Electroencephalogram Signals Analysis by Fuzzy Classifiers based on Cumulative Mutual Information

Jan Rabcan [0000-0003-2835-9114], Elena Zaitseva[0000-0002-9087-0311],Vitaly Levashenko[0000-0003-1932-3603]

University of Zilina, Department of Infromatics, Univerzitna 8215/1,
010 26, Zilina, Slovakia
(jan.rabcan, elena.zaitseva, vitaly.levashenko)@fri.uniza.sk

Abstract. New algorithm for EEG signals classification is proposed in this paper. The typical algorithm for signal classification includes two steps: preliminary transformation and classification. The preliminary transformation modifies investigated signals by the procedures of feature extraction and dimension reduction into data set accepted for the classification. This transformation causes the loss of some information that can improve the classification accuracy. New procedure of fuzzification is added in the preliminary transformation for the proposed algorithm. This procedure allows using the fuzzy classifier at the second step. The Fuzzy Random Forest is used for the classification of EEG signals in the proposed algorithm. The new procedure (fuzzification) in the preliminary transformation and new fuzzy classifier (Fuzzy Random Forest) allows increase the classification accuracy of the EEG signals. The efficiency of new algorithm is evaluated by two investigations. The first one focuses on epilepsy diagnostics and the second one aims at epileptics' seizure detection. The comparison with other studies and shows the increasing of the classification accuracy of EEG signals by the proposed algorithm.

Keywords: Electroencephalogram (EEG), Fuzzification, Fuzzy Decision Tree, Fuzzy Random Forest.

1 Introduction

Epilepsy is considered as one of the most common chronic neurological disorders. Statistics of World Health Organization shows that approximately 50 million people worldwide suffer from epilepsy [1]. The most notable epileptic symptom are recurrent unprovoked seizures which usually occurs without warning, and they can affect each part of the body [2]. Seizures are caused by the electrical activity of the brain, concretely by the unexpected electrical disturbance of brain and excessive neuronal discharge. These brain activities can be monitored by the *Electroencephalogram* (EEG) [3]. Ability to measure brain electrical activity makes EEG as one of the most important tools in neurology diagnostics [5, 6]. EEG measures electrical brain activity by the electrodes placed on the scalp [1, 2]. Captured measurement are signals. These signals are converted into the graphic form represented by curves [4]. The shape and character of

captured curve is strongly affected by current brain activity. Curves of signals obtained from EEG can be evaluated visually by specialized doctors who are able to use them in neural disorder diagnostic. However, such evaluation might not provide enough information for doctors to provide reliable neurological diagnostics. Diagnostic based on EEG can be significantly improved by development of algorithms for automatic analysis of EEG signals. Such algorithms must involve automatic extraction and analysis of information from the captured signal. According to extracted information, the brain state can be predicted or classified. In this paper, the algorithm for EEG signal analysis is proposed. This algorithm is demonstrated by the detection of epileptic seizures and the diagnostics of epilepsy.

To develop reliable and credible algorithms for automatic EEG signal analysis can be developed based on the real data. There is well-known dataset of EEG signals collected and published by R. G. Andrzejak [6]. This dataset consists of 500 samples of EEG signal evenly divided into five disjoint subsets A, B, C, D, and E. Therefore, each subset contains 100 samples of EEG signals. Each of these samples represents an EEG signal captured during 23.6 seconds of EEG recording. Subset A and B contain measurement recorded from persons who do not suffer from epilepsy. In case of subset A, patients had eyes opened during recording while patients in subset B had eyes closed. Whether the eyes are open or closed affects the electrical activity of the brain according to [6]. Samples in subsets C, D, E was recorded from persons suffered from epilepsy. While patients in subsets C and D were investigated during seizure-free intervals, subset E contains samples measured during seizure activity only. Subset C consists of the vectors, were obtained from the hippocampal formation of the opposite hemisphere of the brain. Data vectors in the subset D were recorded from within the epileptogenic zone. Examples of samples from these subsets are shown in Fig. 1.
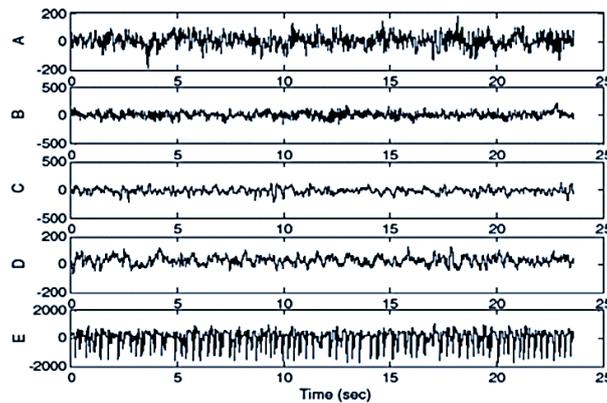

Fig. 1. Raw signals examples, randomly chosen from dataset [6]

From Fig. 1 is observable that signals in A and E are different from signals in B, C, and D. Nevertheless, only visual inspection can be insufficient to recognize healthy (A, B) and epileptics EEG segments (C, D, E). The most significant similarity between this group is oblivious in case of samples from B, C, and D. The samples are very similar

to each other and the use only of visual inspection can lead to failure in diagnostic (patient suffering from epilepsy diagnosed as healthy). This example was used to shows that development of algorithms for automatic EEG signals analysis can be instrumental in diagnostics. It allows deciding whether a person has epilepsy without information if they have had an epileptic seizure or have not. Another task can be focused on detection of seizure activities. Also, development and analysis of these algorithms can enable the growth of decision support systems for early diagnosis of epilepsy [3–5].

Algorithms for automatized EEG signals analysis has been considered in several works. Different classification methods are used in these investigations, for example, K-nearest neighbor classifier [7], neural networks [5, 8, 9], support vector classifier [10, 11], decision trees [12]. The classification accuracy of EEG signals can be increased by the application of evolution methods [7] and clustering analysis [13]. Analysis of these investigations shows that one of effective classification methods of EEG signal for identification of persons with epilepsy is based on the application of decision tree [12] and *Fuzzy Decision Tree* (FDT) [14, 15]. The approach of EEG signal classification by FDT is developed in this paper and *ordered FDT* (oFDT) introduced in [16] are used. This tree has a regular structure that contains exactly one attribute at each level of the tree [17]. This ability allows parallelization of classification. The evaluation of the EEG signals classification in [17] has possible and quite accurate. However, there can be possibility for increasing of the classification accuracy that is considered in this paper. In particular, the application of *Fuzzy Random Forest* (FRF) is considered.

Typical algorithms for EEG signals classification in addition to the classification procedure include the signal preliminary transformation into numerical attributes [18, 19]. For this reason, analyzed signal must be preprocessed to extract useful information necessary for creating classification models. Nowadays, we can find many preprocessing schemes in the literature. Basically, most of them consist of two steps: feature extraction and dimensional reduction. Feature extraction usually includes wavelet or spectral transformation. In [12] discrete wavelet and *Discrete Fourier Transformation* (DFT) was compared. Result of this paper shows better performance of DFT in combination with method C4.5 for decision tree induction. In [20] and [9] the Welch's method was used. These studies indicate that power spectrum density estimation by Welch's method provides strong attributes which represent EEG signals well. The second typical step of preprocessing is dimensional reduction. This step is required when feature extraction produce data matrix of large dimension. Dimensional reduction usually transforms data matrix to matrix of significantly smaller dimension. In [10] three techniques was compared: *Principal Components Analysis* (PCA), linear discriminant analysis, and independent component analysis. The used in this paper PCA is recomended especially in cases when the number of training samples per class is small [21].

In comparison with preprocessing of EEG signals mentioned above, the algorithm presented by us in [14] and [15] involve fuzzy logic to preliminary data transformation as additional step. This step is fuzzification which can significantly reduce possible uncertainty in data. This paper shows a detailed investigation of this step with combination of different fuzzy classifiers: FDTs and FRF. This analysis allows developing of accurate method for EEG signals analysis.

2   Structure of new algorithm for EEG signal classification

The proposed algorithm for epileptics' seizure detection from EEG signals includes two steps (Fig.2): the preliminary transformation of EEG signal and the classification. The differences between traditional algorithms of signal classification and algorithm proposed in this paper are (a) the additional procedure of fuzzification at step of preliminary transformation and (b) the fuzzy classifier at step of classification. The addition of the fuzzification permits to use fuzzy classification and consider the ambiguity of the transformed and reduced initial EEG signal.

In papers [22–24], it has been shown that fuzzy data are useful for representation of ambiguous and imprecise data with higher accuracy and the application of fuzzy data instead of numerical data allows increasing of the accuracy of classification. The principal steps of the proposed algorithm for EEG signal classification are shown in the figure bellow.
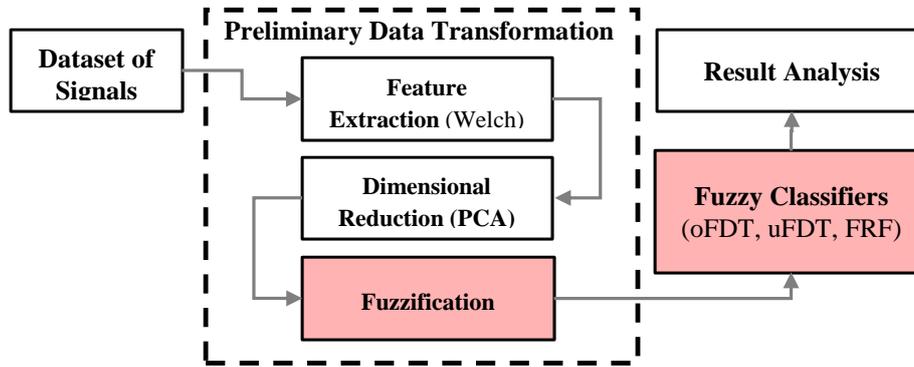

Fig. 2. The main steps of the proposed algorithm

Signal preprocessing and the preliminary signal transformation of EEG signals causes the loss of some information of initial signal. This fact implies the classification of ambiguous data. Fuzzy classifiers can be used for effective classification in this case [25]. The application of these classifiers is possible if the input data for the classification is fuzzy. However, data in standard algorithms of EEG signal classification after the preliminary transformation is not fuzzy even if procedures of attribute extraction and dimension reduction include fuzzy transformation. For example, in paper [26] propose the algorithm for EEG signal classification with the application of adaptive neuro-fuzzy inference system. Other fuzzy techniques are described in [27]. At same time classifications in these algorithms are not implemented based on fuzzy classifier. Classification can be improved by application of a fuzzy classifier that causes the representation of data for classification after the preliminary transformation as the fuzzy data.

Moreover, we lose some amount of information during EEG signal preprocessing during preliminary data transformation. Therefore, the ambiguity and uncertainty of initial data can occur. By applying fuzzification, we can reduce possible ambiguity.

Moreover, fuzzy logic is tolerant to imprecise data and deviations during data measurement. Fuzzy logic can also model nonlinear functions of arbitrary complexity [28]. These facts allow reducing ambiguity and uncertainty of transformed data. According to obtained data after preliminary data transformation, we obtain data which was used to create three fuzzy classifiers. In the next sections, the steps of preliminary data transformation are described in detail.

3   Preliminary Data Transformation

The preliminary data transformation developed in [14] and [15] was evaluated by classification of EEG signals from dataset gathered in [6] (Fig. 1). Description of recording in [6] tells that during the recording of each EEG signal was used the same 128-channels amplifier system. Recorded data were converted to digital form by 12 bit analog-to-digital conversion, and then the data were written continuously onto disk of data acquisition computer system at a sampling rate of 173.61 Hz. The frequency of Band-pass filter was 0.53-40 Hz. Aa was mention in previous section, signal obtained by this recording is function of time. Therefore, the preliminary data transformation is necessary to extract useful information, which can be used to build classifiers and classify new EEG records.

3.1   Feature Extraction

The signal preprocessing which preliminary data transformation do, consist of three steps (Fig. 2). At first, feature extraction by *Welch method* is applied [29]. The method is based on the concept of periodogram spectrum estimates, which is the result of signal conversion from its time to frequency domain. Periodogram is commonly used for examining the amplitude vs frequency characteristics of analyzed signal. Time series of a signal is split into overlapping segments (Fig. 3) of the predefined length which is equal for each preiodogram. Then the periodgram of each segment is computed and used to find important periodic components in the time series. At last step, periodograms are averaged. The result of averaged periodograms is known as power spectral density estimation of analyzed signal [29].

Fig. 3. Illustration of segmentation by Welch's method. For each segment, periodogram is estimated and obtained periodograms are averaged.

After application of *Welch method* to EEG signals, we obtain feature matrix of 128 features. Each row of this matrix represents one EEG signal in frequency domain. Obtained number of features is too big to build reliable classifier (we obtain too many features in ratio with number of data instance). To decrease number of features, we applied dimensional reduction technique. We use PCA technique, which is commonly used in EEG signal analysis [20].

3.2 Dimensionality Reduction

In this paper, we use PCA to reduce dimension of feature matrix obtained by spectral transformation. PCA transform a feature matrix described by n features $Y = (Y_1, \ldots, Y_n)$ into linearly uncorrelated latent variables called "principal components" [30]. Transformation is carried out in such a way that the first component has the biggest variance as possible. Each succeeding component has the biggest variance too as possible but the constraint that it is orthogonal to the preceding components must hold. Bigger variance of component indicates bigger variability in the data. Therefore, variance of component characterizes its importance. Hence, the first component is considered as the most significant component. PCA provides this transformation using eigenvectors $V = (v_1 \quad v_2 \quad \cdots \quad v_n)$ of covariance matrix $M = \frac{YY^T}{n-1}$. Eigenvector of matrix M is defined as vector v_i that $Mv_i = \lambda_i v_i$, where λ_i is a corresponding eigenvalue to the vector v_i given by $|M - \lambda I| = 0$, where I is the unit matrix. The principal components are computed as $X = V^T \times Y^T, Y = (Y_1, \ldots, Y_n)$. This transformation forms orthogonal principal components.

After transformation, it is necessary to choose only several important principal components. Nowadays, many criterions exist to solve this problem. One of the most commonly used criteria to determine appropriate number of principal components is the eigenvalue-one criterion, also known as the Kaiser criterion [31]. This criterion selects a principal component as significant if its variance is higher than 1.00.

As a result of PCA, we obtained 8 principal components, which describe the reduced feature matrix of origin EEG signals. Therefore, each signal is described by 8 principal components. In text bellow, these principal components are noted as numerical input attributes X_i $(i = 1, \ldots, 8)$ of EEG signals.

3.3 Data Fuzzification

We propose to use fuzzification process as part of preliminary data transformation. This process transforms each numeric attribute Xi obtained after dimensionality reduction into fuzzy attribute A_i $(i = 1, \ldots, n)$. After this transformation, each fuzzy attribute A_i will be composed of mi (mi $\geq$ 2) linguistic terms. A j-th linguistic term of attribute A_i is defined by fuzzy set $A_{i,j}$ $(j = 1, \ldots, m_i)$. Fuzzy set $A_{i,j}$ with respect to a universe Xi is defined by a membership function $\mu_{A_{i,j}}(x): X_i \rightarrow \langle 0,1 \rangle$. This function defines a membership degree $\mu_{A_{i,j}}(x)$ for each element x $(x \in X_i)$, which define how strongly element x is the member of fuzzy set $A_{i,j}$. Formally, fuzzy set $A_{i,j}$ is defined as an ordered

set of pairs $A_{i,j} = \left\{\left(x, \mu_{A_{i,j}}(x)\right), x \in X_i\right\}$, where: (a) $\mu_{A_{i,j}}(x) = 0$, if and only if x is not the member of set $A_{i,j}$; (b) $0 < \mu_{A_{i,j}}(x) < 1$, if and only if x is not the full member of set $A_{i,j}$; and (c) $\mu_{A_{i,j}}(x) = 1$, if and only if x is the full member of set $A_{i,j}$.

To perform fuzzification of analyzed data, we used the algorithm proposed in [32]. This algorithm is based on the computation of the fuzzy entropy of fuzzy sets. The algorithm divides values $x \in X_i$ into m_i intervals. Intervals are defined by points $C_1 \ldots C_{m_i}$. We find this borders by K-Means algorithm as in [32]. Number m_i of intervals is determined automatically. This number agrees with the number of linguistic terms mi of fuzzy attribute A_i. The algorithm is initially performed with two linguistic terms, then the algorithm adds linguistic term to attribute. Adding of linguistics term is repeated until the fuzzy entropy of attribute does not raise. The definition of Fuzzy entropy, which with the aforementioned algorithm works, is following.

$$FE\left(A_{i,j}\right) = -\sum_{k=1}^{m_b} D_{A_{i,j}}^b \times log_2\, D_{A_{i,j}}^b$$

where m_b is the number of classes defined by output attribute B. Notation $x_{\in b}$ define, that x belongs to class b. Than, $D_{A_{i,j}}^b$ is defined as follow:

$$D_{A_{i,j}}^b = \frac{\sum_{x_{\in b}} \mu_{A_{i,j}}(x)}{\sum_x \mu_{A_{i,j}}(x)}$$

For each attribute A_i must hold following constrain where $\sum_{j=1}^{m_i} A_{i,j} = 1$. Then the fuzzy entropy of attribute A defined as:

$$FE(A_i) = \sum_{q=1}^{n} FE\left(A_{i,j}\right)$$

Then the fuzzification process assigns membership degree $\mu_{A_{i,1}}(x)$ to each x of numeric attribute X_i. This membership degree is obtained by triangular membership function. The definition of used membership function for first linguistic term $A_{i,1}$ of attribute A_i is following.

$$\mu_{A_{i,1}}(x) = \begin{cases} 1 & x \leq C_1 \\ \dfrac{C_2 - x}{C_2 - C_1} & C_1 < x < C_2 \\ 0 & x \geq C_2 \end{cases}$$

Each non-first and non-last linguistics term of A_i has a membership function $\mu_{A_{i,q}}$ ($q = 2,3,\ldots,m_i - 1$) defined as follow:

$$\mu_{A_{i,q}}(x) = \begin{cases} 0 & x \leq C_{j-1} \\ \dfrac{x - C_{q-1}}{C_q - C_{q-1}} & C_{q-1} < x \leq C_q \\ \dfrac{C_{j+1} - x}{C_{j+1} - C_j} & C_q < x \leq C_{q+1} \\ 0 & x \geq C_{q+1} \end{cases}$$

Finally, the last term A_{i,m_i} of A_i has the membership function with following mathematical form:

$$\mu_{A_{i,m_i}}(x) = \begin{cases} 0 & x \leq C_{m_i-1} \\ \dfrac{x - C_{k-1}}{C_k - C_{k-1}} & C_{m_i-1} < x \leq C_{m_i} \\ 1 & x \geq C_{m_i} \end{cases}$$

Fuzzification is the last step of preliminary data transformation. After fuzzification of reduced data matrix, we obtained data represented by 8 fuzzy attributes. This data was used to induct fuzzy classifiers and evaluate the performance of proposed algorithm for EEG signals analysis.

4   Classification of Fuzzy Data

After fuzzification, we obtained data described by 8 fuzzy input attributes A_i ($i = 1, \ldots, 8$) and one output attribute B. According to this data, we induct three fuzzy classifiers to evaluate new algorithm for EEG signal analysis. We use oFDT which combines into the FRF.

4.1   Fuzzy Decision Trees

Nodes and leaves are the main parts of the Decision trees. Each node is associated with one input attribute (splitting attribute). Domain of this input attribute determines outcoming edges of the node. The domain of attribute defines all its possible values. The classification's task solving begins at the root of the tree. If a node is associated with an input attribute, then the outcome edge for the classified instance is determined according to splitting attribute, and the classification continues using the appropriate sub-tree. When classified instance encounters a leaf, encountered leaf give the predicted class as classification result. In case of fuzzy decision trees, classified instance usually traverses by multiple branches during classification. Therefore, the decision is based on several leaves.

In this paper, we use two type of fuzzy decision trees based on CMI. The first is oFDT. This type of tree has nodes on the same level associated with one input attribute. This fact allows to performing classification as parallel process. The second type is uFDT. uFDT do not require the same splitting attribute on each level of the tree. In Table 1 are shown splitting criteria for these FDTs.

Table 1. Splitting criteria based on CMI.

uFDT	oFDT
$i_q = argmax \left(\dfrac{\mathbf{I}\left(B; U_{q-1}, A_{i_q}\right)}{\mathbf{H}\left(A_{i_q} \mid U_{q-1}\right)} \right)$	$i_q = argmax \left(\dfrac{\mathbf{I}\left(B; \boldsymbol{U_{q-1}}, A_{i_q}\right)}{\mathbf{H}\left(A_{i_q} \mid \boldsymbol{U_{q-1}},\right)} \right)$

In the Table 1 *argmax* returns attribute index i_q with maximal value of CMI, $U_{q-1} = \{A_{i_1,j_1} \times \ldots \times A_{i_{q-1},j_{q-1}}\}$ is the fuzzy set defined by the sequence of fuzzy terms $A_{i_1,j_1}, \ldots, A_{i_{q-1},j_{q-1}}$ of selected attributes $A_{i_1}, \ldots, A_{i_{q-1}}$ from the root to the q-th node. $\boldsymbol{U_{q-1}}$ is the sequence of selected attributes $\{A_{i_1}, \ldots, A_{i_{q-1}}\}$. The CMI in output attribute B about the attribute A_{i_q} and the sequence of values U_{q-1} been introduced in [33] and calculated as:

$$\mathbf{I}\left(B; U_{q-1}, A_{i_q}\right) = \sum_{j_q=1}^{m_{i_q}} \sum_{j=1}^{m_b} M\left(B_j \times U_{q-1} \times A_{i_q,j_q}\right) \times$$

$$\times \left(\log_2 M\left(B_j \times U_{q-1} \times A_{i_q,j_q}\right) + \log_2 M\left(U_{q-1}\right) - \log_2 M\left(B_j \times U_{q-1}\right) - \log_2 M\left(U_{q-1} \times A_{i_q,j_q}\right) \right),$$

where $M\left(B_j \times U_{q-1} \times A_{i_q,j_q}\right)$ is measurement of cardinality of fuzzy set $B_j \times U_{q-1} \times A_{i_q,j_q}$; The conditional cumulative entropy between fuzzy attribute A_{i_q} and the sequence of selected attribute terms U_{q-1} is defined as:

$$\mathbf{H}\left(A_{i_q} \mid U_{q-1}\right) = \sum_{j=1}^{m_{iq}} M\left(A_{i_q,j}, U_{q-1}\right) \times \left(\log_2 \left(M\left(U_{q-1}\right)\right) - \log_2 M\left(A_{i_q,j} \times U_{q-1}\right) \right).$$

Implemented FDTs solve overfitting problem by pruning technique. Overfitting occur leaves covers insuficient number of instances in ratio with input data. To avoid overfitting, we use threshold parameters alpha and beta to determine leaf nodes during FDT induction. These parameters are described in detail in [15].

4.2 Fuzzy Random Forest

Random forest is classifier, which combines set of individual decision trees into one classifier system. Advantage of random forests against individual decision trees is stability. Decision trees are sensitive to variations in data. Generally, random forests are not sensitive to overfitting while decision trees tend to overfit input data [34]. Overfitting occurs when leaves contain a small number of instances in ratio with number of input data. Nevertheless, pruning techniques can solve this problem. On the other hand, decision trees are well interpretable classifier. In case of random forest, it is hard to figure out what hundreds of trees tell. Next advantage of decision trees is classification

performance. Provided classification by many trees is significantly slower then classify by one individual decision tree. Nevertheless, many papers shows bigger classification accuracy of random forests over individual decision trees, eq. in [35].

Random forests are type of ensemble methods. Ensemble methods combines multiple classifiers to obtain better classification accuracy than can acquire any of the constituent classifiers alone [36]. Many methods of ensemble learning based on different base of classifiers was developed. One of the first is bagging [37]. Bagging creates new dataset from origin dataset for each classifier by sampling with replacement. It means, one instance from origin dataset can occur in sampled dataset multiple times. This approach helps to avoid overfitting. Bagging is often combined with decision trees. In [38], bagging is combined extended by random attribute selection. Splitting attribute is chosen for each non-leaf node from randomly selected subset of unused attributes in analyzed branch according to used splitting criteria. In paper [39], we propose to use splitting criterion based on CMI [17, 33] for FRF trees induction.

FRF consists of defined number of individual FDTs. Each of these FDT provide some decision (classification result). To obtain decision of forest, results of individual trees must be combined (Fig. 3). The combining of these results was achieved by the summation of membership degrees of belonging to each class divided by the number of trees. The strategy of classification, which we use, is shown in figure below.
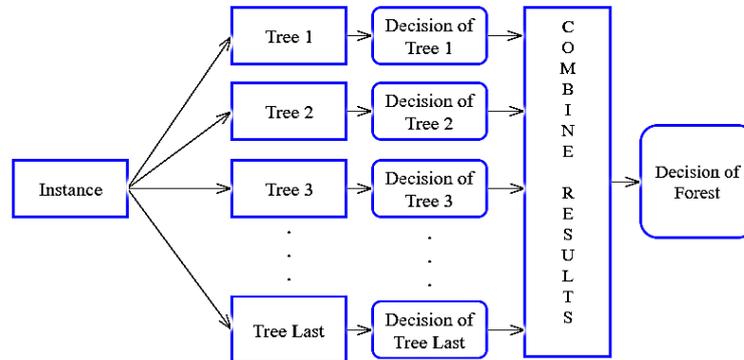

Fig. 2. Strategy of classification by FRF

In case of implemented random forest, the number of decision trees must be specified as input parameter. It is important to find ideal value of this parameter because small number of decision trees can lead to small classification power. In case of big number of decision trees, the algorithm can suffer from slow computation performance caused by thousands of trees. Therefore, we implemented simple iterative procedure to find the smaller value of this parameter in combination with the biggest classification accuracy. This procedure builds initial FRF consists of 3 FDTs. On the start of each iteration, one decision tree is added to FRF and accuracy is evaluated. This procedure finishes when classification accuracy stops growing.

5 Evaluation of classifiers

After preliminary data transformation we obtained data described by 8 fuzzy input attributes and one output attribute. According to this data, we create fuzzy classifiers. In this section is described, how classification performance of used classifiers was evaluated. To evaluate classification performance of implemented classifiers, we used three metrics: accuracy, sensitivity, and specificity. Sensitivity and specificity are important in case, when data is not evenly distributed into classes. The definition of used metrics is showed in the table below.

Table 2. Metrics for evaluation of classification performance

Accuracy	Sensitivity	Specificity
$\dfrac{TP + TN}{TN + TP + FN + FP}$	$\dfrac{TP}{TP + FN}$	$\dfrac{TN}{TN + FP}$

where TP stands for true positive, TN is true negative, FP is false positive, and FN is false negative. True positive and true negative results are obtained after correct classification while false negative and false positive stands for misclassification results. To evaluate classifiers, we divide data into training and testing set in ratio 70:30. Training set was used to create classifier and testing set was used to evaluate it. We repeated this for 10 000 times for each training set.

FRF needs to have defined an exact number of trees. Therefore, we analyzed how number of trees can affect classification accuracy of FRF. In this analysis, the forest from 3 trees was created and its accuracy was evaluated. Then, the number of trees was increased by one, and performance was re-evaluated. Adding of trees was repeated while the accuracy of the forest was raising.

5.1 Evaluation of Algorithms for Classification of EEG Signals using Fuzzy Classifiers

Evaluation of used classifiers was performed by two experiments. In the first experiment, the classification was aimed to recognize patients who suffer from epilepsy. The goal of the second experiment was to find patients during epileptic's seizure. To solve this task, the database was divided into two classes. In case of the first experiment, healthy persons and persons suffering from epilepsy (AB, CDE). In the second experiment, we were aimed to separate persons during seizure activity and persons in seizure-free interval (ABCD, E). For each experiment, the analysis of number of trees necessary for accurate prediction of FRF classifier was performed. Results obtained by proposed algorithm are shown in Table 3.

Table 3. Table of obtained results

Classifier	Predicated	Accuracy	Sensitivity	Specificity
uFDT	Seizure	0.991	0.996	0.986
oFDT	Seizure	0.980	0.992	0.966
FRF	Seizure	0.995	0.999	0.992
uFDT	Epilepsy	0.906	0.944	0.846
oFDT	Epilepsy	0.897	0.942	0.821
FRF	Epilepsy	0.894	0.946	0.818

Implemented FRF requires to have a specified number of trees as input parameter. This parameter affects computational complexity of aforementioned FRF. In case of large number of trees, total time necessary for FRT induction can be large. Also, classification by thousands of trees can be complex. Therefore, the setting of this parameter to proper value is necessary. To find proper value of this parameter, the FRF from 3 trees was inducted and its accuracy was estimated. Then, the number of trees was increased by one, and performance was evaluated again. Adding of trees was repeated while the accuracy of forest was raising. The results obtained by this analysis are shown on plots bellow. There are shown four plots. The plots show classification accuracy of FRF depending on the number of inducted FDT. The x axes of these plots stand for number of trees. The y axes sands for classification accuracy. Plots on the left side show accuracy for FRF consisted of 3 to 100 trees. Plots on the right side are plots created from the same data, but these plots show the accuracy of FRF consisted from 82 to 100 trees.
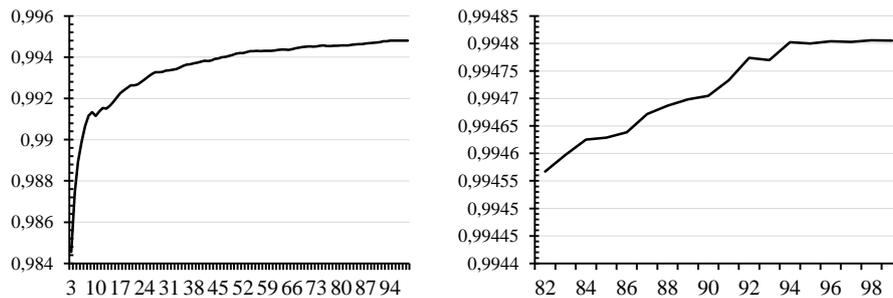

Fig. 3. Plots show accuracy (y axis) of FRF depending on the number of inducted FDT (y axis) for epileptics' seizure prediction.
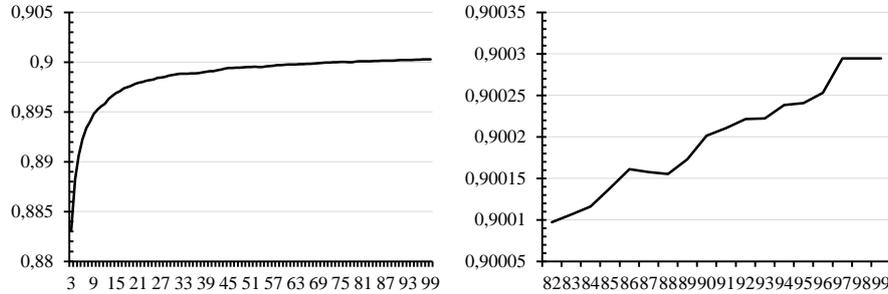
Fig. 4. Plots show accuracy (y axis) of FRF depending on the number of inducted FDT (y axis) for epileptics' detection.

To compare results of our investigation with similar studies, we created the comparison showed in Table 4. These tables include only studies which used the same data as we used. The classification of selected studies was targeted on prediction of occurrences of the epileptic's seizures. Differences between selected studies are in its preliminary data transformation and classification methods, eq. attribute extraction, dimensional reduction or classification algorithm. The comparison is shown in Table 4, where the first column contains first authors of studies and the reference to the paper, the second describes used methods in preliminary data transformation, third columns tells which classifier was used and the last column shows acquired classification accuracy.

Table 4. Table of results from other studies for epileptic's seizure detection

Study	Preprocessing methods	Classifier	Accuracy
K. Polat and S. Güneş [12]	DWT	C4.5	98.76%
L. Guo, D. Rivero [40]	DWT	MLPNN	97.77%
M. A. Naderi, H. Mahdavi-Nasab [9]	FFT, PCA	MLPNN	100%
U. Orhan, M. Hekim [7]	DWT	K-means and MLPNN	96.67%
N. Nicolaou, J.G. Kios [11]	Permutation Entropy	SVM	94.38%
K.Polat and S.Güneş[20]	FFT, PCA	AIRS classifier	99.81%
J. B. Jian, B. Goparaju [41]	CEEMD domain	RF	98.00%
This study	**Welch, PCA, Fuzzification**	**FRF**	**99.48%**

Conclusion

In this paper, the algorithm for EEG signal classification based on fuzzy logic was proposed. Previously, we solve task of EEG signals FDTs in [14] and [15], where the analysis of different fuzzification techniques was performed in detail. In this paper, the algorithm based on fuzzy entropy according to [32] was used. Involving of fuzzy logic permits to take into account the ambiguity of data that can arise in step of preliminary transformation (the feature extraction and the dimension reduction). The modification resulted in adding of fuzzification as new step in the preliminary data transformation. According to data obtained after preliminary data transformation, we compare three type of fuzzy classifiers: oFDT, uFDT, and FRF. These classifiers were inducted based on estimation of the cumulative mutual information [33]. In case of FRF, the impact of number of trees to classification accuracy has been analyzed. Results shows, that usage of proposed preliminary transformation can improve the classification accuracy of the EEG signals. Therefore, involving of fuzzy logic can be considered as recommended step for future development in EEG classification.

In future development of proposed classification algorithm, the impact of another classifier can be evaluated. Also, the modification of preliminary data transformation is possible. There are numerous ways to provide spectral transformation, dimensional reduction or fuzzification. The crucial goal of this development is increasing the classification accuracy of EEG signals.

References

1. WHO, http://www.who.int/mediacentre/factsheets/fs999/en/ last accessed 11/30/2018
2. Engel J. J., Starkman S.: Emergency Medicine Clinics of North America. Emerg. Med. Clin. North Am. 12, 895–923 (1994).
3. Iasemidis, L.D.: Epileptic seizure prediction and control. Biomed. Eng. IEEE Trans. 50, 549–558 (2003).
4. Libenson, M.H.: Practical approach to electroencephalography. (2010).
5. Subasi, A., Erc, E.: Classification of EEG signals using neural network and logistic regression. Comput. Methods Programs Biomed. 87—99 (2005).
6. Andrzejak, R.G., Lehnertz, K., Mormann, F., Rieke, C., David, P., Elger, C.E.: Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state. Phys. Rev. E. Stat. Nonlin. Soft Matter Phys. 64, 061907 (2001).
7. Orhan, U., Hekim, M., Ozer, M.: EEG signals classification using the K-means clustering and a multilayer perceptron neural network model. Expert Syst. Appl. 38, 13475–13481 (2011).
8. Lin, C.-J., Hsieh, M.-H.: Classification of mental task from EEG data using

neural networks based on particle swarm optimization. Neurocomputing. 72, 1121–1130 (2009).

9. Naderi, M.A.: Analysis and classification of EEG signals using spectral analysis and recurrent neural networks. Biomed. Eng. (NY). 3–4 (2010).

10. Subasi, A., Gursoy, M.I.: EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Syst. Appl. 37, 8659–8666 (2010).

11. Nicolaou, N., Georgiou, J.: Detection of epileptic electroencephalogram based on Permutation Entropy and Support Vector Machines. Expert Syst. Appl. 39, 202–209 (2012).

12. Polat, K., Güneş, S.: A novel data reduction method: Distance based data reduction and its application to classification of epileptiform EEG signals. Appl. Math. Comput. 200, 10–27 (2008).

13. Guo, L., Rivero, D., Dorado, J., Munteanu, C.R., Pazos, A.: Automatic feature extraction using genetic programming: An application to epileptic {EEG} classification. Expert Syst. Appl. 38, 10425–10436 (2011).

14. Rabcan, J., Kvassay, M.: Electroencephalogram Signals Classification by Ordered Fuzzy Decision Tree. In: CEUR Workshop Proceedings. pp. 72–87. ICT in Education, Research, and Industrial Applications: Integration, Harmonization, and Knowledge Transfer, Kyiv, Ukraine (2017).

15. Rabcan, J., Kvassay, M.: Identification of Persons with Epilepsy from Electroencephalogram Signals using Fuzzy Decision Tree. In: Communications in Computer and Information Science. Springer (2018).

16. Androulidakis, I., Levashenko, V., Zaitseva, E.: An empirical study on green practices of mobile phone users. Wirel. Networks. 22, 2203–2220 (2016).

17. Zaitseva, E., Levashenko, V., Kvassay, M., Rabcan, J.: Application of ordered fuzzy decision trees in construction of structure function of multi-state system. (2017).

18. Jantan, H., Hamdan, A.R., Othman, Z.A.: Data Mining Classification Techniques for Human Talent Forecasting.

19. Maletic, J., Marcus, A.: Data Mining & Knowledge Discovery (2005).

20. Polat, K., Güneş, S.: Artificial immune recognition system with fuzzy resource allocation mechanism classifier, principal component analysis and FFT method based new hybrid automated identification system for classification of EEG signals. Expert Syst. Appl. 34, 2039–2048 (2008).

21. Martõ Ânez, A.M., Kak, A.C.: PCA versus LDA. IEEE Trans. Pattern Anal. Mach. Intell. 23, 228–233 (2001).

22. Ley, D.: Approximating process knowledge and process thinking: Acquiring workflow data by domain experts. Conf. Proc. - IEEE Int. Conf. Syst. Man Cybern. 3274–3279 (2011).

23. Gueorguieva, N., Georgiev, G.: Fuzzyfication of Principle Component Analysis for Data Dimensionalty Reduction. 2016 IEEE Int. Conf. Fuzzy Syst. 1818–1825 (2016).

24. Tsipouras, M.G., Exarchos, T.P., Fotiadis, D.I.: A methodology for automated fuzzy model generation. Fuzzy Sets Syst. 159, 3201–3220 (2008).

25. Zaitseva, E., Levashenko, V.: Construction of a reliability structure function

based on uncertain data. IEEE Trans. Reliab. 65, 1710–1723 (2016).

26. Vatankhah, M., Yaghubi, M.: Adaptive Neuro-fuzzy Inference System for Classification of EEG Signals Using Fractal Dimension. In: 2009 Third UKSim European Symposium on Computer Modeling and Simulation. pp. 214–218. IEEE (2009).

27. Sudirman, R., Koh, C., Safri, N.M., Daud, W.B., Mahmood, N.H.: EEG different frequency sound response identification using neural network and fuzzy techniques. 2010 6th Int. Colloq. Signal Process. its Appl. 1–6 (2010).

28. Hinojosa, J., Domenech-Asensi, G.: Multiple adaptive neuro-fuzzy inference systems for accurate microwave CAD applications. 2007 18th Eur. Conf. Circuit Theory Des. 767–770 (2007).

29. Gupta, H.R., Mehra, R.: Power Spectrum Estimation using Welch Method for various Window Techniques. Int. J. Sci. Res. Eng. Technol. 2, 389–392 (2013).

30. Smith, L.I.: A tutorial on Principal Components Analysis Introduction. Statistics (Ber). 51, 52 (2002).

31. Jackson, D.A.: Stopping rules in principal components analysis - a comparison of heuristical and statistical approaches. Ecology. 74, 2204–2214 (1993).

32. Lee, H.-M., Chen, C.-M., Chen, J.-M., Jou, Y.-L.: An efficient fuzzy classifier with feature selection based on fuzzy entropy. IEEE Trans. Syst. Man, Cybern. Part B Cybern. 31, 426–432 (2001).

33. Levashenko, V., Zaitseva, E.: Usage of New Information Estimations for Induction of Fuzzy Decision Trees. Lect. Notes Comput. Sci. 2412, 493–499 (2002).

34. Podgorelec, V., Kokol, P., Stiglic, B., Rozman, I.: Decision Trees: An Overview and Their Use in Medicine. J. Med. Syst. 26, 445–463 (2002).

35. De Matteis, A.D., Marcelloni, F., Segatori, A.: A new approach to fuzzy random forest generation. In: 2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). pp. 1–8. IEEE (2015).

36. Rokach, L.: Ensemble-based classifiers. Artif. Intell. Rev. 33, 1–39 (2010).

37. Breiman, L.: Bagging Predictors. 24, 123–140 (1996).

38. Breiman, L.: Random Forests. Mach. Learn. 45, 5–32 (2001).

39. Rabcan, J., Levashenko, V., Zaitseva, E., Chovancova, O.: Generation of Structure Function Based on Ambiguous and Incompletely Specified Data Using Fuzzy Random Forest. In: 2018 IEEE 9th Int. Conf. on Dependable Systems, Services and Technologies (DESSERT). pp. 418–423. IEEE (2018).

40. Guo, L., Rivero, D., Dorado, J., Rab, J.R., Pazos, A.: Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks. J. Neurosci. Methods. 191, 101–109 (2010).

41. Jia, J., Goparaju, B., Song, J., Zhang, R., Westover, M.B.: Automated identification of epileptic seizures in EEG signals based on phase space representation and statistical features in the CEEMD domain. Biomed. Signal Process. Control. 38, 148–157 (2017).