnertz K. Prediction of seizure occurrence by chaos analysis: Technique and therapeutic implications", In: Rosenow F, Lüders HO, editors. *Presurgical assessment of the epilepsies with clinical neurophysiology and functional imaging. Handbook of Clinical Neurophysiology*. Vol. 3. Elsevier: Amsterdam; p. 491-500.
13. McSharry PE, Smith LA, Tarassenko L. Prediction of epileptic seizures: Are nonlinear methods relevant? *Nat Med* 2003;9:241-2.
14. Mormann F, Kreuz T, Andrzejak RG, David P, Lehnertz K, Elger CE. Epileptic seizures are preceded by a decrease in synchronization. *Epilepsy Res* 2003;53:173-85.
15. Elger CE, Lehnertz K. Seizure prediction by non-linear time series analysis of brain electrical activity. *Eur J Neurosci* 1998;10:786-9.
16. Andrzejak RG, Chicharro D, Elger CE, Mormann F. Seizure prediction: Any better than chance? *Clin Neurophysiol* 2009;120:1465-78.
17. Mormann F, Andrzejak RG, Elger CE, Lehnertz K. Seizure prediction: The long and winding road. *Brain* 2007;130:314-33.
18. Stacey W, Le Van Quyen M, Mormann F, Schulze-Bonhage A. What is the present-day EEG evidence for a preictal state? *Epilepsy Res* 2011;97:243-51.
19. Mormann F, Kreuz T, Rieke C, Andrzejak RG, Kraskov A, David P, et al. On the predictability of epileptic seizures. *Clin Neurophysiol* 2005;116:569-87.
20. Cerf R, el-Ouasdad EH. Spectral analysis of stereo-electroencephalograms: Preictal slowing in partial epilepsies. *Biol Cybern* 2000;83:399-405.

21. Park Y, Luo L, Parhi KK, Netoff T. Seizure prediction with spectral power of EEG using cost-sensitive support vector machines. *Epilepsia* 2011;52:1761-70.
22. Bandarabadi M, Rasekhi J, Teixeira CA, Dourado A. Epileptic seizure prediction using relative spectral power features. *Clin Neurophysiol* 2014; in press.
23. Direito B, Ventura F, Teixeira C, Dourado A. Optimized feature subsets for epileptic seizure prediction studies. *Conf Proc IEEE Eng Med Biol Soc* 2011;2011:1636-9.
24. Bandarabadi M, Teixeira CA, Sales F, Dourado A. Wepilet, optimal orthogonal wavelets for epileptic seizure prediction with one single surface channel. *Conf Proc IEEE Eng Med Biol Soc* 2011;2011:7059-62.
25. Costa R, Oliveira P, Rodrigues G, Leitão B, Dourado A. Epileptic seizure classification using neural networks with 14 features. In: Lovrek I, Howlett R, Jain L, editors. *Knowledge-based intelligent information and engineering systems*. Berlin/Heidelberg: Springer; 2008. p. 281–8
26. Rasekhi J, Mollaei MR, Bandarabadi M, Teixeira CA, Dourado A. Preprocessing effects of 22 linear univariate features on the performance of seizure prediction methods. *J Neurosci Methods* 2013;217:9-16.
27. Klatt J, Feldwisch-Drentrup H, Ihle M, Navarro V, Neufang M, Teixeira C, et al. The EPILEPSIAE database: An extensive electroencephalography database of epilepsy patients. *Epilepsia* 2012;53:1669-76.
28. Direito B, Teixeira C, Ribeiro B, Castelo-Branco M, Sales F, Dourado A. Modeling epileptic brain states using EEG spectral analysis and topographic mapping. *J Neurosci Methods* 2012;210:220-9.
29. Welch PD. The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans Audio Electroacoustics* 1967; 15:70-3.
30. van Laar JO, Peters CH, Houterman S, Wijn PF, Kwee A, Oei SG. Normalized spectral power of fetal heart rate variability is associated with fetal scalp blood pH. *Early Hum Dev* 2011;87:259-63.
31. Hjorth B. EEG analysis based on time domain properties. *Electroencephalogr Clin Neurophysiol* 1970;29:306-10.
32. Litt B, Esteller R, Echaz J, D'Alessandro M, Shor R, Henry T, et al. Epileptic seizures may begin hours in advance of clinical onset: A report of five patients. *Neuron* 2001;30:51-64.
33. Chisci L, Mavino A, Perferi G, Sciandrone M, Anile C, Colicchio G, et al. Real-time epileptic seizure prediction using AR models and support vector machines. *IEEE Trans Biomed Eng* 2010;57:1124-32.
34. Priestley MB, editor. *Spectral Analysis and Time Series*. London, New York: Academic Press; 1981.
35. Box GE, Jenkins G. *Time Series Analysis, Forecasting and Control*. Holden-Day, Incorporated; 1990. p. 500.
36. Stanski DR, Hudson RJ, Homer TD, Saidman LJ, Meathe E. Pharmacodynamic modeling of thiopental anesthesia. *J Pharmacokinetic Biopharm* 1984;12:223-40.
37. Petrosian A, Prokhorov D, Homan R, Dashei R, Wunsch D. Recurrent neural network based prediction of epileptic seizures in intra- and extracranial EEG. *Neurocomputing* 2000;30:201-18.
38. Graf AA, Smola AJ, Borer S. Classification in a normalized feature space using support vector machines. *IEEE Trans Neural Netw* 2003;14:597-605.
39. Peng H, Ding C, Long F. Minimum redundancy-maximum relevance feature selection. *IEEE Intell Syst* 2005;20:70-1.
40. Mirowski PW, Yann L, Madhavan D, Kuzniecky R. Comparing SVM and Convolutional Networks for Epileptic Seizure Prediction from Intracranial EEG. *Machine Learning for Signal Processing, MLSP 2008 IEEE Workshop on* 2008. p. 244-9.
41. Teixeira CA, Direito B, Feldwisch-Drentrup H, Valderrama M, Costa RP, Alvarado-Rojas C, et al. EPILAB: A software package for studies on the prediction of epileptic seizures. *J Neurosci Methods* 2011;200:257-71.
42. Teixeira C, Direito B, Bandarabadi M, Dourado A. Output regularization of SVM seizure predictors: Kalman Filter versus the "Firing Power" method. *Conf Proc IEEE Eng Med Biol Soc* 2012;2012:6530-3.
43. Schelter B, Winterhalder M, Maiwald T, Brandt A, Schad A, Schulze-Bonhage A, et al. Testing statistical significance of multivariate time series analysis techniques for epileptic seizure prediction. *Chaos* 2006;16:013108.

How to cite this article: Rasekhi J, Mollaei MRK, Bandarabadi M, Teixeira CA, Dourado A. Epileptic Seizure Prediction based on Ratio and Differential Linear Univariate Features. *J Med Sign Sence* 2015;5:1-11.

Source of Support: This work was partially supported by EU FP7 211713 EPILEPSIAE Project and partially supported by iCIS-CENTRO-07-0224-FEDER-002003, **Conflict of Interest:** None declared

BIOGRAPHIES



Jalil Rasekhi received his M.Sc. degree in Electrical Engineering - Communications in 2004 from the University of Tehran, Iran. Currently, he is a PhD candidate in electrical engineering, bioinformatics at Babol Noshirvani institute of technology, Iran.

His research interests include cellular Epileptic seizure prediction & detection, Artificial Intelligence and robotics.

E-mail: jrasekhi@gmail.com



Mohammad Reza Karami-Mollaei received the B.Sc. in Electrical and Electronic Engineering in 1992, M.Sc. of signal processing in 1994, and PhD in 1998 in Biomedical Engineering from I.N.P.L d’Nancy of France. He is now an associate

professor with the Department of Electrical and Computer Engineering, Babol University of Technology. Since 1998 his research is in signal and speech processing. He published more than 80 articles in journals and conferences. He teaches Digital Signal, Biomedical and speech processing at university. His research interests include Speech, Image and signal processing.

E-mail: mkarami@nit.ac.ir



Mojtaba Bandarabadi received his PhD degree in Information Science and Technology from the University of Coimbra, Portugal, in 2015. His PhD research focused on the development of robust and low-complexity measures for epileptic seizure

prediction and early detection. He obtained his both B.Sc. and M.Sc. degrees in Electronics Engineering from the University of Mazandaran, Iran, in 2007 and 2009, respectively. His research interests include computational neuroscience, signal processing, pattern recognition, and epilepsy.

E-mail: mojtaba@dei.uc.pt



Cesar A. Teixeira is an assistant professor at University of Coimbra, Portugal. He has a graduation (2003) in Systems and Computation Engineering and a PhD (2008) in Electronics Engineering and Informatics both from the University of the Algarve,

Portugal. His expertise is on bio-signal and image processing and classification, more precisely on EEG and ultrasound signal processing.

E-mail: cteixei@dei.uc.pt



António Dourado is Full Professor and Scientific Coordinator of Soft Computing Group of Center for Informatics and Systems of the University of Coimbra (CISUC). He is teaching courses on Systems Engineering, Automatic Control, Fuzzy

Systems, Neural Networks, and Theory of Computation at graduate and undergraduate levels. He has been and is involved in international (namely in EU FP7) and national projects (in collaboration with industry) and was the coordinator of the European project EPILEPSIAE- Evolving Platform for Improving the Living Expectations of Patients Suffering from Ictal Events. He is author or co-author of more than 200 international publications in referred journals, book chapters and conferences. He is member of IEEE (Computational Intelligence Society, System Man and Cybernetic Society, Engineering in Medicine and Biology Society) and has been co-founder of European Control Association and Portuguese Association of Automatic Control (IFAC National Member). His research interests are Computational Intelligence, Signal Processing, data mining for medical and industrial applications, and intelligent control.

E-mail: dourado@dei.uc.pt