Comparison of the EEG Signal Classifiers LDA, NBC and GNBC Based on Time-Frequency Features

Beata Szuflitowska, Przemysław Orłowski
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie, ul. Sikorskiego 37, 70-313, Szczecin

Abstract: EEG signals are non-stationary and used to study the activities of the brain in pathology. Epilepsy belongs to the most common neurological diseases. In the paper, real EEG sequences described by a doctor as normal and epileptic (ictal and interictal) are used. In classification process these sequences are divided into training and testing subsets. The classification are performed using Short-Time Fourier Transform. Based on obtained spectrum four features have been extracted. The study presents experiments based on the analysis and classification of EEG signals using various methods, including Linear Discriminant Analysis, Naive Bayes Classifier and Gaussian Naive Bayes Classifier. The results indicated that used techniques a potential to be applied within an automatic neurologic diseases diagnosis system and could thus further increase the number of correct diagnoses.

Keywords: electroencephalograph classification, machine learning, short-time Fourier transform, time-frequency analysis

1. Introduction

EEG signals are used to study the activities of the brain such in pathology, as detection of epileptic seizure from other types of signal such as Alzheimer's disease (AD) and in Brain-Computer Interface [1–8]. The detection of neurological disease by visual scanning of a patient's EEG data is a time-consuming process. Additionally, long-term treatment with antiepileptic drugs can change the EEG activity. The detection of pathology usually is possible only in long-term EEG recordings, which significantly prolongs the time of correct diagnosis. Automated systems can shorten the waiting time for diagnosis. The main problems relate to the elimination of artifacts and the analysis of EEG signal characteristics. A large part of the artifacts may imitate pathological stages such as epilepsy or dementia. Visual inspection may be insufficient. In modern medical systems more emphasis is put on improvement of hardware and software. Time, frequency and time-frequency analysis of EEG signals have emerged in recent years [9–16]. Biological signals are non-stationary. Therefore, time-frequency decomposition are a central part of EEG data analysis [5, 16]. The time-frequency representation family is large and includes e.g.:

Short-Time Fourier Transform, Gabor, Wigner-Vile Transform and Cone-Shaped Transform. Short-Time Fourier Transform (STFT) enables the extraction of information how its spectrum changes in the time [17]. The intensive research on EEG signal processing contributed to developing new classification method. Naive Bayes Classifier (NBC), Linear Discriminant Analysis (LDA) and Fuzzy Mutual Information (FMI) are the most common classifiers. The comparison of different kernels (Linear, Sigmoid and Grid) of Support Vector Machine (SVM) classifier on three classes (normal, interictal and ictal) with Approximate Entropy features extracted from multichannel EEG signals has been presented in [18]. The highest classification accuracy of 98.9% has been achieved using by Grid SVM kernel. The Approximate Entropy with Wavelet Transform has been used in [19]. The classifier is Surrogate Data Analysis. The obtained results (96.8%) are slightly worse. The Naive Bayesian Classifier is an algorithm based on clustering of data and applying Bayes theorem on the clustered data [17]. Rajaguru at al. proposed the use of the LDA and NBC to reduce the dimensions of the recorded EEG data. The highest accuracy (99.27%) is obtained when the dimensionality reduction technique used is LDA and classified with Particle Swarm Optimization Based Sparse Representation Classifier [20]. Yaik et al. applied Naive Bayesian Classifier for detection ictal (during seizure) and interictal (inter seizure) signals. The prediction accuracy of EEG signals is 95.14% using Naive Bayes Classifiers (94.10% for ictal and 97% for interictal) [21]. Sole-Casals and Vialatte showed that a strategy for semi-automatic cleaning based on blind source separation may improve the specificity of Alzheimer screening using markers of EEG artifacts: kurtosis, sample entropy, zero-crossing rate and fractal dimension [22]. Azami and al. classified the controls versus
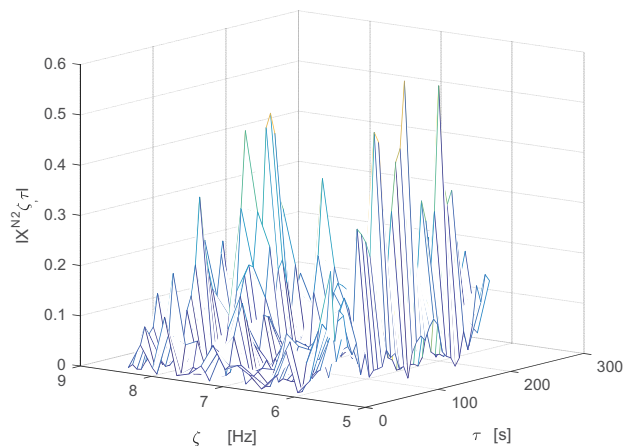
Fig. 1. Time-frequency EEG signal representation using STFT in normal
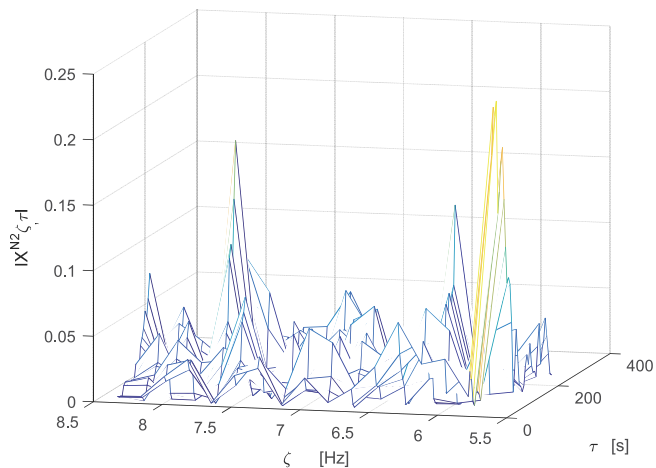Rys. 1. Czasowo-częstotliwościowa reprezentacja sygnału EEG z użyciem STFT w normie

Fig. 3. Time-frequency EEG signal representation using STFT in selected interictal sequence
Rys. 3. Czasowo-częstotliwościowa reprezentacja sygnału EEG z użyciem STFT wybranej międzynapadowej sekwencji
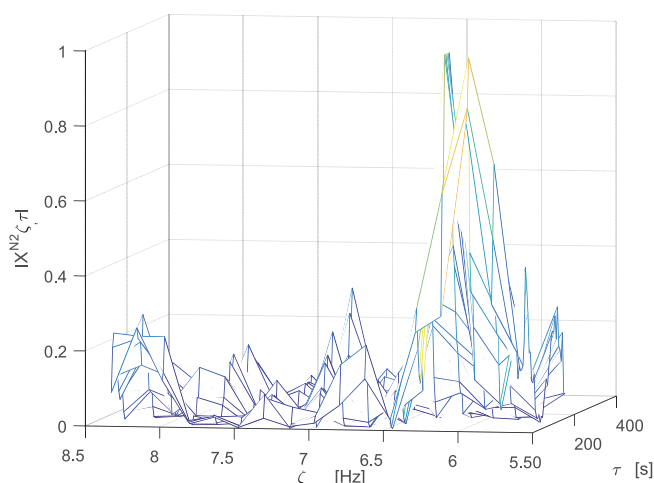

Fig. 2. Time-frequency EEG signal representation using STFT – occurrence of epileptic seizure
Rys. 2. Czasowo-częstotliwościowa reprezentacja sygnału EEG z użyciem STFT – wystąpienie napadu padaczkowego

AD subjects using the NBC. The classification ratios of average of slope values of the multiscale entropy and multivariate entropy profile below 72% have been obtained [23].

An automated system able to accurately differentiate between normal and interictal EEG signals can be used to diagnose epilepsy, while a system that can accurately differentiate between interictal and ictal EEG signals can be used to detect seizures. The purposes of this paper shall be discussing Short-Time Fourier transform of EEG feature extraction methods and comparing Naive Bayes Classifier and Linear Discriminant Analysis to detection three diagnostic cases: normal, ictal (during seizure) and interictal. Different sizes of training and testing group are used.

2. Signal EEG processing

The EEG data used in this study come from the Epilepsy Centre at the University of Bonn. The EEG data sets were recorded with a 128-channel amplifier system. The EEG database including 300 sequences were determined by the doctor as normal and epileptic: ictal and interictal. The EEG dataset corresponding to healthy subjects was taken from the EEG recordings of 5 healthy persons, who were relaxed in an awaken state, using the standardized electrode placement technique.

The epileptic EEG signals were recorded from 5 epileptic patients and were taken from the intracranial EEG recordings during presurgical diagnosis. Electroencephalogram artifacts were rejected by visual inspection. The all EEG sequences were cut out from the continuous EEG recordings and divided into time sequence equal to 23.6 s, that according to the duration of seizure. Sampling time was equal to 0.0057 s. The length of each recording is 4097 samples [24]. The signals are analyzed using Short-Time Fourier Transform. The STFT is obtained by applying the Fourier Transform by a fixed-sized, rectangular moving window to input series. Short-Time Fourier Transform of a signal $X(t)$ in the position (τ, ξ) in the time-frequency surface t/f can be defined as [5]:

$$X\left(\tau_n, \xi_k\right) = T_s \sum_{i=0}^{N-1} x\left(t_i\right) \varphi\left(\tau_n - t_i\right) e^{-j2\pi\xi_k t_i} \tag{1}$$

where: T_s – sampling time $t_i = iT_s$, τ_n – time shift $\tau_n = nT_s$, ξ_k – frequency shift $\xi_k = k/NT_s$.

In this paper, rectangular time window with width $w = 2^{10}T_s$ has been considered. All extracted EEG signal segments are additionally normalized between 0 and 1:

$$X_{\tau,\zeta}^N = \frac{\left|X_{\tau,\zeta}\right|}{\max\left|X_{\tau,\zeta}\right|} \tag{2}$$

After calculation of STFT of EEG samples, square of the absolute value of $X_{\tau,\xi}^N$ has been calculated.

A shift of EEG signal energy from lower to higher frequency bands before and during a seizure is common phenomenon, which is widely described in literature. Some changes in the EEG signal, which are not so obvious in the original full-spectrum signal, can be amplified when sub-band is considered independently [4]. The biggest changes in values of for three analyzed stages: normal, ictal and interictal have been observed for 5.6–8.3 Hz domain. In Figures 1–3 exemplary values of $X_{\tau,\xi}^N$ in three cases: normal, ictal and interictal have been presented. The healthy segment are treated as a reference. The highest values of $X_{\tau,\xi}^N$ are often observed in ictal and the lowest in interictal. These changes are repeated in time-frequency EEG signal representations.

Based on these observations, the following features from the considered time-frequency space have been extracted:

1) Maximal value of integral of average the magnitude in segment from 5.6 Hz to 8.3 Hz in the frequency domain over the time shift segment from 0 s to 230 s of the STFT spectrum:

$$f_{\mathrm{I}} = \max_{1 \leq n \leq 40} \left(\sum_{k=34}^{50} \left| X_{n,k}^N \right|^2 \right) \cong \frac{1}{T_s} \max_{0 \leq \tau \leq 230} \left(\int_{5.6}^{8.3} \left| X_{\tau,\xi}^N \right|^2 d\xi \right) \qquad (3)$$

2) Minimal value of integral of average the magnitude in segment from 5.6 Hz to 8.3 Hz in the frequency domain over the time shift segment from 0 s to 230 s of the STFT spectrum:

$$f_{\mathrm{II}} = \min_{1 \leq n \leq 40} \left(\sum_{k=34}^{50} \left| X_{n,k}^N \right|^2 \right) \cong \frac{1}{T_s} \min_{0 \leq \tau \leq 230} \left(\int_{5.6}^{8.3} \left| X_{\tau,\xi}^N \right|^2 d\xi \right) \qquad (4)$$

3) Variance of integral of average the magnitude in segment from 5.6 Hz to 8.3 Hz in the frequency domain over the time shift segment from 0 s to 230 s of the STFT spectrum:

$$f_{\mathrm{III}} = \operatorname*{var}_{1 \leq n \leq 40} \left(\sum_{k=34}^{50} \left| X_{n,k}^N \right|^2 \right) \cong \frac{1}{T_s} \operatorname*{var}_{0 \leq \tau \leq 230} \left(\int_{5.6}^{8.3} \left| X_{\tau,\xi}^N \right|^2 d\xi \right) \qquad (5)$$

4) Median of integral of average the magnitude in segment from 5.6 Hz to 8.3 Hz in the frequency domain over the time shift segment from 0 s to 230 s of the STFT spectrum:

$$f_{\mathrm{IV}} = \operatorname*{M}_{1 \leq n \leq 40} \left(\sum_{k=34}^{50} \left| X_{n,k}^N \right|^2 \right) \cong \frac{1}{T_s} \operatorname*{M}_{0 \leq \tau \leq 230} \left(\int_{5.6}^{8.3} \left| X_{\tau,\xi}^N \right|^2 d\xi \right) \qquad (6)$$

3. Classification

Classifiers is an algorithm which allocates observations to defined classes based on features of these observations. Two basic steps: training and testing in building classifier have been highlighted. Training is the process of taking content that is known to fit to describe classes and creating a classifier on the basis of that known content. In testing process given model to determined new data has been used. As described in section 2, EEG signals are divided into 300 sequences. Each sequence properly labels from the set {normal (0), ictal (1) and interictal (2)} has been assigned. In total, four training and testing subsets presented in Table 1 are analyzed. Testing subsets: S1B and S2B contain of all EEG sequences. Separable testing and training in sets 3 and 4 are considered.

Distribution of features $f_{\mathrm{I}} f_{\mathrm{II}}$ from subset S1A are shown in Figure 4. Ictal sequences are marked green, interictal class signals are marked blue and normal in red. Bellowing figure presents that in a two-dimensional representation some of the points belonging to the normal and ictal classes are located very close to each other. A lot of points from interictal class are even in
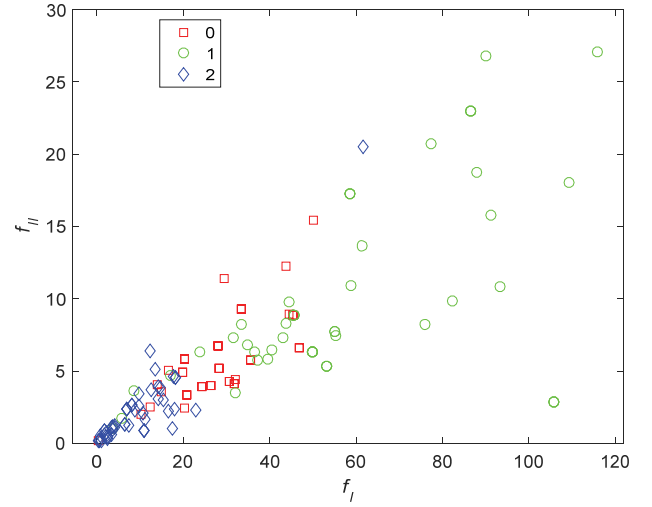

Fig. 4. Distribution of features $f_{\mathrm{I}} f_{\mathrm{II}}$ for subset S1A
Rys. 4. Rozkład cech $f_{\mathrm{I}} f_{\mathrm{II}}$ dla podzbioru S1A

the normal class. The points belong to epileptic class are better separated. A number of points are far from this discrimination function. The distance of these points is indicative of the severity of the brain state.

3.1. Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis belongs to feature reduction techniques. LDA is used to find a linear combination of features that can better separate two or more classes. Let us assume that we have 3 classes, each containing N observations $\boldsymbol{x}_i$. The measure of within-class scatter $\boldsymbol{S}^c$ for the class c can be denoted as:

$$\boldsymbol{S}^c = \sum_{i=1}^{N} \left(\boldsymbol{x}_i^c - \boldsymbol{\mu}^c \right) \left(\boldsymbol{x}_i^c - \boldsymbol{\mu}^c \right)^{\mathrm{T}} \qquad (7)$$

where $\boldsymbol{\mu}^c$ indicates mean of the all observation $\boldsymbol{x}_i$ for c-th class.

Generalization of the within-class scatter $\mathbf{S}_W$ for all 3 classes can be estimated as:

$$\mathbf{S}_W = \sum_{i=1}^{3} \frac{n_i}{N} \boldsymbol{S}^i \qquad (8)$$

where n_i is the number of $\boldsymbol{x}_i$ observations in each class and N is a total number of all observations.

The between-class scatter matrix for class c can be calculated as:

$$\boldsymbol{S}_B^C = \sum_{i=1}^{3} \left(\boldsymbol{x}^i - \boldsymbol{\mu} \right) \left(\boldsymbol{x}^i - \boldsymbol{\mu} \right)^{\mathrm{T}} \qquad (9)$$

where $\boldsymbol{\mu}_i$ indicates the mean of the all observations $\boldsymbol{x}_i$ for i-th class and $\boldsymbol{\mu}$ indicates the mean of the all observations $\boldsymbol{x}_i$ for all classes.

Generalization of between-class scatter $\mathbf{S}_m$ for all three classes can be written as:

Table 1. EEG data used for classification
Tabela 1. Dane EEG zastosowane podczas klasyfikacji

Class	Set 1		Set 2		Set 3		Set 4	
	S1A	S1B	S2A	S2B	S3A	S3B	S4A	S4B
Normal (0)	50	100	25	100	50	50	25	75
Ictal (1)	50	100	25	100	50	50	25	75
Interictal (2)	50	100	25	100	50	50	25	75
Total	150	300	75	300	150	150	75	225

$$S_m = \sum_{i=1}^{3} \frac{n_i}{N} S_B^i \qquad (10)$$

where n_i means the number of x_i observations in each class and N is a total number of all observations.

The total scatter matrix S_T can be denoted as:

$$S_T = S_W + S_m \qquad (11)$$

Orientation W separates the projected feature vectors of a class with other classes. The value of W can be calculated by maximizing the Fisher's criterion function $J(W)$. The Fisher's criterion depends on three factors W, S_W and S_B:

$$J(W) = \left| W^T S_m W \right| / \left| W^T S_W W \right| \qquad (12)$$

where $| \cdot |$ is a determinant.

The orientation matrix W is the solution of below problem:

$$S_W^{-1} S_m w_i = \lambda_i w_i \qquad (13)$$

where w_i (for $i = 1, ..., h$) are the column vectors of W that correspond to the largest eigenvalues (λ_i) [25–27].

Six pairs of features generated as described in Section 2 are used for classification. Classification process takes place in two stages. First, subsets S1A, S2A, S3A and S4A considered in Table 1 are used as training data for LDA classifier. Obtained training results are shown without brackets in the Table 2. Next, subsets S1B, S2B, S3B and S4B are used for testing classifier. Classification accuracy is presented in brackets in Table 2. Classification performance for each considered subset is summarized in the following columns of the Table 2. For two separable pairs of training and testing subsets (S3A and S3B, S4A and S4B) better classification results have been achieved than pairs of subsets S1A and S1B, S2A and S2B. Testing subsets S1B and S2B contain of all EEG sequences. Classification performance is decreased with decreasing training subset size. The rows of the Table 2 contain results obtained for used pairs of features. For most pairs of features comparable and satisfactory results have been obtained. Two pairs of features $f_I f_{III}$ and $f_I f_{IV}$ worsen accuracy significantly. Figure 5 shows distribution of features $f_I f_{II}$ for subset S1A. Additionally, results in form of confusion matrix are presented for separable subsets S3A (without brackets) and S3B (in brackets) in Table 3. Overlapping ictal and normal is clearly visible in Figure 4. Consequently, more errors are committed during classification ictal and normal sequences than interical and ictal signals.
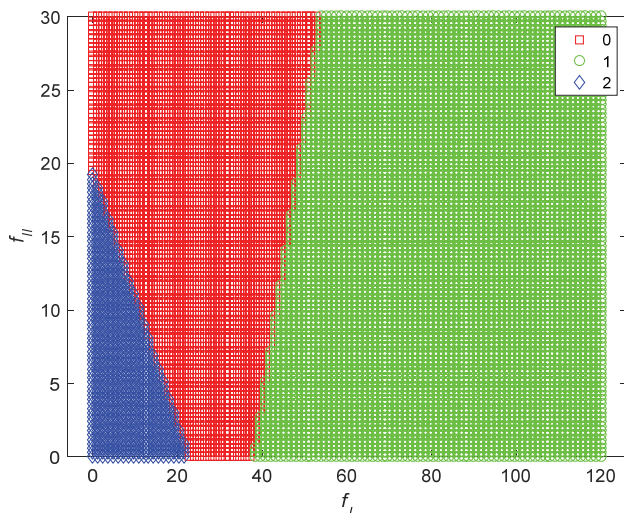

Fig. 5. Distribution of features $f_I f_{II}$ using LDA and subset S3A
Rys. 5. Rozkład cech $f_I f_{II}$ przy zastosowaniu LDA oraz podzbioru S3A

Table 2. Correct classifications (%) obtained for LDA and different pairs of features
Tabela 2. Prawidłowe klasyfikacje (%) otrzymane dla LDA oraz różnych par cech

No.	Features	Correct classifications (%)				
		S1A (S1B)	S2A (S2B)	S3A (S3B)	S4A (S4B)	Mean
1	$f_I f_{II}$	**87** (83)	**79** (78)	**86** (84)	**76** (73)	**82** (79.5)
2	$f_I f_{III}$	79 (78)	68 (63)	77 (74)	67 (65)	73 (70)
3	$f_I f_{IV}$	78 (76)	67 (64)	76 (73)	66 (64)	72 (69)
4	$f_{II} f_{III}$	85 (83)	75 (71)	82 (80)	74 (70)	79 (76)
5	$f_{II} f_{IV}$	83 (80)	73 (72)	81 (80)	72 (71)	77 (76)
6	$f_{III} f_{IV}$	84 (81)	74 (72)	82 (80)	73 (71)	78 (76)
	Mean	**83** (80)	73 (70)	81 (78.5)	71 (69)	–

Table 3. Confusion Matrix obtained using features $f_I f_{II}$ and subsets S3A and S3B (in brackets)
Tabela 3. Macierz błędu otrzymana z użyciem cech $f_I f_{II}$ dla podzbioru S3A oraz S3B (w nawiasach)

Class	Ictal	Normal	Interictal
Ictal	39 (35)	7 (8)	4 (7)
Normal	3 (4)	45 (45)	2 (1)
Interictal	4 (3)	1 (1)	45 (46)

3.2. Naive Bayes Classifier (NBC)

Naive Bayes Classifier is a simple and effective technique for classifier algorithm. The Bayes theorem allows one to calculate the a posteriori probability (the probability of a hypothesis considering a variable's value) based on the a priori probability (the frequency of each hypothesis) of both the data found and the total data according to:

$$P(v_j / A) = \frac{P(A / v_j) \times P(v_j)}{P(A)} \qquad (14)$$

where: v_j – the hypothesis $j = 1, 2, 3$ – in the set of hypotheses V, and A – the set of features $\langle x_1, x_2, ..., x_n \rangle$ describing the data.

In the case, when A has more than one attribute, it is necessary to estimate $P = (x_1, x_2, ..., x_n | v_j)$ in order to calculated $P = (v_j | x_1, x_2, ..., x_n)$. The estimate $P(A | v_j)$ is computationally costly, when the amount of samples is extremely large. It is possible to obtain a good classification performance event if attributes are not totally independent.

$$P = (x_1, x_2, ..., x_n / v_j) = \prod_i P(x_i / v_j) \qquad (15)$$

and the classifier output is given by:

$$v_{MAP} = \operatorname*{argmax}_{v_j \in V} \left\{ P(v_j) \times \prod_i P(x_i / v_j) \right\} \qquad (16)$$

where: v_{MAP} – the maximum *a posteriori* probability calculated within the space of hypotheses V [28, 29].

Table 4 presents classification accuracy using Naive Bayes Classifier calculated in an analogous manner as before. The use of NBC provides better classification accuracy in all analysed cases than LDA classifier. It can be clearly seen for features $f_I f_{III}$ and $f_I f_{IV}$. Distribution of features $f_I f_{II}$ for subset S3A is depicted in Figure 6. Table 5 shows confusion matrix achieved using features $f_I f_{II}$ for subsets S3A and S3B. NBC classifier improves classification of ictal sequences and slightly deteriorates detection interictal sequences compared with LDA.

Table 4. Correct classifications (%) obtained for NBC and different pairs of features
Tabela 4. Prawidłowe klasyfikacje (%) otrzymane dla NBC oraz różnych par cech

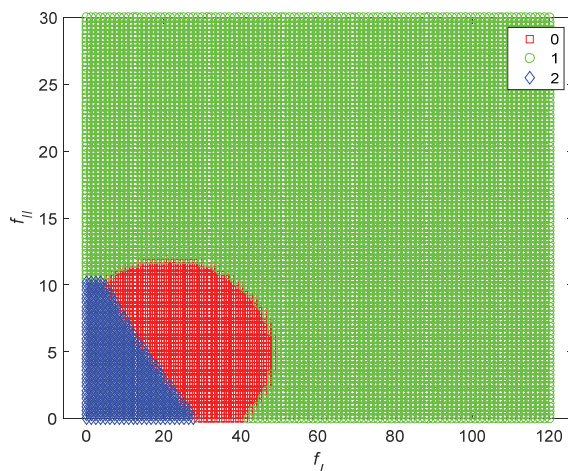| No.	Fea-tures	Correct classifications (%)				
		S1A (S1B)	S2A (S2B)	S3A (S3B)	S4A (S4B)	Mean
1	$f_I f_{II}$	**90** **(87)**	**82** **(79)**	**88** **(85)**	**80** **(77)**	**84** **(81)**
2	$f_I f_{III}$	82 (79)	72 (70)	82 (80)	71 (71)	77 (75)
3	$f_I f_{IV}$	83 (79)	71 (69)	83 (81)	70 (67)	77 (74)
4	$f_{II} f_{III}$	87 (85)	79 (77)	86 (83)	75 (74)	80.5 (78)
5	$f_{II} f_{IV}$	88 (84)	81 (78)	85 (82)	76 (74)	81 (78.5)
6	$f_{III} f_{IV}$	88 (86)	82 (79)	86 (84)	78 (76)	82 (80)
	Mean	**86** **(83)**	74.5 (72)	85 (83)	75 (73)	–

Fig. 6. Distribution of features $f_I f_{II}$ using NBC and subset S3A
Rys. 6. Rozkład cech $f_I f_{II}$ przy zastosowaniu NBC oraz podzbioru S3A

Table 5. Confusion matrix obtained using features $f_I f_{II}$ and subsets S3A and S3B (in brackets)
Tabela 5. Macierz błędu otrzymana z użyciem cech $f_I f_{II}$ dla podzbioru S3A oraz S3B (w nawiasach)

Class	Ictal	Normal	Interictal
Ictal	45 (43)	3 (4)	2 (3)
Normal	1 (4)	45 (43)	4 (3)
Interictal	6 (7)	2 (2)	42 (41)

3.3. Gaussian Naive Bayes Classifier (GNBC)

The extension of Naive Bayes is called Gaussian Naive Bayes Classifier (GNBC). The likelihood of the features is assumed to be Gaussian [30].

$$P\left(x_i \ / \ \nu_j\right) = \frac{1}{\sqrt{2\pi\sigma_{\nu_j}^2}} exp\left[-\frac{\left(x_i - \sigma_{\nu_j}\right)^2}{2\sigma_{\nu_j}^2}\right] \quad (17)$$

Table 6 shows classification results using Gaussian Naive Bayes Classifier. More classification errors are committed than for NBC. Comparable accuracy is achieved in the most cases as for LDA classifier. GNBC improves significantly average classification performance calculated for subset 4 and pairs of features $f_I f_{III}$ and $f_I f_{IV}$ by comparing with LDA. The distribution of features $f_I f_{II}$ for subset S1A is presented in Figure 7. Table 7 contains results in form confusion matrix for subsets S3A (in brackets) and S3B (without brackets). Based on the Table 7, it can be concluded that detection ictal sequences is better than for LDA. GNBC gives the worst classification results for sequences belong to interictal class.

Table 6. Correct classifications (%) obtained for GNBC and different pairs of features
Tabela 6. Prawidłowe klasyfikacje (%) otrzymane dla GNBC oraz różnych par cech

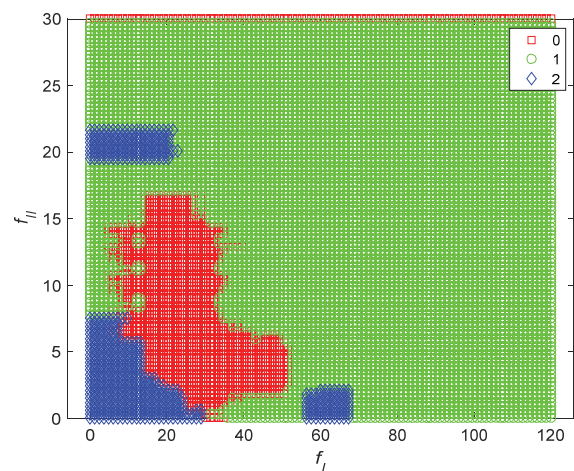| No.	Features	Correct classifications (%)				
		S1A (S1B)	S2A (S2B)	S3A (S3B)	S4A (S4B)	Mean
1	$f_I f_{II}$	**86** **(83)**	**81** **(78)**	**84** **(81)**	**80** **(79)**	**83** **(80)**
2	$f_I f_{III}$	81 (78)	76 (75)	80 (76)	75 (74)	78 (76)
3	$f_I f_{IV}$	82 (80)	74 (73)	80 (77)	73 (74)	77 (76)
4	$f_{II} f_{III}$	84 (82)	81 (80)	82 (79)	79 (76)	81.5 (79)
5	$f_{II} f_{IV}$	84 (81)	80 (78)	83 (79)	79 (73)	81.5 (78)
6	$f_{III} f_{IV}$	85 (83)	79 (76)	84 (80)	78 (78)	81.5 (79)
	Mean	**84** **(81)**	78.5 (77)	82 (79)	77 (76)	–

Fig. 7. Distribution of features $f_I f_{II}$ using GNBC and subset S3A
Rys. 7. Rozkład cech $f_I f_{II}$ przy zastosowaniu GNBC oraz podzbioru S3A

Table 7. Confusion matrix obtained using features f_I and f_{II} and subsets S3A and S3B (in brackets)
Tabela 7. Macierz błędu otrzymana z użyciem cech $f_I f_{II}$ dla podzbioru S3A oraz S3B (w nawiasach)

Class	Ictal	Normal	Interictal
Ictal	43 (42)	4 (5)	3 (3)
Normal	2 (3)	46 (44)	2 (4)
Interictal	8 (10)	5 (3)	37 (37)

4. Conclusion

The proposed methods are promising for three-class EEG signal classification. The main observations of this study are presented in following points:

1. The STFT features extraction procedure may appear as effective for the classification performance even the training data size is much smaller than size of testing data.

2. The best classification results are obtained using Naive Bayes Classifier for each analyzed set and each combination of used features, i.e. for features $f_I f_{II}$ and subset S3B classification accuracy is equal to 85%, 84% and 81% for NBC, LDA and GNBC, respectively. This dependency can be observed also for mean values of classification accuracy for all combinations of features in each considered subset (i.e. for S3B: 83%, 78.5% and 79% for NBC, LDA and GNBC, respectively) and for each pair of features in all subsets (i.e. for testing subsets and features $f_I f_{II}$: 81%, 79.5% and 80% for NBC, LDA and GNBC, respectively).

3. Gaussian Naive Bayes Classifier turned out to be the most stable classifier. For this classifier the highest classification accuracy equal to: 83%, 78%, 81% and 79% and the lowest – 78%, 73%, 76% and 73% have been achieved for subsets: S1B, S2B, S3B and S4B, respectively. The calculated differences between the best and the worst classification results for each considered set are equal to: 5%, 5%, 5% and 7%. In turn using LDA the best classification results are equal to: 83%, 78%, 84% and 73%, and the worst – 76%, 63%, 73% and 64% for subsets: S1B, S2B, S3B and S4B, respectively. The achieved differences are: 7%, 15%, 11% and 9% for each, consecutive subset. For NCB the calculated differences are equal to: 8%, 10%, 5% and 10%. Similar results calculated for average classification accuracy obtained for all combinations of features in each considered subset have been achieved: 4% (the highest accuracy – 80% and the lowest – 76%), 9.5% (the highest accuracy – 79.5% and the lowest – 69%) and 7% (the highest accuracy – 81% and the lowest – 74%) for GNBC, LDA and NBC, respectively. The differences calculated for all pairs of features in all subsets are equal to: 5% (the highest accuracy – 81% and the lowest – 76%), 10.5% (the highest accuracy – 80% and the lowest – 69.5%), and 11% (the highest accuracy – 83% and the lowest – 72%) for GNBC, LDA and NBC, respectively. These observations are confirmed during training classifiers.

4. Based on results presented in point 3, it can be concluded that Linear Discriminant Analysis is the most sensitive to the choice of combination of used features and considered subsets.

5. The achieved results indicate the possibility of using Short-Time Fourier transform successfully in the process of feature extraction and classification for neurological disorders.

6. Creating modern diagnostic tools based on accurate algorithms can have a significant impact on the improvement of making diagnosis in shorter time.

Bibliography

1. Oweis R.J., Abdulhay E.W., *Seizure classification in EEG signals utilizing Hilbert-Huang transform*, "Biomedical Engineering Online", 2011; 10: 38, DOI: 10.1186/1475-925X-10-38.

2. Dauwels J., Vialatte F., Cichocki A., *Diagnosis of Alzheimer's disease from EEG signals: where are we standing?* "Current Alzheimer Research", 6, 2010, 487–505.

3. Gajic D., Djurovic Z., Di Gennaro S., Gustafsson F., *Classification of EEG signals for detection of epileptic seizures based on wavelets and statistical pattern recognition.* "Biomedical Engineering: Applications, Basis and Communications", Vol. 26, Iss. 2, 2014, DOI: 10.4015/S1016237214500215.

4. Solé-Casals J., Benoît V., *An automatic detection of focal EEG signals using new class of time–frequency localized orthogonal wavelet filter banks,* "Knowledge-Based Systems", Vol. 118, 2017, 217–227, DOI: 10.1016/j.knosys.2016.11.024.

5. Kołodziej M., Majkowski A., Rak R., *Optymalizacja doboru okien czasowych do przetwarzania sygnału EEG w interfejsach mózg-komputer*, „Przegląd Elektrotechniczny", R. 87, Nr 9a, 2011, 142–144.

6. Szuflitowska B., Orłowski P., *Classification of electroencephalograph signals using time-frequency decomposition and linear discriminant analysis*, paper no 10445-253, accepted for publication in SPIE (the International Society for Optics and Photonics), Vol. 1045, ISBN: 9781510613546.

7. Durka P.J., *Matching Pursuit and Unification in EEG Analysis*, Artech House, 2007.

8. Nguyen T., Khosravi A., Creighton D., Nahavandi S., *EEG signal classification for BCI applications by wavelets and interval type-2 fuzzy logic systems,* "Expert Systems with Applications", Vol. 42, Iss. 9, 2015, 4370–4380, DOI: 10.1016/j.eswa.2015.01.036.

9. Orłowski P., *Simplified Design of Low-Pass, Linear Parameter-Varying, Finite Impulse Response Filters.* "Information Technology and Control", Vol. 39, No. 2, 2010, 130–137.

10. Orłowski P., *Fractional Indexes Impulse Responses Approximation for Discrete-Time Weyl Symbol Computation.* "Elektronika ir Elektrotechnika", Vol. 104, No. 8, 2010, 9–12, 2010.

11. Sobolewski A., Kublik E., Świejkowski D.A., Łęski S., Kamiński J.K., Wróbel A., *Cross-trial correlation analysis of evoked potentials reveals arousal related attenuation of thalamo-cortical coupling*, "Journal of Computational Neuroscience", Vol. 29, Iss. 3, 2010, 485–493, DOI: 10.1007/s10827-010-0220-0.

12. Łęski S., Lindén H., Tetzlaff T., Pettersen K.H., Einevoll G.T., *Frequency dependence of signal power and spatial reach of the local field potential*, "PLoS Computational Biology", Vol. 9, e1003137, 2013, DOI: 10.1371/journal.pcbi.1003137.

13. Patan K., Rutkowski G., *Analysis and classification of EEG data: an evaluation of methods*, "Lecture Notes in Artificial Intelligence", Vol. 7268, 2012, 310-317.

14. Iscan Z., Dokur Z., Tamer D., *Classification of electroencephalogram signals with combined time and frequency features*, "Expert Systems with Applications", Vol. 38, Iss. 8, 2011, 10499–10505, DOI: 10.1016/j.eswa.2011.02.11.

15. Oliva J.T., Garcia Rosa J.L., *How an epileptic EEG segment, used as reference, can influence a cross-correlation classifier?* "Applied Intelligence", Vol. 47, Iss. 1, 2017, 1–19, DOI: 10.1007/s10489-016-0891-y.

16. Rak R., Majkowski A., *Czasowo-częstotliwościowa analiza sygnałów*, „Przegląd Elektrotechniczny", R. 80, Nr 5, 2004, 515–520.

17. Bhattacharyya S., Khasnobish A., Chatterjee A., Konar A., *Performance analysis of LDA, QDA and KNN algorithms in left-right limb movement classification from EEG data*, [in:] International Conference on Systems in Medicine and Biology (ICSMB), Kharagpur, India. 16–18, Dec. 2010, 126–131, DOI: 10.1109/ICSMB.2010.5735358.

18. Vakludal V., Srinath S., Vijayapuriga S., *Automated Detection of Epileptic EEG Using Approximate Entropy in Elman Networks*, "International Journal on Recent Trends in Engineering and Technology", Vol. 1, No. 1, 2009, 307–312, DOI: 01.IJRTET.2009.01.01.1501.

19. Ocak H., *Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy.*

"*Expert Systems* with *Applications*", *Vol.* 36, Iss. 2, Part 1, 2009, 2027–2036, DOI: 10.1016/j.eswa.2007.12.065.

20. Rajaguru V., Harikumar T., Prabhakar S.K., Kumar S., *A Framework for Epilepsy Classification Using Modified Sparse Representation Classifiers and Naïve Bayesian Classifier from Electroencephalogram Signals*, "Journal of Medical Imaging and Health Informatics", Vol. 6, No. 8, 2016, 1829–1837, DOI: 10.1166/jmihi.2016.1935.

21. Yayik A., Yildirim E., Kutlu Y., Yildirim S., *Epileptic State Detection: Pre-ictal, Inter-ictal, Ictal*, "*International Journal of Intelligent Systems and Applications in Engineering*", Vol. 3, No. 1, 2015, 14–18.

22. Solé-Casals J., Vialatte F.B., *Towards Semi-Automatic Artifact Rejection for the Improvement of Alzheimer's Disease Screening from EEG Signals*, "Sensors", Vol. 15(8), 2015, 17963–17976, DOI: 10.3390/s150817963.

23. Azami H., Abásolo D., Simons S., Escudero H., *Univariate and Multivariate Generalized Multiscale Entropy to Characterise EEG Signals in Alzheimer's Disease*, "Entropy", Vol. 19, Iss. 1, 2017, 31–47, DOI: 10.3390/e19010031.

24. Andrzejak R.G., Lehnertz K., Rieke C., Mormann F., 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*, "Physical Review E", Vol. 64, 061907, 2001, DOI: 10.1103/PhysRevE.64.061907.

25. Sharma A., Paliwal K.K., *Linear discriminant analysis for the small sample size problem: an overview*, "*International Journal of Machine Learning and Cybernetics*", Vol. 6, Iss. 3, 2015, 443–454, DOI: 10.1007/s13042-013-0226-9.

26. Zhang R., Xu P., Guo L., Zhang Y., Li P., Yao D., Z-*Score Linear Discriminant Analysis for EEG Based Brain-Computer Interfaces*, PloS One, 9, 2013, 1-7, DOI: 10.1371/journal.pone.0074433.

27. Mitchel T., *Machine Learning*, McGraw-Hill Science: New York, NY, USA, 1997.

28. Machado J., Balbinot A., *Executed Movement Using EEG Signals through a Naive Bayes Classifier*, "Micromachines", Vol. 5, Iss. 4, 2014, 1082–1105, DOI: 10.3390/mi5041082.

29. Bielza C., Larrañaga P., *Bayesian networks in neuroscience: a survey*, "Frontiers in *Computational Neuroscience*", Vol. 8, 2014, 1–23, 10.3389/fncom.2014.00131.

30. Srivastava S., Gupta M.R., Frigyik B.A., *Bayesian Quadratic Discriminant Analysis*, "Journal of Machine Learning Research", Vol. 8, 2007, 1277–1305.

Porównanie klasyfikatorów LDA, NBC i GNBC sygnału EEG stosujących cechy czasowo-częstotliwościowe

Streszczenie: Sygnały EEG są z definicji niestacjonarne i stosowane do badania aktywności mózgu w patologii. Epilepsja należy do najczęstszych chorób neurologicznych. W pracy użyto rzeczywistych sekwencji EEG określonych przez lekarza jako stan normalny oraz padaczka (stany napadowe oraz międzynapadowe). W procesie klasyfikacji sygnały zostały podzielone na dwa podzbiory – uczący oraz testujący. Klasyfikacja została przeprowadzona za pomocą krótkotrwałej transformaty Fouriera. Na podstawie otrzymanego widma dokonano ekstrakcji czterech cech. Badanie przedstawia eksperymenty oparte na analizie i klasyfikacji sygnałów EEG za pomocą różnych metod, w tym Liniowej Analizy Dyskryminacyjnej, Naiwnego Klasyfikatora Bayesa oraz Naiwnego Klasyfikatora Bayesa dla rozkładu Gaussa. Wyniki pokazują, że użyty algorytm może być potencjalnie stosowany w automatycznej diagnostyce schorzeń neurologicznych i może w przyszłości zwiększyć liczbę poprawnie stawianych diagnoz.

Słowa kluczowe: elektroencefalografia, klasyfikacja, uczenie maszynowe, krótkoczasowa transformata Fouriera, analiza czasowo-częstotliwościowa

Beata Szuflitowska, MSc
beataszuflitowska@wp.pl

She graduated from the University of Szczecin and Pomeranian Medical University in Szczecin. Currently she is PhD student at West Pomeranian University of Technology, Szczecin. Research topics: electromagnetic field, brain-computer interface and model predictive control.

Przemysław Orłowski, PhD, DSc
przemyslaw.orlowski@zut.edu.pl

Associate Professor in the Department of Control and Measurements at West Pomeranian University of Technology Szczecin. The research topics are concerned on the analysis and synthesis of control systems, discrete-time systems, time-varying systems, nonlinear systems, uncertain systems and hybrid systems.