

1 Improving Classification of Epileptic and Non-Epileptic EEG Events

2 by Feature Selection

3
4 * Evangelia Pippa¹, Evangelia I. Zacharaki¹, Iosif Mporas¹

5 Vasiliki Tsirka², Mark P. Richardson², Michael Koutroumanidis² and Vasileios
6 Megalooikonomou¹

7 ¹Multidimensional Data Analysis and Knowledge Management Laboratory

8 Dept. of Computer Engineering and Informatics, University of Patras

9 26500 Rion-Patras, Greece

10 ²Dept. of Clinical Neurophysiology and Epilepsies

11 Guy's & St. Thomas' and Evelina Hospital for Children, NHS Foundation Trust/ King's College,

12 London, United Kingdom

13 *Corresponding Author's email: pippa@ceid.upatras.gr

14
15 **Abstract:** Correctly diagnosing generalized epileptic from non-epileptic episodes, such as
16 psychogenic non epileptic seizures (PNES) and vasovagal or vasodepressor syncope (VVS), despite its
17 importance for the administration of appropriate treatment, life improvement of the patient, and cost
18 reduction for patient and healthcare system, is rarely tackled in the literature. Usually clinicians
19 differentiate between generalized epileptic seizures and PNES based on clinical features and video-
20 EEG. In this work, we investigate the use of machine learning techniques for automatic classification
21 of generalized epileptic and non-epileptic events based only on multi-channel EEG data. For this
22 purpose, we extract the signal patterns in the time domain and in the frequency domain and then
23 combine all features across channels to characterize the spatio-temporal manifestation of seizures.
24 Several classification algorithms are explored and evaluated on EEG epochs from 11 subjects in an
25 inter-subject cross-validation setting. Due to large number of features feature ranking and selection is
26 performed prior to classification using the ReliefF ranking algorithm within two different voting
27 strategies. The classification models using feature subsets, achieved higher accuracy compared to the
28 models using all features reaching 95% (Bayesian Network), 89% (Random Committee) and 87%
29 (Random Forest) for binary classification (epileptic versus non-epileptic). The results demonstrate the
30 competitiveness of this approach as opposed to previous methods.

1 **Keywords:** epileptic seizures, PNES, vasovagal syncope, classification, machine learning.

2

3 **1. Introduction**

4 One of the most common and challenging medical cases in everyday clinical practice is that of
5 patients reporting one or more episodes of paroxysmal loss of consciousness or altered awareness. The
6 management of these medical cases may be proven to be demanding, time consuming and expensive
7 and finally, in spite of the extensive and exhaustive investigation, the underlying diagnosis may remain
8 elusive [1,2,3]. The differential diagnosis that a clinician usually faces is mainly that of an epileptic
9 seizure, a possible psychogenic non epileptic seizure (PNES) and a probable vasovagal syncope (VVS).

10 The diagnosis of epilepsy and its differentiation from other causes of TLoC is typically based on
11 historical information and is assisted by specific tests [2]. However clinical information is commonly
12 fragmented or even missing because patients may have limited or no recall of the event and a witness
13 account might not be available to describe diagnostically decisive clinical phenomena [1,2]. Even when
14 a witness is available, diagnosis may be difficult and often remains uncertain because convulsive
15 syncope, a seizure-like reaction resulting from global cerebral hypoperfusion, can mimic epileptic
16 seizures [3,4]. Agreement between physicians as to the nature of a single event may also be limited [5].
17 Such diagnostic uncertainty has cost both in terms of mortality and ongoing morbidity and in terms of
18 the financial burden associated with hospitalization and repeated investigations.

19 Epileptic seizures are brief episodes of abnormal excessive or synchronous neuronal activity in
20 the brain of patients suffering from epilepsy [6]. During an epileptic seizure there are several specific
21 changes recorded in the electroencephalogram (EEG) which is a sensitive and important test used to
22 evaluate patients with suspected epilepsy. There are certain characteristic ictal neurophysiological
23 patterns that support the identification and detection of epileptic events and postictal and/or interictal
24 abnormalities that can provide supplementary information. Fig. 1 shows generalized spike wave
25 abnormalities from an epileptic patient. Specifically, there is a burst of generalized 3-5 Hz spike and
26 slow wave complex lasting approximately 5 secs.

27 **FIGURE 1**

28 Psychogenic non-epileptic seizures (PNES) are sudden paroxysmal changes in behavior or
29 consciousness, that resemble epilepsy but are not accompanied by the electrophysiological changes that
30 characterize an epileptic seizure [7]. Although the clinical history can help differentiate these episodes,

1 it is not unlikely to have inconclusive and insufficient event description by the patient and witnesses,
2 not being able to confidently exclude an underlying epileptic disorder. In these cases the diagnosis of
3 PNES can be supported by video-EEG monitoring, especially if a psychogenic event is captured, since
4 in the case of PNES there are no specific EEG changes. Fig 2 shows an EEG fragment during a PNES.
5 No EEG correlates can be seen and the recording is frequently marred by muscular artifacts.

6 **FIGURE 2**

7 Vasovagal or vasodepressor syncope is a common type of syncope and various mechanisms
8 have been postulated for explaining the characteristic association of hypotension and bradycardia. The
9 term "vasovagal" was introduced by Lewis [8] to indicate that both blood vessels and heart were
10 implicated and since atropine reversed the bradycardia but not the hypotension he considered
11 vasodilatation as the primary responsible factor. During a vasovagal syncopal attack there may be some
12 characteristic EEG changes starting with progressive generalized theta slowing of background rhythms
13 followed by sometimes hypersynchronous delta activity of high voltage, (beta / alpha → theta → delta)
14 and appearance of progressively lower voltage rhythms until isoelectric suppression [9,10] (see Fig. 3).
15 This pattern is progressively reversed after the patient's fall, during his/her recovery. These changes do
16 not include any ictal activity.

17 **FIGURE 3**

18 Several methods have been proposed for the classification of EEG captured events into epileptic
19 or normal [11,12,13,14,15]. The problem of the discrimination between ictal and interictal EEG signals
20 has been studied [16], too. However, only a few studies deal with the differentiation between epileptic
21 and other paroxysmal episodes of loss of consciousness such as PNES and vasovagal syncope. It is
22 worth to note that the discrimination between different types of non-epileptic events is considerably
23 more useful in diagnostic procedure given the semiological resemblance between the aforementioned
24 paroxysmal attacks. Furthermore, according to [7] the one third of PNES patients may have clinical
25 convincing GrandMal like seizures. This makes discrimination between PNES and epileptic seizures a
26 challenging task, especially in an online monitoring system for automatic detection of epileptic events,
27 such as [17], where false alarms caused by events similar to epilepsy are undesired.

28 To the best of our knowledge, only a few studies have been proposed in the literature for
29 automated classification between epileptic and non-epileptic pathological events from EEG. Poulos et
30 al. [18] proposed an algorithm which estimates a number of auto-correlated coefficients extracted from

1 an appropriately selected epileptic EEG segment and examines whether these coefficients are
2 correlated with the coefficients of the unknown EEG segments in order to classify the latest into
3 epileptic or non-epileptic. Their algorithm obtained a sensitivity of 83% for 90% specificity.
4 Papavlasopoulos et al. [19] trained a LVQ1 neural network on an appropriately extracted set of auto-
5 correlation coefficients (codebook) and used the resulting model to classify the corresponding feature
6 vectors of the unknown EEG segments. The LVQ1 network achieved 86% accuracy. The feature
7 extraction methods of the aforementioned classification frameworks, as well as the achieved results,
8 can be found in [20]. Statistical analysis of the results based on chi-square test showed that the LVQ
9 neural network method is superior than the cross-correlation one [20].

10 Regarding the features used for the classification of EEG segments the relevant works in the
11 literature are considerably more. In the majority of them, the analysis is based on the estimation of the
12 EEG channels' spectral magnitude [11, 15, 21, 22]. Other EEG features that have been reported are the
13 autoregressive filter coefficients, the continuous and discrete wavelet transform, as well as energy per
14 brain wave (delta, theta, alpha, beta, gamma) bands [15,21, 23]. Finally, time domain features have
15 been proposed, such as zero-crossing rate [24] and statistics of the EEG samples per channel [15,21].

16 In this study, we evaluate a large set of time and frequency domain features which have been
17 widely used for the analysis of EEG signals in the literature. In addition to the reported evaluations
18 found in the literature, we extend the non-epileptic class to both PNES and VVS events. The diagnosis
19 of epilepsy is more challenging compared to the detection of seizure onset due to the semiological
20 resemblance between epileptic and non-epileptic events, especially when video-EEG monitoring is not
21 incorporated [25]. Also, the classification of abnormal episodes into different types requires a broad
22 knowledge of EEG patterns across patients, while seizure detection can rely on patient-specific models
23 which are easier to learn, especially for generalized seizures [26]. For the evaluation, we examined a
24 number of different classification algorithms. Our classification methodology can be used as part of
25 our previous seizure detection architecture [26,27] in order to discriminate the detected events into
26 epileptic or non-epileptic.

27 In a further step, feature ranking investigation using two different strategies (one based on
28 frequency of feature appearing in a specific rank and the other based on sum of the weights assigned by
29 the ReliefF ranking algorithm) was performed. The classification models using subsets of N best
30 features were evaluated and revealed the most significant features for the classification task.

1 The rest of this paper is organized as follows. In Section 2 the classification methodology is
2 presented and details about the evaluation data are provided. Section 3 describes the experimental
3 protocol followed and presents the achieved results. Finally, in Section 4 we conclude this work.

4 5 **2. Material and methods**

6 **2.1 Methodology for classification of generalized epileptic and non-epileptic events**

7 The presented architecture for classification between generalized epileptic and non-epileptic
8 EEG events is part of an end-to-end system for monitoring and analysis of brain disorders, the
9 ARMOR framework [17]. Within the ARMOR framework patients suffering from seizures are
10 monitored through sensors and the multi-parametric data are processed automatically (real-time by
11 software tools) or semi-manually (offline with the support of software tools and visualizations) by
12 neurology experts [26,27].

13 The proposed classification methodology can be used as additional module after the seizure
14 detection and focal-vs-generalized events classification components [26,27] in order to discriminate the
15 detected events into generalized epileptic, manifested by Generalized Spike Wave discharges (GSW) or
16 non-epileptic, such as PNES and VVS. The block diagram of the overall architecture is illustrated in
17 Figure 4.

18 **FIGURE 4**

19 Initially, the multidimensional EEG data are preprocessed by applying notch filtering (at 50Hz),
20 baseline correction and re-sampling at 250 Hz (in order to obtain a common resolution level for all data
21 coming from different patients and acquisition systems). Frame blocking of the incoming EEG streams
22 to epochs of constant length w (= 2 seconds) is performed with constant time-shift and without time-
23 overlap between successive epochs. Each epoch is a $N \times w$ matrix, where N is the number of selected
24 EEG electrodes. A large number of features is extracted for each one of the N electrodes to
25 characterize the temporal patterns and frequency content of each epoch. The extracted time domain and
26 frequency domain features from all electrodes are concatenated to a single feature vector as a
27 representative signature for each epoch. Details on the type of extracted features are provided in section
28 2.3.

29 All epochs are used as input to ARMOR's seizure detection module which detects paroxysmal
30 events. The epochs classified as normal are ignored whereas epochs classified as seizure are further

1 entered to the seizure type classifier. In this final step, models for binary classification between
2 generalized epileptic or non-epileptic events (PNES or VVS), which have been previously built in a
3 training phase, are used in order to label the epochs. Each epoch is classified independently and no
4 temporal constraints (across epochs) are applied, such as taking into consideration the class label of the
5 precedent or subsequent epoch or the total event duration.

6 During the training phase of the classification architecture, epochs with known class labels
7 (labeled manually by medical experts) are used to train binary classification models, i.e. generalized
8 epileptic (GSW) vs non-epileptic (PNES and VVS).

9 During the test phase the unknown multidimensional EEG signal is preprocessed and
10 parameterized with similar setup as in the training phase. Each extracted feature vector is provided as
11 input to the seizure detector and according to the decision to the seizure classifier.

12

13 **2.2 Data**

14 The previously described classification methodology was evaluated on multi-parametric
15 recordings performed within the ARMOR project [17]. The recordings were performed in the
16 Department of Clinical Neurophysiology and Epilepsies in St Thomas' Hospital in London and
17 acquired from 11 patients in total. All participants had at least one of their typical epileptic or non
18 epileptic events captured during the recording procedure. The epileptic group, consisted of patients
19 with known diagnosis of idiopathic generalized epilepsy, manifested clinically with absence seizures
20 and they had at least one clinical episode captured during the recording associated with generalized
21 spike wave discharges on the EEG. The non epileptic group included patients who had sustained a
22 vasovagal syncope (2 participants) or a psychogenic non epileptic attack (5 participants) during their
23 monitoring. The epilepsy group contains 105 generalized seizures while the non-epilepsy groups
24 include 21 events (19 PNES and 2 VVS). Patients with focal seizures were excluded from this analysis.

25 The recordings were performed using conventional AgCl EEG electrodes positioned according to
26 the extended international 10–20 system. A subset of the main EEG channels was selected for analysis
27 which included the following channels: Fp2, F8, F4, T4, C4, A2, P4, T6, O2, Fp1, F7, F3, A1, C3, T3,
28 P3, T5, O1, Fz, Cz, Pz. The recordings were manually annotated by expert Neurologists of the King
29 College London. Only epochs during paroxysmal events were considered for training and for testing.
30 All data were stored in EDF+ formatted files [28].

1

2 **2.3 Feature Extraction and Classification Algorithms**

3 After preprocessing, time domain and frequency domain features were extracted for each epoch.
4 In particular, each of the EEG channels was parameterized using the following features: (i) time-
5 domain features: minimum value, maximum value, mean, variance, standard deviation, percentiles
6 (25%, 50%-median and 75%), interquartile range, mean absolute deviation, range, skewness, kurtosis,
7 energy, Shannon's entropy, logarithmic energy entropy, number of local maxima and local minima,
8 zero-crossing rate, and (ii) frequency-domain features: 6-th order autoregressive-filter (AR)
9 coefficients, power spectral density, frequency with maximum and minimum amplitude, the power of
10 continuous wavelet transform using symlet 5 mother wavelet of scale 25 and 32, the power of discrete
11 wavelet transform with mother wavelet function Daubechies 16 and decomposition level equal to 8.
12 This resulted to 55 variables for each of the $N=21$ EEG channels producing a feature vector of
13 dimensionality equal to 1155 in total.

14 The computed feature vectors, V , were used to train classification models. In order to evaluate the
15 ability of the above features to discriminate between epileptic and non-epileptic epochs we examined
16 several classification algorithms, including BayesNet [29,30], RandomCommittee, Random Forest
17 [31], IBk [32] and SMO [33,34] with RBF kernel, which were implemented by the WEKA machine
18 learning toolkit [35]. The classifiers in our study were selected in an attempt to evaluate representative
19 algorithms for each one of the main categories of machine learning classification methods including
20 probabilistic networks (BayesNet), decision trees (RandomForest), support vector machines (SMO),
21 ensemble classifiers (RandomCommittee and RandomForest) but also simple methods such as k
22 nearest neighbors (IBk).

23 During the test phase, the EEG recordings were pre-processed and parameterized as during
24 training. Each classification model was used to label each of the detected seizure epochs. In the present
25 evaluation no additional rules (e.g. knowledge based rules regarding events duration) were applied on
26 the classification decision .

27 Evaluation was performed in a leave-one-out cross-validation setting. Specifically, each time one
28 subject was left-out for testing, while the rest of the subjects were used for training. For the left-out
29 subject, all epochs between seizure onset and offset were used as testing samples. Table 1 shows the
30 number of epochs that were extracted from each subject during the seizure(s).

1 **TABLE 1**

2 The purpose of this study was to evaluate the seizure classification module, thus only paroxysmal
3 events were used for training and testing of the classifiers. Evaluation of the total ARMOR framework
4 may include the combined use of seizure detection and seizure classification in future work.

5
6 **2.4 Feature Ranking and Feature Subsets Evaluation**

7 In a further step we examined the discriminative power of the extracted features for the
8 classification of epileptic and non-epileptic EEG events. The ReliefF algorithm [36] (which is an
9 extension of an earlier algorithm called Relief [37]) was used for estimating the importance of each
10 feature in binary classification (generalizing to polynomial classification by decomposition into a
11 number of binary problems). In the ReliefF algorithm the weight of any given feature decreases if the
12 squared Euclidean distance of that feature to nearby instances of the same class is more than the
13 distance to nearby instances of the other class. ReliefF is considered one of the most successful feature
14 ranking algorithms due to its simplicity and effectiveness [38, 39,40] (only linear time in the number of
15 given features and training samples is required), noise tolerance and robustness in detecting relevant
16 features effectively, even when these features are highly dependent on other features [38,41].
17 Furthermore, ReliefF avoids any exhaustive or heuristic combinatorial search compared with
18 conventional wrapper methods and usually performs better compared to filter methods due to the
19 performance feedback of a nonlinear classifier when searching for useful features [40].

20 In this study, ranking is performed by following a leave-one-out strategy on the available subjects.
21 Specifically, for each leave-one-out experiment, feature ranking is performed using the ReliefF
22 algorithm in each training subset. We combine the rankings of all leave-one-out experiments and
23 calculate the total rank of features using two different strategies. The first strategy calculates the total
24 rank of features according to the frequency of a feature appearing in a specific rank. For example the
25 top-ranked feature is assumed to be the one that more frequently has the highest ranking score,
26 regardless of the distribution of the scores it receives across experiments. The second strategy
27 calculates the total rank of features according to the sum of the weights assigned by ReliefF in each
28 training set. We examined the performance of the method, in terms of accuracy, sensitivity and
29 specificity, for different number of N-best features (N =10, 20, 30, ...100, 200, 300, ... ,1100), with
30 respect to the above strategies of feature ranking.

1

2 **3. Results and Discussion**

3 The classification methodology presented in Section 2.1 was evaluated using the classification
4 algorithms and the cross-validation scheme described in Section 2.3. The accuracy, sensitivity and
5 specificity are defined as:

$$6 \quad accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (2)$$

$$7 \quad sensitivity = \frac{TP}{TP+FN} \quad (3)$$

$$8 \quad specificity = \frac{TN}{FP+TN} \quad (4)$$

9 where TP denotes the true positives, TN the true negatives, FP the false positives and FN the false
10 negatives. The results of the method using all features are shown on the left of Table 2. Here we
11 consider the epileptic class as the positive and the non-epileptic class (PNES or VVS) as the negative.

12 **TABLE 2**

13 As can be seen in Table 2, the overall highest accuracy of the proposed methodology for
14 classification between epileptic and non-epileptic EEG events is 86% for BayesNet classification.
15 RandomCommittee and Random Forest classification models follow with 83% and 74% accuracy,
16 respectively. For the classifier with the highest accuracy (BayesNet), the sensitivity (or recall), i.e. the
17 fraction of actual epileptic events which are correctly identified as such, is 92% and the specificity, i.e.
18 the proportion of non-epileptic events (either PNES or VVS) which are correctly classified as such, is
19 78%.

20 In a further step, we applied feature ranking using the ReliefF algorithm and the two strategies
21 described in Section 2.4. The performance of the classification, in terms of accuracy, for different
22 number of N-best features (N =10, 20, 30 ..., 100, 200, 300, ... ,1100) and for each algorithm
23 separately are shown in Fig. 5 for the 1st ranking strategy and in Fig. 6 for the 2nd ranking strategy.

24 **FIGURE 5**

25 **FIGURE 6**

26 As can be seen in the above figures the highest classification accuracy is achieved when a small
27 subset of discriminative features are used. Specifically, when the 1st ranking strategy is used the
28 highest accuracy is achieved for a subset of 10 best features with a percentage of 95% for the Bayesian

1 Network, which is sufficiently high in comparison to the accuracy achieved when all features are used.
2 Similarly, Random Committee and Random Forest achieve their highest accuracies for a subset of 300
3 and 200 best features respectively. The reported accuracies for these subsets of features are 92% and
4 87% for each algorithm respectively. IBk and SMO follow with an accuracy of 86% when a subset of
5 40 best features is used and 87% for a subset of 200 best features, respectively.

6 The 2nd ranking strategy shows similar behavior. For the Bayesian Network the highest accuracy
7 (94%) is achieved for a subset of 50 best features. Random Committee, Random Forest, IBk and SMO
8 follow with 85% for a subset of 70 best features, 89% for a subset of 50 best features, 84% for a subset
9 of 70 best features and 90% for a subset of 60 best features, respectively.

10 In general, Random Forest and Random Committee seem to be the less stable algorithms, SMO
11 on the other hand, although not the most accurate classifier, shows a more stable behavior as function
12 of number of retained features, with the accuracy decreasing significantly for more than 200 features.

13 Tables 3 and 4 show the 50 best features according to the ranking strategy 1 and 2 respectively.

14 **TABLE 3**

15 **TABLE 4**

16 As can be seen, in general the two ranking strategies overall agree. Both of them rank features *nmin*
17 (number of local minima), *nmax* (number of local maxima), *aryule3* (the 3rd coefficient of 6th order
18 autoregressive filter), *minfreq* (frequency with minimum power), *cwt25* and *cwt32* (the coefficients of
19 continuous wavelet transform using symlet 5 mother wavelet of scale 25 and 32) in the top 50 features.

20 The number of local minima (*nmin*) and the number of local maxima (*nmax*) seem to be the
21 features with the highest discriminative ability. Since these features measures the smoothness of the
22 signal it seems that the smoothness of the epileptic epochs is different from the one of non-epileptic
23 epochs and aids the discrimination among them. Such a claim can be verified from the distributions of
24 the values of the *nmin* (see Fig. 7) and *nmax* (see Fig. 8) features for the epileptic and non-epileptic
25 class.

26 **FIGURE 7**

27 **FIGURE 8**

28 In both figures 7 and 8, the blue boxes indicate the distribution of the feature values on the epileptic
29 class and the black boxes the distribution of the feature values on the non-epileptic class. As can be
30 seen, there is a perfect discrimination between the epileptic and non-epileptic main boxes with the non-

1 epileptic epochs having a significantly larger number of local minima and maxima indicating less
2 smooth signal compared to the generalized spike waves. The only overlap is observed between the
3 extreme values of two classes (whiskers of the boxplots). The evaluation of our framework using only
4 these two features (*nmin* and *nmax*) extracted from all the available channels resulted in 91% accuracy
5 for 92% sensitivity and 89% specificity. The performance in terms of accuracy increases slightly when
6 *nmin* and *nmax* are extracted from the 5 best channels (Fp1, T4 , T5, F7, Fp2), reaching 92%. This
7 increase indicates that the frontotemporal regions in the brain covered by the aforementioned channels
8 might be more important in discriminating generalized spike waves from PNES or VVS. The next most
9 important features for discriminating epileptic from non-epileptic events are *aryule3*, *minfreq* and
10 *cwt25* and *cwt32*. The autoregressive model specifies whether the EEG epoch depends linearly on its
11 own previous values by expressing the signal with lagged terms of itself. In particular, the AR model
12 residual (i.e. the prediction error) shows how possible is to model each sample as a linear combination
13 of its previous ones. The lower absolute values of the AR coefficients of the non epileptic class (see
14 Fig. 9 for the *aryule3* feature values) indicates that the signal of the non-epileptic class is much more
15 noisy and stochastic-like compared to the epileptic signals which seem to be more structured and
16 deterministic-like. Such an experimental result is consistent with our intuition about the two types of
17 signals and the clinicians description of the events.

18 **FIGURE 9**

19 Differentiation is also observed on the frequency with the minimum power (*minfreq*) in the
20 spectrogram of epileptic and non-epileptic epochs (see Fig. 10), with the *minfreq* of the epileptic class
21 having a much greater range of values compared to the non-epileptic class in which the *minfreq* values
22 are clustered around 50Hz. Note that this finding is not due to notch filtering since the same
23 preprocessing was applied to all data (both epileptic and non-epileptic).

24 **FIGURE 10**

25 Finally, the expression of each epoch as a linear combination of the chosen wavelet basis functions
26 captures the frequency content of the epoch in a localized area of the signal which seem to highlight
27 the differences between the two classes (see Fig. 11 and 12).

28 **FIGURE 11**

29 **FIGURE 12**

1 Finally, in order to examine the ability of the BayesNet classifier to discriminate each type of
2 pathological events (GSW, PNES or VVS) from the others, we performed binary classification of all
3 possible pairs of pathological events (GSW-PNES, GSW-VVS and PNES-VVS). The results in terms
4 of classification accuracy for different number of N-best features (10, 20, ..50) are shown in Fig. 13.

5 **FIGURE 13**

6 As can be seen, the PNES-VVS classification problem is the most difficult case for the
7 classifier. The best classification accuracy (76%) for PNES-VVS pair is achieved when all features
8 (1155) are used. On the other hand, GSW-PNES and GSW-VVS pairs are much easier cases for the
9 classifier obtaining their maximum accuracy for the 10 best features. Specifically, GSW-PNES
10 classification achieves 96% accuracy for 96% sensitivity and 100% specificity while GSW-VVS
11 classification results in slightly lower percentages, i.e. 93% accuracy for 96% sensitivity and 87%
12 specificity. Since generalized spike waves are very specific ictal neurophysiological patterns, they
13 present much more consistent features (compared to the other types) which makes their detection an
14 easier task. On the other hand, PNES has no specific EEG patterns but is frequently accompanied by
15 muscular artifacts which present a variability across subjects. Similar variability appears even between
16 consecutive epochs of VVS examples since there are several changes that happen successively in time
17 during such an episode (beta / alpha \rightarrow theta \rightarrow delta \rightarrow lower voltage rhythms \rightarrow isoelectric
18 suppression). It seems that the variability in the feature values of the PNES and VVS epochs is high
19 (in respect to the available training data) impeding the learning of a discrimination model.

20 While 19 PNES appear to make a rather limited dataset, we believe that are sufficient given the
21 lack of ictal EEG changes and the fact that their variability reflects only muscle and movement
22 activities. The main problem is the really small sample of the 2 VVS-patients. However, VVS typically
23 occur very rarely, in most patients annually, and only in very few patients more frequently, say
24 monthly. It is therefore extremely unlikely to record them on standard EEG that is a 20min to one hour
25 "snapshot" of brain activity. Still, because of the rather predictable sequence of EEG changes (alpha-
26 theta delta etc) we believe that reasonable learning of a discrimination model is achievable / possible.

27 The proposed methodology takes into account features extracted from all the available channels
28 by concatenating them in a single feature vector. The spatial localization of the features is encoded in
29 their location within the feature vector presented to the classifier. Since the seizure onset patterns in
30 focal seizures appear over a small subset of channels close to or at the epileptic focus, a strategy to

1 overcome the problem that different focal seizures appear on different channels is required. Such a
2 strategy that successfully tackles the aforementioned problem has already been proposed in the
3 literature [42]. In order to remove the information about the spatial location of the seizure from the
4 training set, the authors in [42] proposed a sorting operation on the extracted features that reorders the
5 features from the different channels in the feature vector before feeding it to the classifier.

6 However, since a seizure with focal onset (as manifested electroencephalographically) is
7 always epileptic, here we have implemented a simplified version of a focal-vs-generalized seizure
8 classification rule (as part of the ARMOR project) that automatically detects and labels the focal
9 seizures. The focal-vs-generalized seizure classification rule is part of the online seizure detector
10 [26][27] which performs a per channel analysis followed by the imposition of spatiotemporal
11 constraints before taking the final decision (clear, focal, generalized) for each tested epoch. The
12 classification rule is based on a mimetic approach requiring the seizure to be detected in at least 65% of
13 the channels in order to be characterized as generalized; otherwise it is characterized as focal. Due to
14 the different type of analysis (fusion of channel-based decisions versus fusion of features per channel to
15 reach a decision), we are not presenting results of focal seizure classification in this paper, but rather
16 focus on the classification of generalized events, which is the last component in our seizure analysis
17 framework.

18 For a clinician the differentiation between focal and generalized events is important because it
19 will play a crucial role in the medication/ treatment and general management choices. Such a rule
20 (appearance in at least 65% of channels) is mainly useful when the events are focal, since focal EEG
21 onset always indicates focal epileptic seizure activity. This rule has no clinical utility to the other event
22 types, since the EEG expression of both psychogenic non-epileptic seizures and vasovagal syncope
23 which leads to impairment of consciousness are “generalized”. However, this step was introduced to
24 facilitate the solution methodologically. Upon the characterization of focal events, the method
25 presented here can be used to discriminate the remaining events into epileptic or non-epileptic.

26 On the clinical usefulness front, it is true that a competent seizure detection algorithm or set of
27 algorithms should be able to detect both focal and generalized seizures, and either of these from non-
28 epileptic events. The reason is that impairment of consciousness can be seen in temporal lobe seizures
29 or seizures with secondary generalization. However, initial prodromal clinical symptoms and some
30 typical EEG characteristics can be used for the differential diagnosis. Due to the big variability of

1 seizure presentation there needs to be a detailed analysis of adequate number of representative cases of
2 different evolving patterns and this will be the part of our next work.

3 Until however we will be able to evaluate the method more extensively on a large dataset with an
4 adequate number of representative cases for focal seizures we applied the proposed methodology on a
5 dataset of 9 patients (2 subjects with focal seizures, 5 subjects with PNES and 2 subjects with VVS).
6 We developed also a different algorithm to remove the spatial content from the features. The algorithm
7 sorts the features for each channel according to feature type and then extracts the standard deviation of
8 each feature type across channels, and the difference between maximum and minimum values of each
9 feature (max-min). We introduced the standard deviation and the max-min values to the BayesNet
10 classifier and achieved 74% accuracy, 70% sensitivity and 76% specificity when the 20 best features
11 are used. The above results were obtained with a leave-one-patient-out strategy for validation. Since
12 the dataset size is small we assessed the method also in a leave-one-epoch-out strategy, as performed in
13 some other studies and achieved 90.2% accuracy, 86.7% sensitivity and 91.5% specificity. The
14 achieved accuracy in this case is much higher, as expected. However, we do not emphasize the
15 importance of these results since they might not generalize to other data.

16 Although direct comparison with other studies is not possible due to the different characteristics
17 of each dataset (e.g. different seizure types, lack of PNES or VVS examples in most studies or use of
18 single channel data), the achieved classification accuracy is higher than the one reported in the
19 literature. In particular, the achieved accuracy in [22] is 86%, lower than the accuracy of BayesNet in
20 our methodology (95%). Furthermore, in [21] the reported sensitivity (83%) and specificity (98%) are
21 lower than the sensitivity of the majority of the classification methods evaluated in our work and the
22 specificity achieved by our framework (98%) when a subset of 10 discriminative features are used with
23 respect to BayesNet classification.

24 Finally, although an initial work was held to reliably solve the problem of discrimination between
25 different types of paroxysmal event and reveal the most discriminative features from a large set of
26 time and frequency domain features given a dataset of 11 patients, there are some limitations that
27 should be taken into account. The number of non-epileptic examples especially those of VVS) is
28 limited and might not capture well the variability of the corresponding EEG events while the available
29 generalized spike waves seem to be enough to describe such a consistent group of patterns. Under this
30 scope we plan to start EEG recordings during tilt table test, which provokes VVS and therefore we

1 shall have a substantial number for further analysis. Furthermore, we aim to perform a more in depth
2 analysis of focal seizures.

3 **4. Conclusions**

4 In this paper, we investigated the problem of classification between epileptic and non-epileptic
5 events from multi-channel EEG data using a large number of time-domain and frequency domain
6 features. The proposed methodology was evaluated in EEG data from 11 subjects. Examination of
7 several classification algorithms showed that the best accuracy is achieved by BayesNet. Feature
8 ranking investigation and evaluation of the classification models using subsets of features were
9 performed and revealed the most significant features for the classification task. The use of the most
10 discriminative features ($N = 10$) increased significantly the performance of BayesNet classification at
11 95% accuracy (94% sensitivity for 98% specificity). The method has been evaluated using cross-
12 validation across subjects and showed that it can generalize satisfactorily providing the means for
13 diagnosis support.

14

15 **5. Acknowledgement**

16 This study is partially funded by the EC under the FP7/2007-2013 with grant ARMOR,
17 Agreement Number 287720. This research has been co-financed by the European Union (European
18 Social Fund – ESF) and Greek national funds through the Operational Program "Education and
19 Lifelong Learning" of the NSRF - Research Funding Program: Thales. Investing in knowledge society
20 through the European Social Fund.

21

22 **6. References**

- 23 [1] Al. McKeon, C. Vaughan and N. Delanty, "Seizure versus syncope," vol. 5, *Lancet Neurol.*, 2006,
24 pp. 171-80.
- 25 [2] F. Brigo, R. Nardone and L.G. Bongiovanni, "Value of tongue biting in the differential diagnosis
26 between epileptic seizures and syncope," vol. 21, *Seizure*, 2012, pp. 568-572.
- 27 [3] J.T. Lin, D.K. Ziegler, C.W. Lai, W. Bayer, "Convulsive syncope in blood donors," vol. 11,
28 *Annals of Neurology*, 1982, pp. 525–28.

- 1 [4] W. Wieling, R.D. Thijs, N. van Dijk, A.A. Wilde, D.G. Benditt and J.G. van Dijk, “Symptoms and
2 signs of syncope: a review of the link between physiology and clinical clues,” vol. 132, *Brain*,
3 2009, pp. 2630–2642.
- 4 [5] W.A.J. Hoefnagels, G.W. Padberg, J. Overweg and R.A.C. Roos, “Syncope or seizure? A matter
5 of opinion,” vol. 94, *Clinical Neurology and Neurosurgery*, 1992, pp.153–156.
- 6 [6] R. S. Fisher, W. V. E. Boas, W. Blume, C. Elger, P. Genton, P. Lee and J. Engel, “Epileptic
7 seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE)
8 and the International Bureau for Epilepsy (IBE),” vol. 46, *Epilepsia*, 2005, pp. 470-472.
- 9 [7] L. Szabo, “Semiologic analysis and classification of childhood psychogenic non-epileptic seizures
10 based on Video-EEG monitoring data,” PhD thesis, 2013.
- 11 [8] T. Lewis, “Vasovagal syncope and the carotid sinus mechanism,” vol. 1, *The British Medical*
12 *Journal*, 1932, pp. 873-876.
- 13 [9] F. Ammirati, F. Colivicchi, G. Di Battista G, F.F. Garelli, C. Pandozi and M. Santini, “Variable
14 cerebral dysfunction during tilt induced vasovagal syncope,” vol. 21, *Pacing and Clinical*
15 *Electrophysiology*, 1998, pp. 2420-5.
- 16 [10] O. Mecarelli, P. Pulitano, E. Vicenzini, N. Vanacore, N. Accornero and M. De Marinis,
17 “Observations on EEG patterns in neurally-mediated syncope: an inspective and quantitative
18 study,” vol. 34, *Clinical Neurophysiology*, 2004, pp. 203-207.
- 19 [11] B. Greene, G. Boylan, R. Reilly, P. Chazal and S. Connolly, “Combination of EEG and ECG for
20 improved automatic neonatal seizure detection,” vol. 118, *Clinical Neurophysiology*, 2007, pp.
21 1348-1359.
- 22 [12] Meier, H. Dittrich, A. Schulze-Bonhage and A. Aertsen, “Detecting epileptic seizures in long-term
23 human EEG: a new approach to automatic online and real time detection and classification of
24 polymorphic seizure patterns,” vol. 25, *Clinical Neurophysiology*, 2008, pp. 119-131.
- 25 [13] A. Shoeb, H. Edwards, J. Connolly, B. Bourgeois, S.T. Treves and J. Guttag, “Patient-specific
26 seizure onset detection,” vol. 5, *Epilepsy Behavior*, 2004, pp. 483-98.
- 27 [14] A. Shoeb and J. Guttag, “Application of machine learning to epileptic seizure onset detection,”
28 27th International Conference on Machine Learning (ICML), 2010.

- 1 [15] M. Valderrama, S. Nikolopoulos, C. Adam, V. Navarro and M. Van Quyen, "Patient-specific
2 seizure prediction using a multi-feature and multi-modal EEG-ECG classification," In Proc. of
3 IFMBE, 2010.
- 4 [16] M.Z. Parvez and M. Paul, "Epileptic seizure detection by analyzing EEG signals using different
5 transformation techniques," vol. 145, Neurocomputing, 2014, pp.190-200.
- 6 [17] ARMOR project. Web-site: [http:// www.armor-project.eu/](http://www.armor-project.eu/).
- 7 [18] M. Poulos, F. Geogiadodis, V. Chrissicopoulos and A. Evangelou, "Diagnostic test for the
8 discrimination between interictal epileptic and non-epileptic pathological EEG events using auto-
9 cross-correlation methods," vol. 43, American Journal of Electroneurodiagnostic Technology,
10 2003, pp. 228-264.
- 11 [19] S. Papavlasopoulos, M. Poulos and A. Evangelou, "Feature extraction from interictal epileptic and
12 non-epileptic pathological EEG events for diagnostic purposes using LVQ1 neural network,"
13 Proceedings of 7th International Conference on Mathematics Methods in Scattering Theory and
14 Biomedical Technology, BIOTECH'7, 2005.
- 15 [20] S. Papavlasopoulos, M. Poulos, G. Bokos and A. Evangelou, "Classification control for
16 discrimination between interictal epileptic and non-epileptic pathological EEG events," vol. 1,
17 International Journal of Medical World Academy of Science, Engineering and Technology,
18 Health, Pharmaceutical and Biomedical Engineering, 2007.
- 19 [21] H.R. Mohseni, A. Maghsoudi and M.B. Shamsollahi, "Seizure detection in EEG signals: a
20 comparison of different approaches," In Conf. Proc IEEE Eng Med Biol Soc. ,2006, 6724-7.
- 21 [22] S. Nasehi and H. Pourghassem , " Seizure Detection Algorithms Based on Analysis of EEG and
22 ECG Signals: a Survey," vol. 44, Neurophysiology, 201, pp. 174-186.
- 23 [23] S. Tong and N.V. Thankor , " Quantitative EEG Analysis Methods and Applications,"
24 Engineering in Medicine & Biology, 2009.
- 25 [24] B. Hunya, M. Signoreto, W. Van Paesschen, J.A. Suykens, S. Van Huffel and M. De Vos,
26 "Incorporating structural information from the multichannel EEG improves patient-specific
27 seizure detection," vol. 123, Clinical Neurophysiology , pp. 2352-2361.
- 28 [25] E. Asanoa, C. Pawlak, A. Shah, J. Shah, A.F. Luat, J. Ahn-Ewing, H.T. Chugani, "The diagnostic
29 value of initial video-EEG monitoring in children—Review of 1000 cases," Epilepsy Research
30 66, 2005, pp. 129–135.

- 1 [26] I. Mporas, V. Tsirka, E. Zacharaki, M. Koutroumanidis and V. Megalooikonomou, "Online
2 Seizure Detection from EEG and ECG signals for Monitoring of Epileptic Patients," 8th Hellenic
3 Conference on Artificial Intelligence (SETN), Lecture Notes in Computer Science, vol. 8445,
4 2014, pp 442-447.
- 5 [27] I. Mporas, V. Tsirka, E.I. Zacharaki, M. Koutroumanidis and V. Megalooikonomou, "Evaluation
6 of Time and Frequency Domain Features for Seizure Detection from Combined EEG and ECG
7 signals," 7th Int. Conf. on Pervasive Technologies Related to Assistive Environments (PETRA),
8 2014.
- 9 [28] B. Kemp and J. Olivan, "European data format 'plus' (EDF+), an EDF alike standard format for
10 the exchange of physiological data," vol. 114, Clinical Neurophysiology, 2003 pp. 1755–1761.
- 11 [29] G.F. Cooper and E. Herskovits, "A Bayesian method for constructing Bayesian belief networks
12 from databases," 1990.
- 13 [30] G. Cooper and E. Herskovits, "A Bayesian method for the induction of probabilistic networks
14 from data," vol. 9, Machine Learning, 1992, pp. 309-347.
- 15 [31] L. Breiman, "Random Forests," vol. 45, Machine Learning, 2001, pp.5-32.
- 16 [32] D. Aha and D. Kibler, "Instance-based learning algorithms," vol. 6, Machine Learning, 1991, pp.
17 37-66.
- 18 [33] S.S. Keerthi, S.K. Shevade, C. Bhattacharyya and K.R.K. Murthy, "Improvements to Platt's SMO
19 algorithm for SVM classifier design," vol. 13, Neural Computation, 2001, pp. 637-649.
- 20 [34] J. Platt, "Fast Training of Support Vector Machines using Sequential Minimal Optimization,"
21 Advances in Kernel Methods - Support Vector Learning, 1998.
- 22 [35] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I. H. Witten, "The WEKA Data
23 Mining Software: An Update," vol. 11, SIGKDD Explorations, 2009.
- 24 [36] I. Kononenko, "Estimating attributes: Analysis and extension of RELIEF," vol. 784, Proc. Euro.
25 Conf. Mach. Learn., 1994, pp. 171– 182.
- 26 [37] K. Kira and L. A. Rendell, "A practical approach to feature selection," Proc. 9th Int. Conf. Mach.
27 Learn., 1992, pp. 249 – 256.
- 28 [38] T. G. Dietterich, "Machine learning research: Four current directions," vol. 18, Artificial
29 Intelligence Magazines, 1997, pp. 97–136.

1 [39] Y. Sun and J. Li, "Iterative RELIEF for feature weighting," Proc. 21st Int. Conf. Mach. Learn.,
2 2006, pp. 913–920.

3 [40] Y. Sun and D.Wu, "A RELIEF based feature extraction algorithm," In Proceedings of SIAM
4 International Conference on Data Mining, 2008.

5 [41] I. Kononenko, E.Simec and M. Robnik-Sikonja, "Overcoming the Myopic of Inductive Learning
6 Algorithms with RELIEF-F," Applied Intelligence, 1997

7 [42] B. R. Greene, W. P. Marnane, G. Lightbody, R. B. Reilly and G. B. Boylan, "Classifier models
8 and architectures for EEG-based neonatal seizure detection," vol. 29, Physiological
9 measurement, 2008

10
11
12

| Subject | Class | Number of Epochs | Number of Seizures |
|----------------|--------------|-------------------------|---------------------------|
| 1 | GSW | 59 | 52 |
| 2 | GSW | 29 | 19 |
| 3 | GSW | 16 | 14 |
| 4 | GSW | 19 | 20 |
| 5 | PNES | 1 | 1 |
| 6 | PNES | 1 | 1 |
| 7 | PNES | 1 | 1 |
| 8 | PNES | 13 | 13 |
| 9 | PNES | 3 | 3 |
| 10 | VVS | 45 | 1 |
| 11 | VVS | 18 | 1 |

Table 1 Number of seizures and number of seizure epochs (2 seconds) per subject

1
2

| Classification Model | Statistical Measures before Feature Selection | | | Statistical Measures after Feature Selection | | |
|-----------------------------|--|--------------------|--------------------|---|--------------------|--------------------|
| | <i>Accuracy</i> | <i>Sensitivity</i> | <i>Specificity</i> | <i>Accuracy</i> | <i>Sensitivity</i> | <i>Specificity</i> |
| BayesNet | 86% | 92% | 78% | 95% | 94% | 98% |
| RandomCommittee | 83% | 88% | 77% | 92% | 89% | 77% |
| RandomForest | 74% | 77% | 70% | 87% | 92% | 79% |
| IBk | 69% | 86% | 43% | 86% | 94% | 76% |
| SMO (RBF kernel) | 68% | 55% | 87% | 87% | 85% | 91% |

Table 2 Classification performance before and after Feature Selection.

3
4

| Feature | Channels |
|----------------|--|
| nmin | Fp1, T4, T5, Fz, F7, Fp2, A1, T6, Cz, O1, F4, F3, F8, P4 |
| nmax | Fp1, T4, T5, F7, Fp2, P3, Fz, T3, T6, A1, Cz, O1, F4, F3, F8, P4 |
| max | Fp1 |
| std | Fp1 |
| aryule3 | T6, T4, F4, C4, F7, T5 |
| maxfreq | T3, O2 |
| aryule2 | T4, T5, P4, Fp1, O1 |
| loge | C4 |
| cwt25 | A1 |
| cwt32 | A1, Fp2 |
| minfreq | F7 |

1 **Table 3 Best features according to the top 50 ranking of STRATEGY 1 and the channels they appear on.**

2

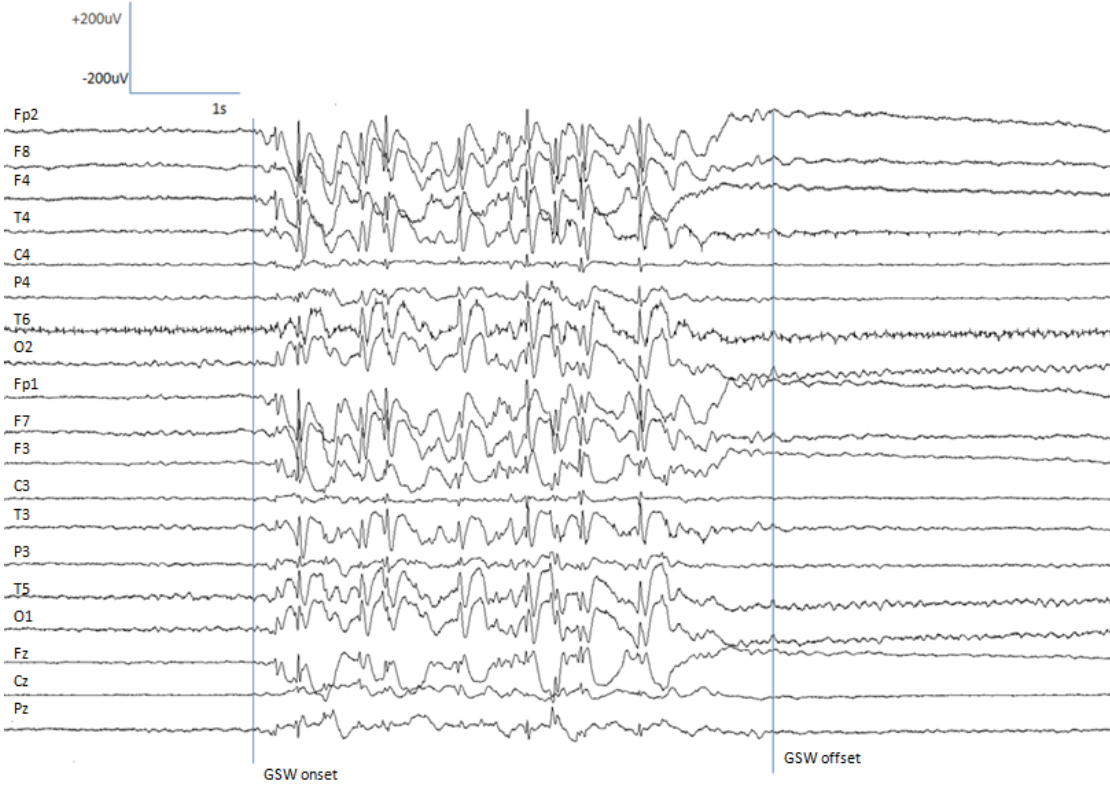
| Feature | Channels |
|----------------|--|
| nmax | T4, Fp1, T5, F7, Fp2, Fz, T3, O1, P3, T6, F3, F8, F4, A1, Cz, P4 |
| nmin | T4, Fp1, T5, F7, Fp2, Fz, T3, O1, P3, T6, F3, F8, F4, A1, Cz, P4 |
| aryule3 | T4, F4, T6, C4, T5, F7, O1 |
| aryule2 | T4, T3, T5, Fp2, F4, P4 |
| Minfreq | T4 |
| cwt32 | A1, Fp2 |
| cwt25 | A1, Fp2 |

3 **Table 4 Best features according to the top 50 ranking of STRATEGY 2 and the channels they appear on.**

4

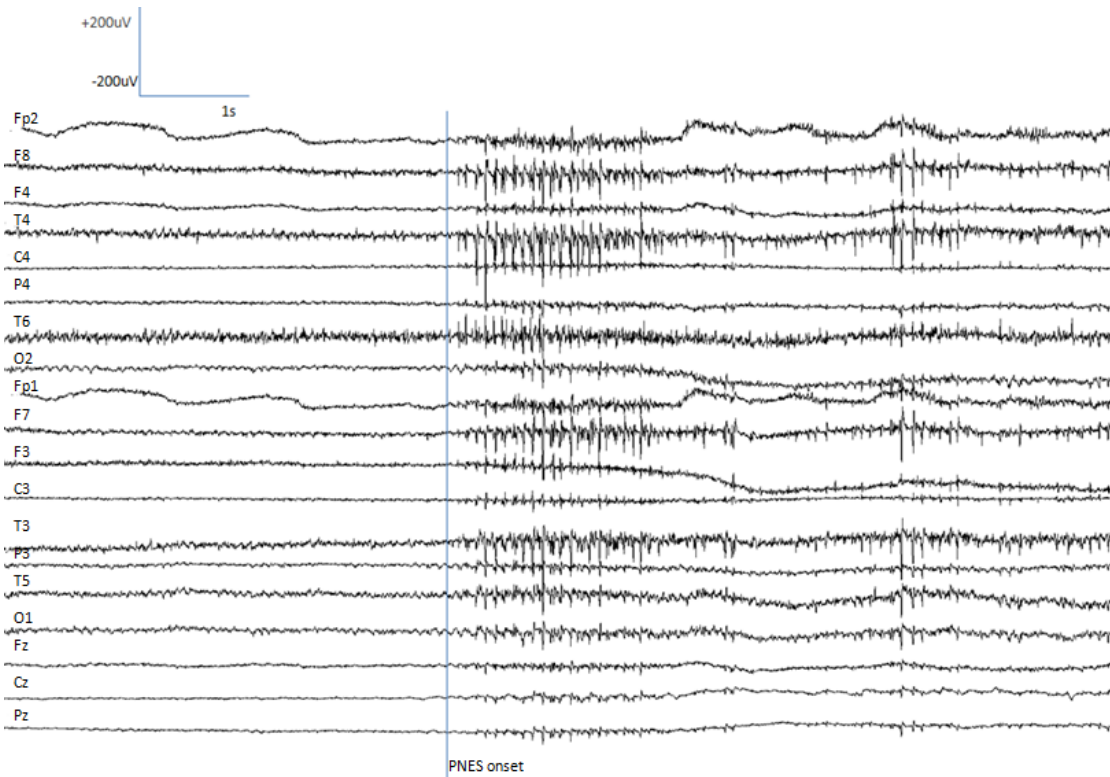
5

1



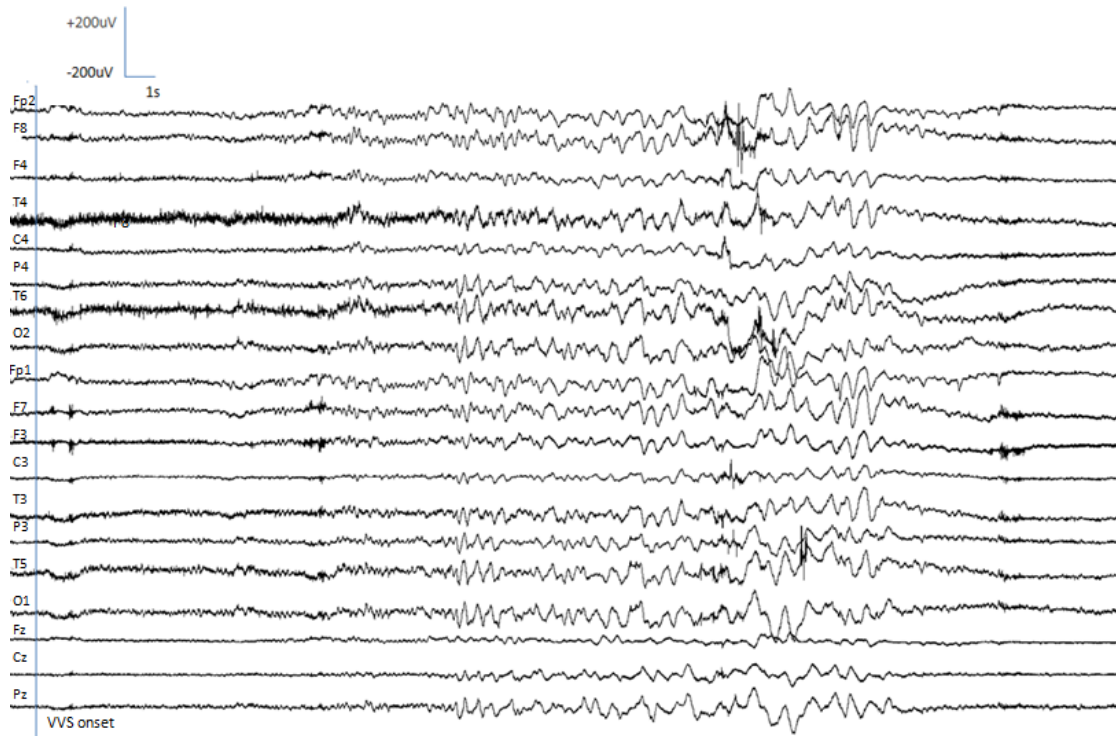
2

3 **Figure 1 Generalized Spike Wave (GSW) example. The first marker indicates the beginning of the GSW**
4 **event and the second marker its end.**



5

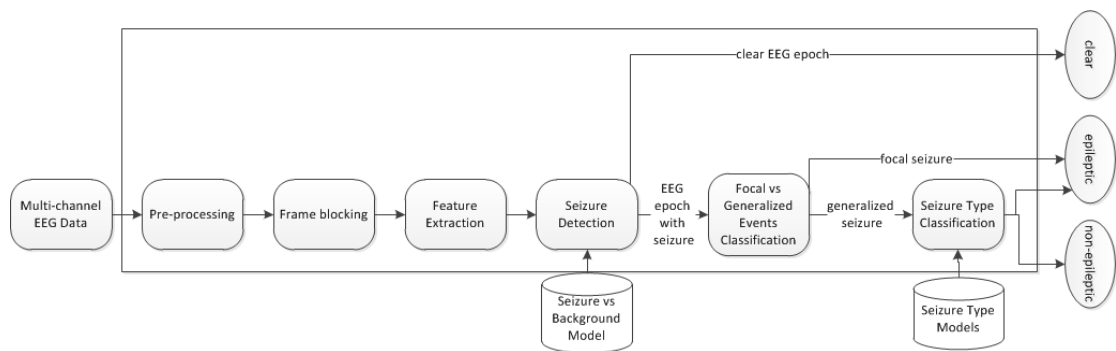
6 **Figure 2 Psychogenic Non Epileptic Seizure (PNES) example. The marker indicates the beginning of the**
7 **PNES event.**



1

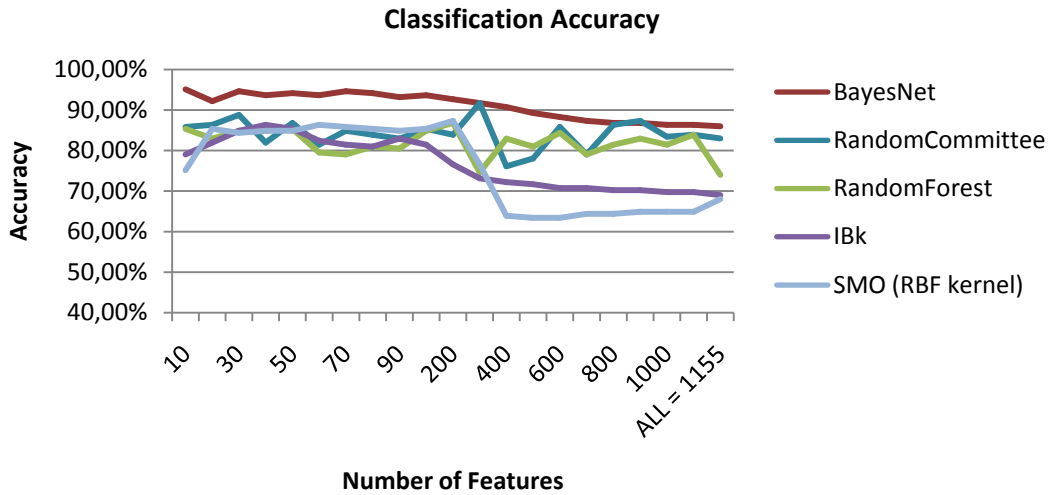
2 **Figure 3 Vasovagal Syncopal Event (VVS) example. The marker indicates the beginning of the VVS event.**

3

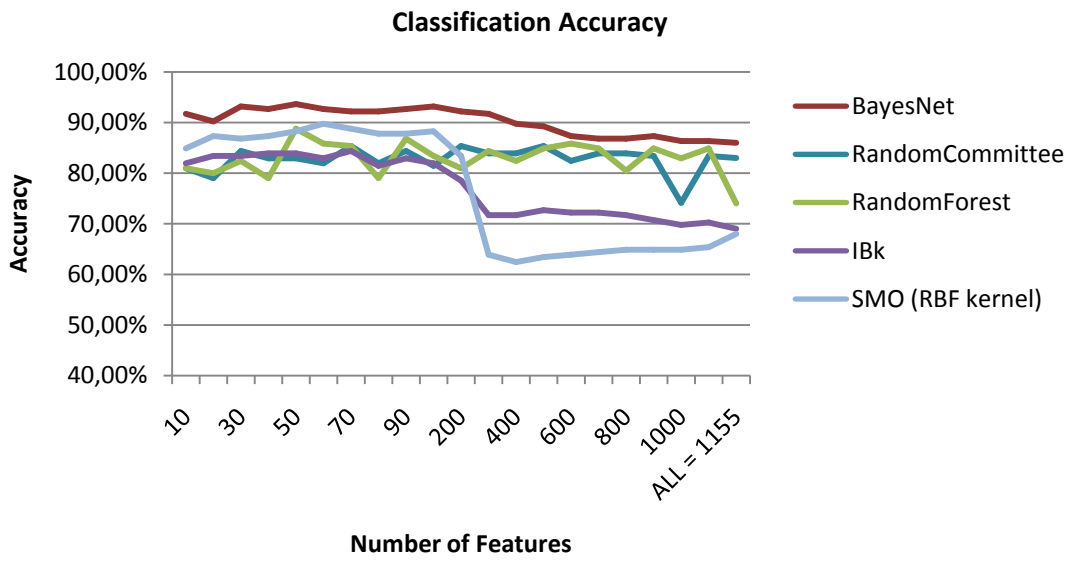


4

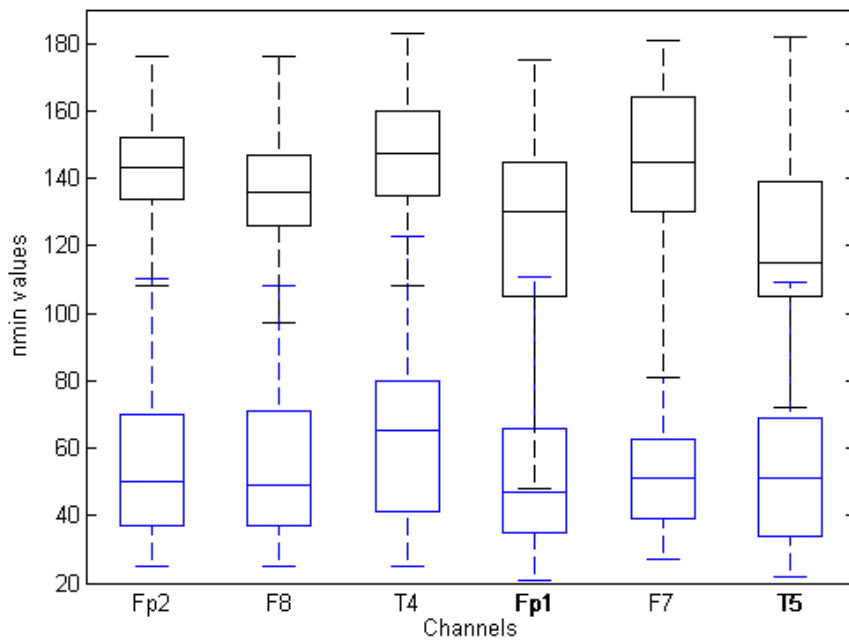
5 **Figure 4. The concept of seizure detection and classification within the ARMOR framework.**



1
2 **Figure 5 Classification Accuracy when 1st Ranking Strategy (based on Frequency) is used.**

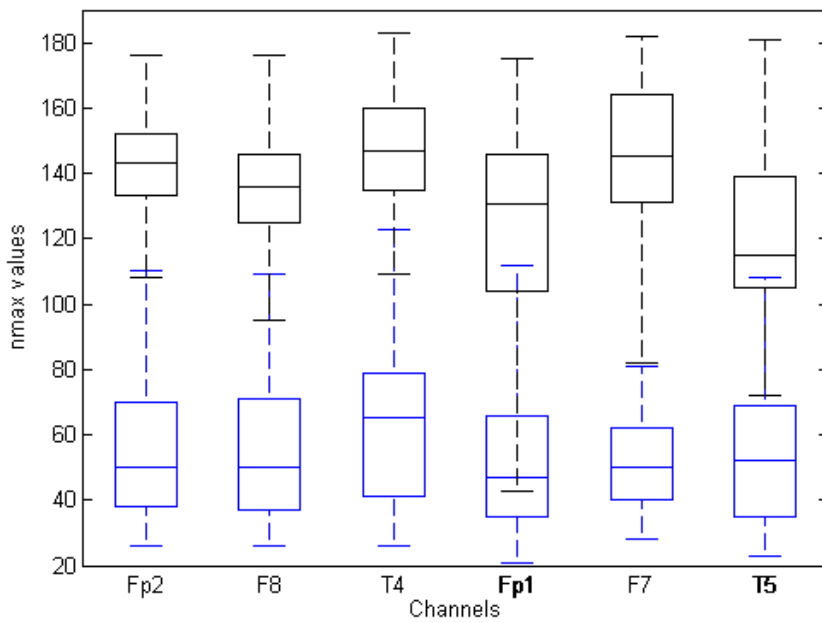


3
4 **Figure 6 Classification Accuracy when 2nd Ranking Strategy (based on Sum of ReliefF Weights) is used.**



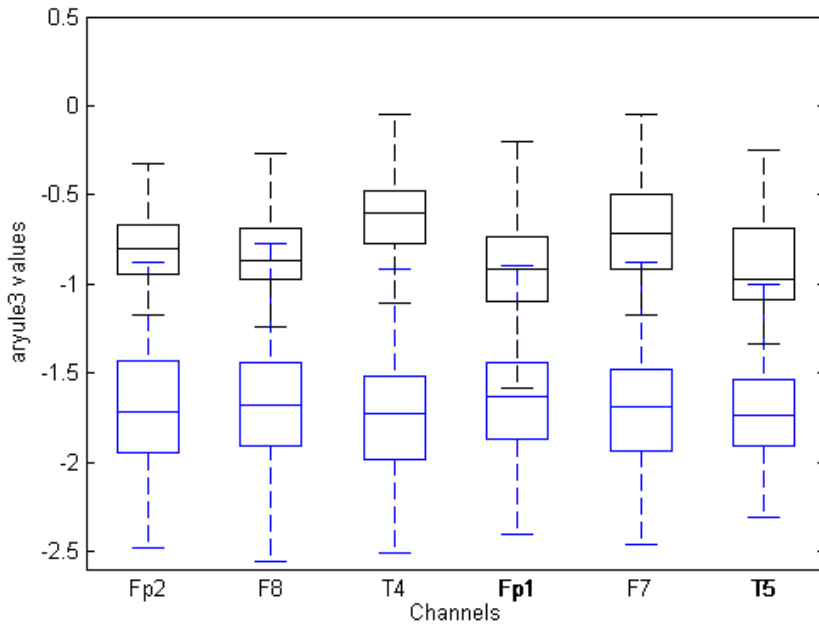
1

2 **Figure 7 Values of the feature *nmin* of epileptic (blue boxes) and non-epileptic (black boxes) EEG epochs.**



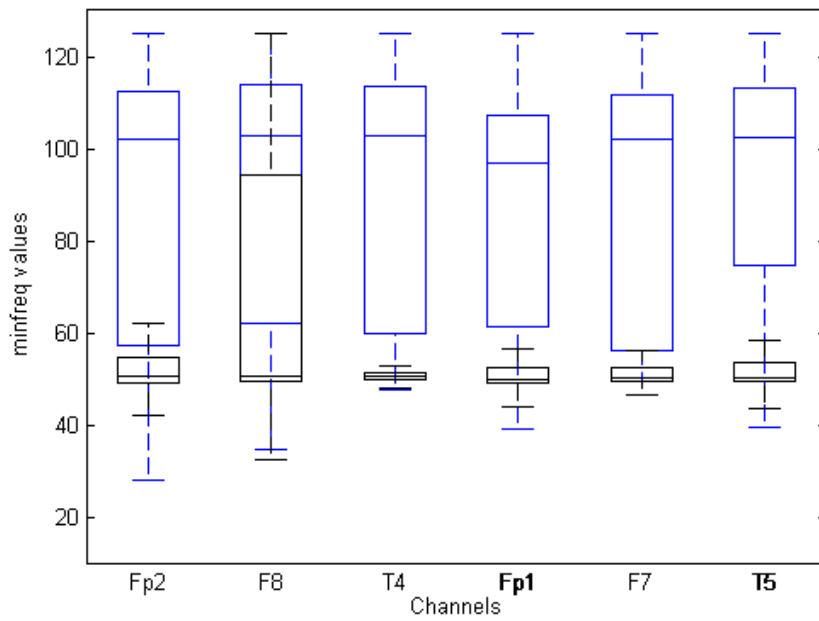
3

4 **Figure 8 Values of the feature *nmax* of epileptic (blue boxes) and non-epileptic (black boxes) EEG epochs.**



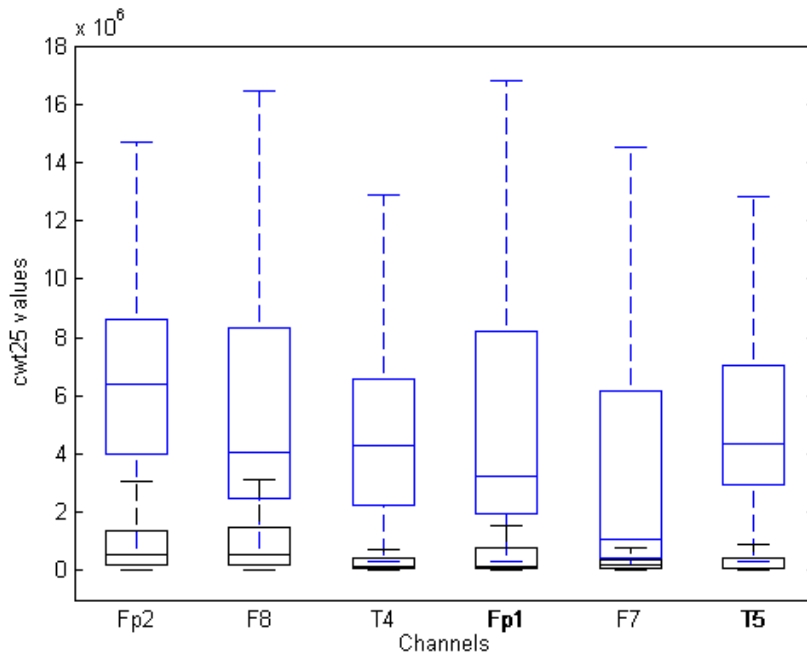
1

2 **Figure 9** Values of the feature *aryule3* of epileptic (blue boxes) and non-epileptic (black boxes) EEG epochs.



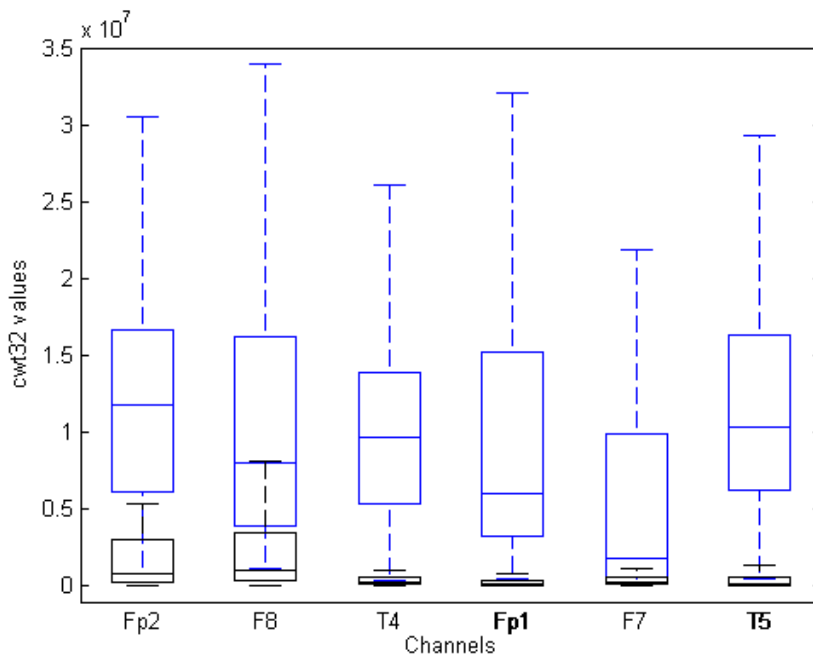
3

4 **Figure 10** Values of the feature *minfreq* of epileptic (blue boxes) and non-epileptic (black boxes) EEG
5 epochs.



1

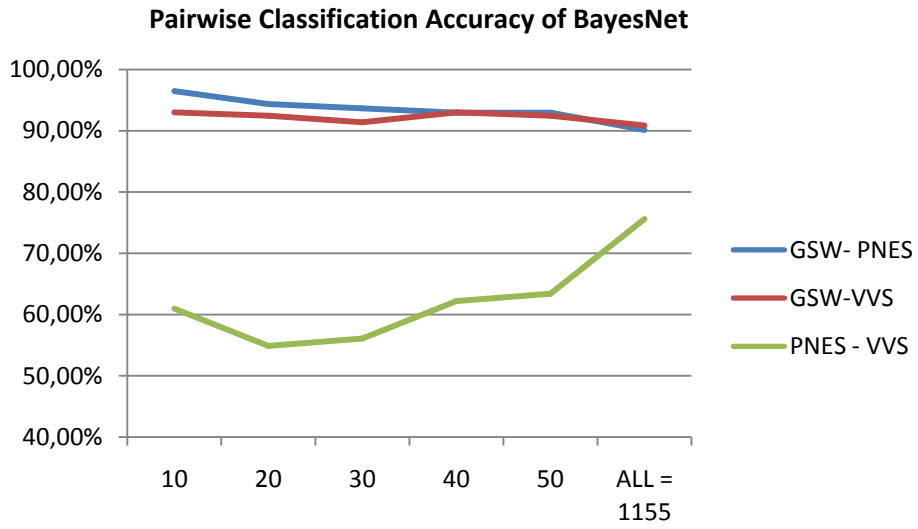
2 **Figure 11** Values of the feature *cwt25* of epileptic (blue boxes) and non-epileptic (black boxes) EEG epochs.



3

4 **Figure 12** Values of the feature *cwt32* of epileptic (blue boxes) and non-epileptic (black boxes) EEG epochs.

5



1

2 **Figure 13** Pairwise classification accuracy of BayesNet classifier for different number of best features.

3