EEG Transient Event Detection and Classification Using Association Rules

Article in IEEE Transactions on Information Technology in Biomedicine · August 2006

DOI: 10.1109/TITB.2006.872067 · Source: PubMed

5 authors, including:

Alexandros T Tzallas
University of Ioannina

Dimitrios I Fotiadis
University of Ioannina

Sotirios Giannopoulos
University of Ioannina

EEG Transient Event Detection and Classification Using Association Rules

Themis P. Exarchos, Alexandros T. Tzallas, Dimitrios I. Fotiadis, *Member, IEEE*, Spiros Konitsiotis, and Sotirios Giannopoulos

***Abstract*—In this paper, a methodology for the automated detection and classification of transient events in electroencephalographic (EEG) recordings is presented. It is based on association rule mining and classifies transient events into four categories: epileptic spikes, muscle activity, eye blinking activity, and sharp alpha activity. The methodology involves four stages: 1) transient event detection; 2) clustering of transient events and feature extraction; 3) feature discretization and feature subset selection; and 4) association rule mining and classification of transient events. The methodology is evaluated using 25 EEG recordings, and the best obtained accuracy was 87.38%. The proposed approach combines high accuracy with the ability to provide interpretation for the decisions made, since it is based on a set of association rules.**

***Index Terms*—Association rules, clustering, electroencephalographic (EEG), epilepsy, spike detection, transient events.**

I. Introduction

EPILEPSY IS a neurological disorder which results in recurrent seizures. Epilepsy is characterized by the existence of abnormal synchronous discharges in large ensembles of neurons in brain structures. These discharges are often referred to as "paroxysmal activity," and appear either during seizures (ictal periods) or between seizures (interictal periods). Short recordings of interictal periods are commonly examined by the neurologists in daily practice. The most common forms of interictal activity, realized in the electroencephalographic (EEG), are the individual spike, the sharp wave, and the spike-and-wave complex. The aforementioned spikes are observed in the majority of patients with epilepsy, so spike detection and assessment can be very supportive to the neurologists in the diagnosis of epilepsy [1], [2].

In this paper, classification using association rules [3] is proposed for the first time in the literature for the classification of epileptic spikes and other transient events in EEG recordings. Classification with the use of association rules is a technique which provides high accuracy and interpretation for the obtained decisions [3], [4]. The proposed methodology performs automated detection and classification of four types of transient events in EEG recordings: epileptic spikes (ESs), muscle activity (EMG), eye blinking activity (EOG), and sharp alpha activity (SAA), and involves signal preprocessing, transient events detection, transient events clustering, feature extraction, feature discretization, feature subset selection, and classification using association rules.

Our approach introduces several novel features: 1) the employment of association rules for EEG analysis offers the potential of discovering new knowledge in the form of rules; 3) it is able to handle the undesirable activity (artifacts, SAA) by detecting and classifying it to the appropriate categories successfully; and 4) the use of clustering emulates the clinical procedure of EEG reading in daily practice followed by a neurologist. In this routine reading of EEGs, the clinician rapidly browses through the recording, discarding groups of artifacts like EMG, EOG, SAA, etc. based solely on their appearance (shape, duration, etc.) [1], [5].

Moreover, the rule based nature of our methodology gives the ability to provide interpretation for the classification decisions and makes the decision-making process transparent. In medical applications, the ability to explain the reason for a decision is of great value for domain experts [6]. Approaches where the decision is performed in a "hidden" way (e.g., neural networks) could reduce the confidence of both medical experts and patients in the method.

II. Related Work

Mimetic methods, which are based on the general concept that automatic EEG analysis should mirror the visual analysis performed by an expert in the daily practice [7]; techniques which distinguish transients from ongoing background activity [8]; and template based techniques which detect events matching previously selected spikes [9] have been proposed in the literature to address the detection of epileptic spikes in EEG recordings. In addition, some methods use techniques that roughly follow the definition of a spike which was loosely defined by Gloor [10], and emphasize the local context [5] and morphology [11]. Approaches employing artificial neural networks [1], [12], [13], decision trees [14], Bayesian classifiers [14], and other mathematical approaches have also been proposed [15].

The application of association rule mining techniques [16] in the biomedical signal analysis domain has been previously addressed. Specifically, a methodology for mining reproducible activation patterns in epileptic intracerebral EEG signals has

Manuscript received July 25, 2005; revised November 30, 2005, December 9, 2005, and December 15, 2005. This work was supported in part by the European Commission NOESIS Project under Grant IST-2002-507960, Platform for Wide Scale Integration and Visual Representation of Medical Intelligence.

T. P. Exarchos and A. T. Tzallas are with the Unit of Medical Technology and Intelligent Information Systems, Department of Computer Science, University of Ioannina, GR 45110 Ioannina, Greece (e-mail: me01238@cc.uoi.gr; me00716@cc.uoi.gr).

D. I. Fotiadis is with the Unit of Medical Technology and Intelligent Information Systems, Department of Computer Science, University of Ioannina, and also with the Biomedical Research Institute, FORTH, GR 45110 Ioannina, Greece (e-mail: fotiadis@cs.uoi.gr).

S. Konitsiotis and S. Giannopoulos are with the Department of Neurology, Medical School, University of Ioannina, GR 45110 Ioannina, Greece (e-mail: skonitso@cc.uoi.gr; sgiannop@cc.uoi.gr).

Digital Object Identifier 10.1109/TITB.2006.872067

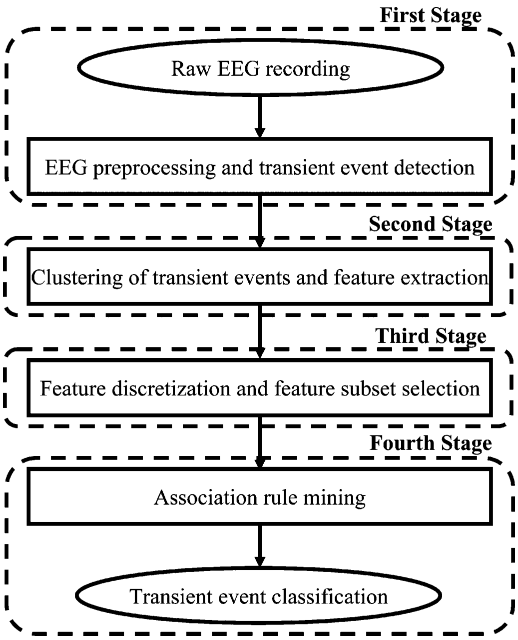


Fig. 1. Flowchart of the proposed methodology.

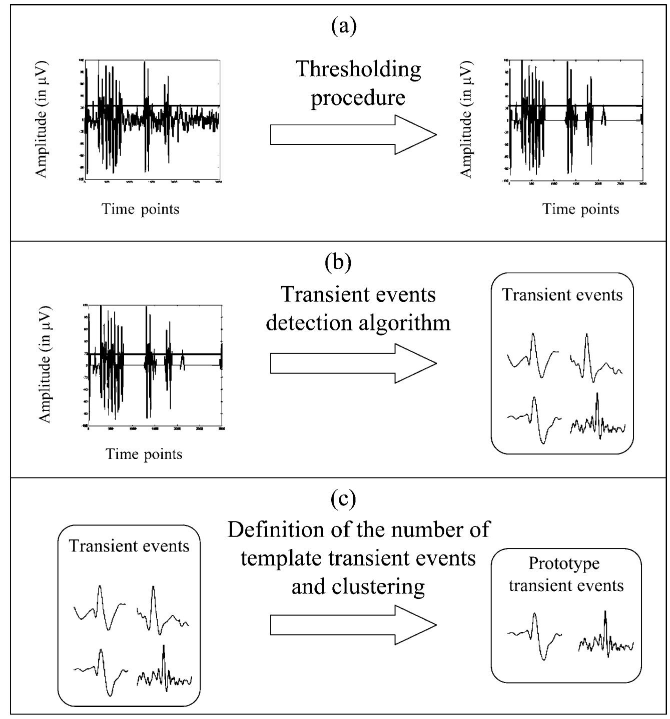


Fig. 2. (a) A magnified segment of F3-C3 channel from an epileptic patient before and after the use of thresholding procedure. (b) Schematic of the first stage of the proposed methodology: signal preprocessing and transient event detection. (c) Schematic of the second stage: clustering of transients events.

been presented in [17], and a rule discovery algorithm appropriate for mining sequential rules in the ECG domain has been reported in [18]. A review of data mining algorithms applied to EEG is presented in [19].

Clustering in the field of EEG analysis has also been addressed. Hierarchical agglomerative methods and self organizing maps have been used for clustering EEG segments [20]. The nearest mean algorithm [21] and the fuzzy k-means [22] algorithm have been employed in order to cluster ESs. In addition, the k-means algorithm has been used in order to cluster spikes and other types of transient events [23]. However, most of the proposed approaches deal only with spike detection, and only a few deal with the classification of transient events [23], [24]. Our approach addresses classification of transient events.

III. Methodology for Rule-Based Classification of Transient Events in EEG

A. Methodology Description

A methodology based on a four-stage schema was developed for transient event detection and classification (Fig. 1). In the first stage, preprocessing of the EEG signal was performed in order to detect the transient events. In the second stage, the transient events were clustered resulting in prototype transient events, and sixteen representative features were extracted from each prototype. In the third stage, the extracted continuous valued features were transformed into discrete ones, and the most consistent feature subset was identified. In the last stage, classification using association rules was realized for transient event categorization.

1) First Stage–Signal Preprocessing and Transient Event Detection: A segmentation algorithm was used to eliminate the areas of low background activity and detect transient events in the EEG recordings [13]. This procedure is data-driven, and enhances the adaptability of our methodology to different EEG signals. The areas of low background activity were eliminated using a threshold t which depends on the mean absolute value of the original EEG signal in each single channel of each montage. The threshold [Fig. 2(a)] was calculated as

$$t = \frac{1}{\mathrm{N}} \sum_{i=1}^{\mathrm{N}} |x_i| \tag{1}$$

where x_i represents the values of the signal and N is the number of samples of the signal. Peaks over the calculated threshold t were considered as transient events. A window with constant length of 91 data points (∼355 ms) [14], [23] was applied, centered at each identified peak (the reason for using the specified window length is that spikes have durations of 20–70 ms, sharp waves last 70–200 ms, EMG waves are less than 30 ms, and EOG waves last more than 150 ms [2]). Finally, if a larger peak (vertex) was found in the window, the window was centred at this peak; otherwise, the window of 91 signal points was considered as a transient event [Fig. 2(b)]

2) Second Stage: Clustering of Transient Events and Feature Extraction:

a) Clustering of Transient Events: The transient events existing in the EEG recordings are the input in the clustering procedure and the prototype transient events are the output [22], [23]. This clustering assigns each transient event to a cluster and is used in order to emulate the procedure that the neurologists follow during the examination of EEG recordings [1], [5]. In addition, with this clustering, we discover the distinct groups of transient events in an EEG recording [22]. In order to identify the number of clusters of transient events and to cluster them into prototype groups, the minimization of the regularized cost function $C(x,y)$ [25] is employed

$$C(x,y) = \sum_{i=1}^{p}\sum_{j=1}^{k} f(y^{(j)}|x^{(i)})\|x^{(i)} - y^{(j)}\|^2 + \textstyle\sum_{i=1}^{p}\sum_{j=1}^{k} \tilde{I}_j \tilde{f}(y^{(j)}|x^{(i)})\|y^{(j)} - y^{(\omega)}\|^2 \quad (2)$$

where $\tilde{I}_j$ is a multiplier unique to each cluster center, and $f(y^{(j)}|x^{(i)}), \tilde{f}(y^{(j)}|x^{(i)})$ are indicator functions given as

$$f(y^{(j)}|x^{(i)}) = \begin{cases} 1, & \text{if} \quad j = \arg\min_m \|x^{(i)} - y^{(m)}\|^2 \\ 0, & \text{otherwise} \end{cases}$$

and

$$\tilde{f}(y^{(j)}|x^{(i)}) = \begin{cases} 1, & \text{if} \quad y^{(i)} \in N_{y^{(\omega)}}, \omega = \arg\min_m \|x^{(i)} - y^{(m)}\|^2 \\ 0, & \text{otherwise} \end{cases}$$

where $N_{y^{(\omega)}}$ is the neighborhood of the cluster center $y^{(\omega)}, x^{(i)} \in \Re^n$ is a pattern (in our case a transient event), p is the number of patterns $\{x^{(i)} : i = 1, 2, \ldots, p\}$, k is the number of clusters, and $\arg\min_m \|x^{(i)} - y^{(m)}\|^2$ is the value of m which minimizes the norm $\|x^{(i)} - y^{(m)}\|^2$.

The first term in (2) distributes the cluster centers in order to minimize the sum-of-squared distance from each input pattern (transient event) to the nearest cluster center. The second term of the cost function minimizes the sum-of-squared distances between nearby clusters.

The whole procedure is iterative, and the cost function with the updated cluster centers is calculated at each iteration (clustering epoch). At the end of each clustering epoch, we combine clusters which are close to each other, and remove clusters which do not include a significant number of patterns. This allows the discrimination of the number of clusters for a particular size of the neighborhood $N_{y^{(\omega)}}$ (details can be found in [25]). Repeating the process for various values of the neighborhood, we can identify the number of clusters k. After clustering, each cluster is represented by a prototype, which is computed as the centroid of patterns belonging to that cluster, and only the prototypes are used at further stages [Fig. 2(c)].

b) Feature Extraction: Various features have been suggested to represent EEG waveforms in the literature [23], [26]. In our work, we use sixteen features in order to characterize every prototype transient event. Six of them, namely the duration, the area, the average slope, the sharpness, the standard deviation, and the dominant frequency have been previously used [23]. In our work, ten more features were extracted from the power spectrum density (PSD) of each prototype transient event. More specifically, the PSD of each prototype transient event was divided into ten distinct frequency ranges. The average of the PSD for each range (discrete power spectrum density) is derived and ten features are formed (named *disc PSD 1* to *disc PSD 10*).

3) Third Stage: Discretization and Feature Selection:

a) Discretization: In this stage, the continuous valued features are transformed to discrete ones using a discretization technique [27]. Discretization is required because of the use of the selected classification association rule mining technique in the fourth stage. The discretization approach we used is based on entropy minimization and the minimum description length principle [27] and, as reported in the literature [28], is slightly superior to other discretization methods. We obtained the following discrete values: six feature values for standard deviation, six for average slope, nine for sharpness, five for duration, six for dominant frequency, five for area, six for *disc PSD* 1, five for *disc PSD 2*, five for *disc PDS 3*, three for *disc PSD 4*, three for *disc PDS 5*, three for *disc PSD 6*, eight for *disc PSD 7*, three for *disc PSD 8*, five for *disc PSD 9* and two for *disc PSD 10*.

b) Feature Selection: After the discretization of the sixteen features, in order to achieve dimensionality reduction, we employed a feature selection method [29], [30]. The problem of feature selection is defined as follows: find a minimum set of W relevant features which describes the dataset as well as the original Z features do, where $W \leq Z$ ($Z = 16$ in our case). In many cases, this procedure can provide better classification accuracy [29]. A typical feature subset selection process consists of four basic steps: subset generation, subset evaluation, application of a stopping criterion and result validation.

In general, feature selection methods can be classified into two categories: the filter approach [29], where the feature selector is independent of a learning algorithm and removes irrelevant and/or redundant features; and the wrapper approach [31], where the feature selector works as a wrapper with a learning algorithm, and the subset of relevant features is determined based on its accuracy. We prefer the filter approach, since: 1) the wrapper approach selects a feature subset that depends heavily on the bias of the learning algorithm (choosing relevant features according to a specific learning algorithm is similar to fitting the data to the learning algorithm); 2) the wrapper approach depends on the time complexity of the learning algorithm [31]; and 3) when the dataset is too large, the training of some learning algorithms may be extremely time consuming and require high computational effort.

Another useful taxonomy can be drawn by dividing algorithms into those which evaluate individual features [32] and those which evaluate subsets of features [30]. The first category evaluates the features independently using some criteria, whereas the second evaluates subsets of features. Feature subset evaluation is considered more appropriate, since it takes into account the correlation between the features.

The selection of a specific subset evaluation metric depends heavily on the problem. In our work, we tested several filter based subset evaluation metrics [30]. The best accuracy was reported using the consistency based subset evaluation method [33]. This method selects a feature subset which has the best consistency with the class [33].

Most feature selection algorithms perform a search through the space of feature subsets. Since the size of this space is exponential in the number of features, an exhaustive search is usually intractable. Greedy hill climbing search strategies such as forward selection [34] are often applied to search the feature subset space in reasonable time. Although simple, these searches often yield good results comparable to more sophisticated search strategies. For this reason, forward selection was used in our methodology.

The subset which is generated from the feature selection consists of eight $(W = 8)$ features. These features are: standard deviation, average slope, dominant frequency, duration, area, *disc PSD 3, disc PSD 4* and *disc PSD 9*. These features are the most useful and widely used by the experts to decide on the classification of a transient event [35], [36].

4) Fourth Stage: Classification Using Association Rules: Classification using association rules employs those rules whose consequent is the class [3]. These rules are called class association rules (CARs) and, after their generation, they are used to generate a classification model.

Several algorithms were tested: 1) The classification based on associations (CBA) algorithm [3], which first generates all the CARs with minimum support and minimum confidence thresholds, provided by the user, as candidate rules. A small set of rules from them is selected using rule pruning [37] in order to create the classifier. When predicting the class label for an example, the best rule whose body is satisfied by the example is chosen for prediction. The best rule is the one which has the highest confidence. 2) The classification based on multiple association rules (CMAR) algorithm [4], which generates and evaluates rules in a similar way as CBA. A major difference is that the classification is performed based on a weighted chi-square analysis using multiple rules. 3) The classification based on predictive association rules (CPAR) algorithm [38], which generates a small set of predictive rules from the dataset based on rule prediction and the instances covered by a rule (i.e., instances that satisfy the conditions of a rule). The rule prediction is the accuracy of the rule, and is measured with the Laplace expected error estimate.

The preceding algorithms apply in a similar fashion. They extract class association rules having as antecedent a subset (or all) of the eight features (reduced subset) of every prototype with its respected discrete value and as consequent the class of the prototype. The form of those CARs is

$$\text{IF } \mathrm{F}_1 = \mathrm{DV}_1 \text{ AND} F_2 = \mathrm{DV}_2 \text{ AND} \ldots \text{AND}$$
$$\mathrm{F_M} = \mathrm{DV_N} \text{ THEN Transient Event is } \mathrm{C_i}$$

where $\mathrm{F_i}$ denotes one of the eight features of the most consistent feature subset, $\mathrm{DV_i}$ denotes one of the discrete values of this feature, and $\mathrm{C_i}$ denotes one of the four classes of the transient events (ES, EMG, EOG, SAA).

In addition to these three algorithms, two other algorithms are examined. The repeated incremental pruning to produce error reduction (RIPPER) [39] and the CN2 algorithm [40]. These algorithms extract rules from the training set without requiring discretization of the continuous valued features. Thus, the algorithms are applied to the most consistent feature subset, using the continuous values of those features.

IV. Dataset and Implementation

Our methodology is evaluated using a dataset of EEG recordings from 25 subjects (12 normal and 13 epileptic) acquired at the Neurology Department at University Hospital of Ioannina, Greece. More specifically, the EEG database consists of 858 prototype transient events: 274 prototype ESs, 254 prototype EMGs, 81 prototype EOGs, and 249 prototype SAAs, all characterized by two experienced neurologists (no disagreement among experts about characterization was realized). In our experiments, we use prototype transient events, and not "raw" transient events, and our goal is to evaluate the feature extraction and the classification steps of our methodology. The "Global Training" approach [2] was followed, and the training and test sets were constructed using the "stratified sampling" method [41]. Half from each category were used for training, while the rest were used for testing. The training set is used for the extraction of the discretization intervals, for finding the most consistent feature subset, and for generating the classification association rules.

The patients' ages ranged from 16 to 76 years. The duration of the EEGs was approximately 15 min. The EEG data was acquired using Ag/AgCl disk electrodes placed according to the International 10–20 Placement System [42]. Sixteen channels are recorded synchronously from five bipolar montages, where each electrode is referenced to an adjacent electrode. The EEGs were bandpass filtered between 1.6 and 70 Hz, including a 50-Hz notch filter, and sampled at 256 Hz with 12-bit resolution. The recordings were acquired while the patients were awake but resting, and include periods of open eyes, closed eyes, hyperventilation, and photic stimulation. Amplification was provided by a Medelec Profile EEG System.

The five classification algorithms previously mentioned have been tested using our dataset. Specifically, they were employed to extract rules from the training set and were further used to classify the cases in the test set. Several values for the parameters of the algorithms have been tested. The minimum support threshold ranged from 0.5 to 30%, and the minimum confidence threshold ranged from 50 to 100%.

V. Results

In order to evaluate the proposed methodology, four different measures were used: sensitivity (Se), specificity (Sp), selectivity (Sel), and accuracy (Acc). Table I depicts the results obtained. All the tested algorithms performed comparably, but the CBA algorithm was found to be the most effective. The transient event classification methodology was tested on our dataset and demonstrated sensitivity, specificity, and selectivity 86.13%, 91.75%, 83.1% for ES; 91.34%, 97.34%, 93.55% for EMG; 77.5%, 99.23%, 91.18% for EOG; and 87.90%, 93.75%, 85.16% for SAA, respectively.

In Table II, several methodologies for epileptic events detection are presented. Direct comparison among ours and other methodologies cannot be performed since 1) different datasets were employed and 2) the reported works address only spike detection, whereas our methodology detects and classifies the undesirable activity as well. To our knowledge, only a few

TABLE I
CLASSIFICATION RESULTS OBTAINED USING VARIOUS RULE-BASED CLASSIFICATION ALGORITHMS

Algorithm	Transient Event	Se[1] %	Sp[2] %	Sel[3] %	Acc[4] %
CBA	ES	86.13	91.75	83.1	**87.38**
	EMG	91.34	97.34	93.55	
	EOG	77.5	99.23	91.18	
	SAA	87.9	93.75	85.16	
CMAR	ES	80.29	92.1	82.71	**85.98**
	EMG	91.34	97.67	94.31	
	EOG	87.5	99	89.74	
	SAA	86.3	91.45	80.45	
CPAR	ES	85.4	91.41	82.39	**86.45**
	EMG	90.55	97.01	92.74	
	EOG	82.5	99.74	97.06	
	SAA	84.68	92.43	82.03	
RIPPER	ES	86.86	87.97	77.27	**85.05**
	EMG	89.79	98.01	95	
	EOG	87.5	99.48	94.59	
	SAA	77.42	93.09	82.05	
CN2	ES	92.7	79.73	68.28	**82.48**
	EMG	86.61	99.34	98.21	
	EOG	80	99.74	96.97	
	SAA	67.74	95.72	86.6	

[1]Se: Sensitivity
[2]Sp: Specificity
[3]Sel: Selectivity
[4]Acc: Accuracy

TABLE II
COMPARISON OF THE PERFORMANCE OF SEVERAL METHODS FOR SPIKE DETECTION

Reference	*# EEG*	*Se[1] (%)*	*Sp[2] (%)*	Sel[3] (%)
Hosteler et al. [7] 1992	5	59	-	89
Dingle et al. [26] 1993	11	53	-	100
Webber et al. [5] 1994	10	74	-	74
Park et al. [11] 1998	32	97	-	89
Tarassenko et al. [1] 1998	-	83-97	86-96	-
James et al. [12] 1999	43	55	-	82
Pang et al. [2] 2003	13	82-86	94-95	82-86
Acir et al. [13] 2005	29	89	-	86
Adjouadi et al. [15] 2005	31	82	-	92
This work	**25**	**86**	**92**	**83**

[1]Se: Sensitivity
[2]Sp: Specificity
[3]Sel: Selectivity

TABLE III
COMPARISON OF THE PERFORMANCE OF SEVERAL METHODS FOR CLASSIFICATION OF TRANSIENT EVENTS

Reference	*# EEG*	*Method*	Acc[1] (%)
Saastamoinen et al. [43] 1998	-	RBF Neural Networks	75
Castellaro et al. [24] 2002	50	Neural Networks, Expert System	80
Tzallas et al. [23] 2004	25	Neural Networks	76
This work	**25**	Association rule based classifier	87

[1]Acc: Accuracy

methods [23], [24], [43] have been previously applied to classify transient events in a wide range of categories. For these reasons, only qualitative comparisons and conclusions can be drawn.

Compared to other transient event classification methods [23], [24], [43] (Table III), our methodology is advantageous since it reports higher accuracy. It should also be mentioned that most of the methods presented in Tables II and III are based on neural networks [1], [2], [5], [7], [11]–[13], [23], [24], [43]. Such methods exhibit a serious drawback compared with our association rule approach, due to their inability to provide explanations for their classification decisions. In contrast, due to the rule-based nature of our methodology, the proposed approach satisfies this important requirement, and it is able to provide for each transient event the reason (rule) leading to each decision.

It should be mentioned that the best results of the CBA algorithm were obtained using minimum support of 1% and minimum confidence of 50%. CBA, in the rule generation process, generated 112 rules. 39 of them were rules which predicted ES, 32 predicted EMG, 8 predicted EOG, and the rest 33 predicted SAA. From the 112 rules extracted with the CBA algorithm, four rules are presented in Fig. 3. One rule from each category and, more specifically, the one with 100% confidence and highest support, is shown.

VI. INTERPRETATION AND USABILITY

Medical experts generally prefer rule-based classifiers to other classification models; e.g., neural networks. Neural networks and other classifiers that cannot provide interpretation are not comprehensible for experts, and thus not desirable [44]. On the other hand rule, extraction techniques are able to provide the desired comprehensibility [3], [4].

In addition, medical experts in their daily practice, when rapidly examining an EEG, use for their diagnosis some empirical rules and patterns. They do not use rules with specific thresholds like the ones reported in our work, but rather more "empirical" ones. The approach presented in this work discovers quantitative rules, which are considered better for use by inexperienced doctors [6].

It should also be mentioned that two experienced neurologists were asked to annotate several transient events for all four categories, whose features satisfied the antecedents (the "if" part) of the rules presented in Fig. 3. Both neurologists agreed with the decisions of the rules and concluded that these rules could be trusted. Other rules with lower support and confidence were presented to the neurologists, which were considered by them as "risky" ones.

Rule 1: **IF** Disc PSD 3 >= 17084.6
AND 195.2975 < Area <= 1835.13
THEN Prototype Transient Event is ES
(support=14.651%, confidence=100.000%)

Rule 2: **IF** Disc PSD 3 <= 123.3515
AND 0.054067 < Average slope <= 0.077462
THEN Prototype Transient Event is EMG
(support 13.023%, confidence=100.000%)

Rule 3: **IF** Disc PSD 9 <= 0.480739
AND 8.75 < Dominant Frequency <= 12.5
AND 39.37575 < Duration <= 60.98395]
THEN Prototype Transient Event is SAA
(support=11.860%, confidence=100.000%)

Rule 4: **IF** Disc PSD 4 <= 304.9815
AND Duration > 101.912
AND 0.054067 < Average Slope <= 0.077462
THEN Prototype Transient Event is EOG
(support=2.326%, confidence=100.000%)

Fig. 3. Four indicative rules extracted using the CBA algorithm.

Finally, according to the neurologists, another highly desirable feature of our methodology, which also motivates the use of association rule mining, is the large number of rules which were extracted. In this way, there were many transient events in the test set covered by multiple rules. Some rules predicted different decisions for the same transient event, but there were many transient events classified by many rules with the same consequent. This is of great importance for both medical experts and patients, since a method that performs automated diagnosis is more reliable if the decision could be confirmed in more than one way (multiple rules) [45]. On the other hand, decision trees make their decision for a test instance based solely on one "rule" [37] (rules can be extracted from decision trees, simply by parsing the tree from the root to a leaf). The preceding is considered as an advantage of classification using association rules against decision trees.

VII. Conclusion

We introduced a four-stage methodology based on data mining techniques for the detection and classification of transient events in EEG recordings. All steps in the proposed methodology are automated. In the first stage, the EEG signal was preprocessed and transient event detection took place. In the second stage, the transient events were clustered in order to obtain the prototypes, and feature extraction was realized. In the third stage, the extracted features were discretized and a subset of them was selected. Finally, in the last stage, classification algorithms based on association rules were applied in order to extract rules for classifying the new transient events. The best overall accuracy obtained was 87.38%.

A limitation of our methodology emerges from the employment of association rules for the classification. The utilization of association rules, besides finding valid, causal relationships in the clinical data, will also find all the spurious and particular relationships among the data in the specific dataset. For this reason, results of any association rule mining procedure should be considered as exploratory and hypothesis-generating.

Further improvement might focus on the utilization of more features extracted from the EEG describing better each transient event, and also the employment of patient's demographic and history data. In addition, spatial information from adjacent channels or temporal information could be used. Neurologists make considerable use of spatial and temporal information when visually perform spike detection. The performance of the transient event classification methodology was high; however, other evaluation methods (e.g., hold-patients-out testing) and clinical assessment might be considered in order to fully reveal its potential.

References

[1] L. Tarassenko, Y. U. Khan, and M. R. G. Holt, "Identification of interictal spikes in the EEG using neural network analysis," *Proc. Inst. Elect. Eng.—Sci. Meas. Technol.*, vol. 145, pp. 270–278, 1998.
[2] C. C. C. Pang, A. R. M. Upton, G. Shine, and M. V. Kamath, "A comparison of algorithms for detection of spikes in the electroencephalogram," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 4, pp. 521–525, Apr. 2003.
[3] B. Liu, W. Hsu, and Y. Ma, "Integrating classification and association rule mining," in *Proc. 4th Int. Conf. Knowledge Discovery and Data Mining (KDD-98)*, New York, 1998, pp. 80–86.
[4] W. Li, J. Han, and J. Pei, "CMAR: Accurate and efficient classification based on multiple class-association rules," in *Proc. 2001 IEEE Int. Conf. Data Mining*, San Jose, CA, 2001, pp. 369–376.
[5] W. R. Webber, B. Litt, K. Wilson, and R. P. Lesser, "Practical detection of epileptiform discharges (EDs) in the EEG using an artificial neural network: A comparison of raw and parameterized EEG data," *Electroen. Clin. Neurophysiol.*, vol. 91, pp. 194–204, 1994.
[6] E. Niedermeyer and F. L. da Silva, *Electroencephalography, Basic Principles, Clinical Applications and Related Fields*. Philadelphia, PA: Williams & Wilkins, 1993.
[7] W. Hosteler, H. Doller, and W. Homan, "Assessment of a computer program to detect epileptiform spikes," *Electroencephalogr. Clin. Neurophysiol.*, vol. 83, pp. 1–11, 1992.
[8] F. H. L. da Silva, "Detection of nonstationarities in EEGs using the autoregressive model-an application to EEGs of epileptics," in *CEAN: Computerized EEG Analysis*, G. Dolce and H. Kunkel, Eds. Stuttgart, Germany: Fischer, 1975, pp. 180–199.
[9] T. Kalayci and O. Ozdamar, "Wavelet pre-processing for automated neural network detection of EEG spikes," *IEEE Eng. Med. Biol. Mag.*, vol. 14, no. 2, pp. 160–166, Mar./Apr. 1995.
[10] P. Gloor, "Contributions of electroencephalography and electrocorticography in the neurosurgical treatment of the epilepsies," *Adv. Neurol.*, vol. 8, pp. 59–105, 1975.
[11] H. S. Park, Y. H. Lee, N. G. Kim, D. S. Lee, and S. I. Kim, "Detection of epileptiform activities in the EEG using neural network and expert system," *Medinfo*, vol. 9, pp. 1255–1259, 1998.
[12] C. J. James, R. D. Jones, P. J. Bones, and G. J. Carroll, "Detection of epileptiform discharges in the EEG by a hybrid system compromising mimetic, self-organized artificial neural network, and fuzzy logic stages," *Clin. Neurophysiol.*, vol. 110, pp. 2049–2063, 1999.
[13] N. Acir, I. Oztura, M. Kuntalp, B. Balkan, and C. Guzelis, "Automatic detection of epilepticform events in EEG by a three stage procedure based on artificial neural networks," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 1, pp. 30–40, Jan. 2005.
[14] P. Valenti, E. Cazamajou, M. Scarpettini, A. Aizemberg, W. Silva, and S. Kochen, "Automatic detection of spikes using data mining models," *J. Neurosci. Meth.*, to be published.
[15] M. Adjouadi, M. Cabrerizo, M. Ayala, D. Sanchez, I. Yaylali, P. Jayakar, and A. Barreto, "Detection of interictal spikes and artefactual data through

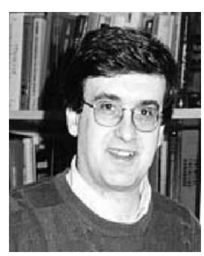
orthogonal transformations," *J. Clinical Neuroph.*, vol. 22, pp. 53–64, 2005.
[16] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in *Proc. 20th Int. Conf. Very Large Data Bases*, Santiago, Chile, 1994, pp. 487–499.
[17] J. Bourien, J. J. Bellanger, F. Bartolomei, P. Chauvel, and F. Wendling, "Mining reproducible activation patterns in epileptic intracerebral EEG signals: Application to interictal activity," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 2, pp. 304–315, Feb. 2004.
[18] S. Konias and N. Maglaveras, "A rule discovery algorithm appropriate for ECG signals," in *Proc. 31st Comput. Cardiology*, Chicago, IL, 2004, pp. 57–60.
[19] A. Flexer, "Data mining and electroencephalography," *Stat. Methods Med. Res.*, vol. 9, pp. 395–413, 2000.
[20] D. Sommer and M. Golz, "Clustering EEG-segments using hierarchical agglomerative methods and self-organising maps," in *Proc. Annu. Conf. Eur. Neural Network Society*, Vienna, Austria, 2001, pp. 642–649.
[21] P. Wahlberg and G. Salomonsson, "Feature extraction and clustering of EEG epileptic spikes," *Comput. Biomed. Res.*, vol. 29, pp. 382–394, 1996.
[22] P. Wahlberg and G. Lantz, "Methods for robust clustering of epileptic EEG spikes," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 7, pp. 857–876, Jul. 2000.
[23] A. T. Tzallas, C. D. Katsis, P. S. Karvelis, D. I. Fotiadis, S. Konitsiotis, and S. Giannopoulos, "Classification of transient events in EEG recordings," in *Proc. 2nd Int. Conf. Adv. Biomed. Signal Inf. Process.*, Malta, 2004, pp. 14–20.
[24] C. Castellaro, G. Favaro, A. Castellaro, A. Casagrande, S. Castellaro, D. V. Puthenparampil, and C. F. Salimbeni, "An artificial intelligence approach to classify and analyse EEG traces," *Clin. Neurophysiol.*, vol. 32, pp. 193–214, 2002.
[25] R. Kothari and D. Pitts, "On finding the number of clusters," *Pattern Recognit. Lett.*, vol. 20, pp. 405–416, 1999.
[26] A. A. Dingle, R. D. Jones, G. J. Carroll, and W. R. Fright, "A multistage system to detect epileptiform activity in the EEG," *IEEE Trans. Biomed. Eng.*, vol. 40, no. 12, pp. 1260–1268, Dec. 1993.
[27] U. M. Fayyad and K. B. Irani, "Multi-interval discretization of continuous-valued attributes for classification learning," in *Proc. 13th Int. Joint Conf. Artif. Intell.*, 1993, pp. 1022–1027.
[28] R. Kohavi and M. Sahami, "Error-based and entropy-based discretization of continuous features," in *Proc. 2nd Int. Conf. Knowledge Data Mining*, Menlo Park, CA, 1996, pp. 114–119.
[29] H. Liu and H. Motoda, *Feature Extraction, Construction and Selection: A Data Mining Perspective*. Boston, MA: Kluwer, 1998.
[30] M. A. Hall and G. Holmes, "Benchmarking attribute selection techniques for discrete class data mining," *IEEE Trans. Knowl. Data Eng.*, vol. 15, no. 6, pp. 1437–1447, Nov./Dec. 2003.
[31] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artif. Intell.*, vol. 97, pp. 273–324, 1997.
[32] I. Guyon and A. Elissee, "An introduction to variable and feature selection," *J. Mach. Learning Res.*, vol. 3, pp. 1157–1182, 2003.
[33] H. Liu and R. Setiono, "A probabilistic approach to feature selection—a filter solution," in *Proc. 13th Int. Conf. Machine Learning*, Bari, Italy, 1996, pp. 319–327.
[34] J. Kittler, "Feature set search algorithms," in *Pattern Regognition and Signal Processing*, C. H. Chen, Ed. Alphen aan den Rijn, The Netherlands: Sijthoff and Noordhoff, 1978
[35] W. T. Blume and J. F. Lemieux, "Morphology of spikes in spike-and-wave complexes," *Electroencephalogr. Clin. Neurophysiol.*, vol. 69, pp. 508–515, 1988.
[36] M. Elton, M. Scherg, and D. Von Cramon, "Effects of high-pass filter frequency and slope on BAEP amplitude, latency and wave form," *Electroencephalogr. Clin. Neurophysiol*, vol. 57, pp. 490–494, 1984.
[37] J. R. Quinlan, *C4.5: Programs for Machine Learning*. Sam Mateo, CA: Morgan Kaufmann, 1993.
[38] X. Yin and J. Han, "CPAR: Classification based on predictive association rules," in *Proc. 3rd SIAM Int. Conf. Data Mining*, San Francisco, CA, 2003, pp. 331–335.
[39] W. W. Cohen, "Fast effective rule induction," in *Proc. ICML-95*, Tarragona, Spain, 1995, pp. 115–123.
[40] P. Clark and R. Boswell, "Rule induction with CN2: Some recent improvements," in *Proc. Machine Learning-EWSL-91*, Berlin, Germany, 1991, pp. 151–163.
[41] M. Kantardzic, *Data Mining: Concepts, Models, Methods and Algorithms*. Hoboken, NJ: Wiley-IEEE Press, 2002.
[42] H. H. Jasper, "The ten twenty electrode system of the international federation," *Clin. Neurophysiol.*, vol. 10, pp. 371–375, 1958.
[43] A. Saastamoinen, T. Pietilä, A. Värri, M. Lehtokangas, and J. Saarinen, "Waveform detection with RBF network—Application to automated EEG analysis," *Neurocomputing*, vol. 20, pp. 1–13, 1998.
[44] V. Schetinin and J. Schult, "The combined technique for detection of artefacts in clinical electroencephalograms of sleeping newborns," in *IEEE Trans. Inform. Technol. Biomed.*, vol. 8, no. 1, pp. 28–35, Mar. 2004.
[45] A. Kusiak, J. A. Kern, K. H. Kernstine, and B. T. L. Tseng, "Autonomous decision-making: A data mining approach," *IEEE Trans. Inform. Technol. Biomed.*, vol. 4, no. 4, pp. 274–284, Dec. 2000.

Themis P. Exarchos was born in Ioannina, Greece, in 1980. He received the Diploma degree in computer engineering and informatics from the University of Patras, Patras, Greece, in 2003. He is currently working toward the Ph.D. degree in medical physics at the University of Ioannina, Ioannina, Greece.

His research interests include medical data mining, decision support systems in healthcare, and biomedical applications.

Alexandros T. Tzallas was born in Athens, Greece, in 1977. He received the B.Sc. degree in physics from the University of Ioannina, Ioannina, Greece, in 2001, where he is currently working toward the Ph.D. degree in medical physics.

His research interests include biomedical signal processing, analysis of electroencephalograms (EEG), epilepsy diagnosis, and biomedical applications.



Dimitrios I. Fotiadis (M'01) was born in Ioannina, Greece, in 1961. He received the Diploma degree in chemical engineering from The National Technical University of Athens, Athens, Greece, and the Ph.D. degree in chemical engineering from the University of Minnesota, Twin Cities.

Since 1995, he has been with the Department of Computer Science, University of Ioannina, Ioannina, Greece, where he is currently an Associate Professor. He is the Director of the Unit of Medical Technology and Intelligent Information Systems. His research interests include biomedical technology, biomechanics, scientific computing, and intelligent information systems.

Spiros Konitsiotis was born in Corfu, Greece, in 1963. He received the M.D. degree from the University of Ioannina Medical School, Ioannina, Greece, in 1987, and the Ph.D. degree from the National Institutes of Health, Bethesda, MD, in 1990, "on the neuropharmacology of basal ganglia."

Since 2000, he has been with the University of Ioannina Medical School, where he is currently an Assistant Professor of Neurology. His research interests include movement disorders, Parkinson's disease, neuropharmacology, and drug-induced movement disorders.

Sotirios Giannopoulos was born in Ioannina, Greece, in 1963. He received the M.D. degree from Athens University Medical School, Athens, Greece, in 1988, and the Specialty Certificate in Neurology in 1994.

Since 2002, he has been with the University of Ioannina Medical School, Ioannina, Greece, where he is a Lecturer in Neurology. His research interests include stroke, neurosonology, and use of neurosonology in clinical practice.