High-performance detection of epilepsy in seizure-free EEG recordings: A novel machine learning approach using very specific epileptic EEG sub-bands

Completed Research Paper

Ricardo Buettner
Aalen University
73430 Aalen, Germany
ricardo.buettner@hs-aalen.de

Janek Frick
Aalen University
73430 Aalen, Germany
janek.frick@visanox.com

Thilo Rieg
Aalen University
73430 Aalen, Germany
thilo.rieg@visanox.com

Abstract

We applied machine learning to diagnose epilepsy based on the fine-graded spectral analysis of seizure-free (resting state) EEG recordings. Despite using unspecific agglomerated EEG spectra, our fine-graded spectral analysis specifically identified the two EEG resting state sub-bands differentiating healthy people from epileptics (1.5-2 Hz and 11-12.5 Hz). The rigorous evaluation of completely unseen data of 100 EEG recordings (50 belonging to epileptics and the other 50 to healthy people) shows that the approach works successfully, achieving an outstanding accuracy of 99 percent, which significantly outperforms the current benchmark of 70% to 95% by a panel of up to three experienced neurologists. Our epilepsy diagnosis classifier can be implemented in modern EEG analysis devices, especially in intensive care units where early diagnosis and appropriate treatment are decisive in life and death scenarios and where physicians' error rates are particularly high. Our approach is accurate, robust, fast, and cost-efficient and substantially contributes to Information Systems research in healthcare. The approach is also of high practical and theoretical relevance.

Keywords: healthcare, medical data science, epilepsy, electroencephalography

Introduction

Since Information Systems scholars currently achieve accuracy levels far above 80% in the machine learning based detection of epileptic seizure epochs within electroencephalographic (EEG) recordings (Acharya et al. 2018; Wang et al. 2018b; Tjepkema-Cloostermans et al. 2018; Martinez-del-Rincon et al. 2017; Abdulhay et al. 2017; Li et al. 2017; Tiwari et al. 2017; Hassan et al 2016a, 2016b), these results have drawn a lot of attention especially from physicians, neuroscientists and neurophysiologists.

However, when a patient suffers from a current epileptic seizure state, this state is also obviously manifested in patients' unconsciousness, loss of awareness, abnormal sensory sensations, physical numbness or cramps. The application of detecting seizure epochs in seizure-loaded recordings is limited.

The ability to use an automated and accurate detection of epileptics based on seizure-free (resting state) EEG recordings instead of seizure-loaded recordings – which currently continues – would be of great use. Since most of the time epileptics are in seizure-free states, an automated EEG resting state diagnosis would be a breakthrough in epilepsy research, diagnosis and treatment. Worldwide there is no system or at least any approach that can accurately detect epilepsy based on seizure-free (resting state) EEG recordings.

Due to different data interpretations in a seizure-loaded EEG recording by physicians, the accuracy of the diagnosis of epilepsy varies from a misdiagnosis rate of 5% in childhood epilepsy – in which the diagnosis was made by a panel of three experienced pediatric neurologists – to at least 23% in adult populations (van Donselaar et al. 2006). The human error rate in epilepsy diagnosis based on larger recordings is even higher – up to 30% (van Donselaar et al. 2006; Zaidi et al. 2000). Therefore, automatic detecting of epilepsy in seizure-free EEG recordings with accuracy above 70% would be a highly useful contribution.

That is why we are interested in the following research question: Can a machine learning based algorithm quickly and accurately distinguish epileptics from healthy persons through the automated analysis of seizure-free EEG recordings?

As a result, we not only surpass the current human 70% threshold mentioned before, but the novel machine learning approach presented here beats the human 95% accuracy threshold (highest panel threshold of three experienced neurologists according to van Donselaar et al. [2006]), but based on seizure-free (resting state) EEG recordings instead of analyzing EEG recordings loaded with epileptic seizure activity.

We achieve this breakthrough by combining a very recent empirical finding from clinical neurophysiology with a machine learning based feature of a subset selection and subsequent algorithm design:

1. Pittau et al. (2018, 2016) recently found that epileptics also have very few (hidden) spikes in seizure-free (resting state) EEG recordings compared to healthy people, but the related spectral sub-bands are still unknown.
2. Using a Random Forest based subset feature selection we identified two EEG resting state sub-bands that differentiate healthy people from epileptics: While epileptics have more spectral power in the Mid-Delta sub-band 1.5-2 Hz (1.358 vs. 0.828, Cohen's $d = 2.155$, $p<0.01$), they have much less spectral power in the High-Alpha sub-band 11-12.5 Hz (0.270 vs. 0.783, Cohen's $d = 2.324$, $p<0.01$).
3. Based on the spectral power of the four fine-graded equidistant slices of the related sub-bands (1.5-2 Hz, 11-11.5 Hz, 11.5-12 Hz, 12-12.5 Hz) of an established high-quality epilepsy EEG dataset (Phys. Rev. E, 64, 061907) we trained and rigorously evaluated a decision tree to quickly and accurately distinguish epileptics from healthy persons.

The five most important contributions of our work are:

- We are the first to build a highly effective epilepsy diagnosis algorithm that achieves very good results based only on the analysis of seizure-free (resting state) EEG recordings.
- Our epilepsy diagnosis algorithm achieves an accuracy of 99%, which significantly outperforms the current benchmark of 95% by three experienced neurologists (van Donselaar et al. 2006).
- We extend the work by Pittau et al. (2018, 2016) in terms of identifying the two EEG resting state sub-bands differentiating healthy people from epileptics (1.5-2 Hz and 11-12.5 Hz).
- Our epilepsy diagnosis algorithm could be implemented in modern EEG analysis devices, especially in intensive care units where early diagnosis and appropriate treatment is imperative for life and death decisions and physicians' error rates are particularly high (van Donselaar et al. 2006; Zaidi et al. 2000; Donchin et al. 1995; Kirwan 1994).
- Our novel algorithm is accurate, robust, fast, and cost-efficient, which substantially contributes to Information Systems research in healthcare (see Romanow et al. 2012).

The novel algorithm diagnosing epilepsy based on seizure-free EEG recordings is of high practical relevance since

- epilepsy is one of the most common and serious brain diseases worldwide, causing a high level of individual suffering, discrimination, and social stigmatization (Moshé et al. 2015; Quintas et al. 2012)
- the high human error rate in epilepsy diagnosis based on EEG recordings (up to 30%) leads to incorrect treatments with heavy consequences for patients (Devinsky et al. 2018; Duncan et al. 2006; van Donselaar et al. 2006; Zaidi et al. 2000).

Beyond this high practical relevance the work is also of theoretical interest since the identification of the two bands 1.5-2 Hz and 11-12.5 Hz in resting state EEG which distinguish epileptics from healthy persons is useful to our understanding of epilepsy itself and stimulates further research.

The paper is organized as follows: Next we present an overview of the research background before providing the research methodology, including the applied machine learning methods and the data used for evaluation. After that we show the machine learning results including the performance evaluation. We then discuss the results and include practical and theoretical implications, before concluding with limitations and suggestions for future research.

Research Background

According to the updated OECD report (http://www.oecd.org/health/healthdata; issued on April 14, 2019), U.S. spending on health per capita increased from $8,223 USD (2010) to over $10,000 USD today (+ 21.6%). This immense amount of public health spending not only in the U.S. but also in all of the 34 OECD countries reinforced the call in the MISQ Editorial 36(2) by Romanow et al. (2012) for substantially more Information Systems research in healthcare. Information systems and technology outcomes continue to show a positive effect on medical outcomes (Aoe et al. 2019; Kruse and Beane 2018; Wang et al. 2018a; Öksüz et al. 2018; Abouzahra et al. 2015).

Clinical Picture of Epilepsy

Epilepsy is one of the most common and serious brain disorders. It is a neurological disease that leads to a permanent vulnerability towards epileptic seizures (Devinsky et al. 2018). The International League against Epilepsy (ILAE) defines an epileptic seizure as "a temporary occurrence of signs and/or symptoms due to abnormal excessive or synchronous neuronal activity in the brain" (Fisher et al. 2005). Epileptic seizures lead to a temporary behavioral disorder. A seizure can lead to unconsciousness, loss of awareness, abnormal sensory sensations, physical numbness or cramps (Devinsky et al. 2018; Duncan et al. 2006). Almost 10% of people suffer an epileptic seizure at some point in their lives (Devinsky et al. 2018). A total of 65 million people worldwide are affected by epilepsy (Devinsky et al. 2018). Epilepsy can lead to a drastic reduction in life expectancy. For example, patients with cryptogenic or idiopathic epilepsy have a reduced life expectancy of two years (Moshé et al. 2015). A reduction of up to ten years was observed in symptomatic epilepsy (Moshé et al. 2015). Due to the intensity of epileptic seizures and the shortened lifespan, epilepsy is associated with a high degree of suffering. Because of the lack of social education about epilepsy, discrimination, social stigmatization and misunderstandings are experienced by patients (Quintas et al. 2012).

The number of epileptics in poorer countries is significantly higher than in wealthy countries. In principle, epilepsy can be successfully treated with medication in most cases. However, these drugs are often not available in poorer countries or cannot be financed by those affected (Moshé et al. 2015). In addition, the disease is often not diagnosed or not diagnosed correctly in poorer countries. This diagnostic gap and the lack of medical care were also highlighted by the WHO (Duncan et al. 2006). The epilepsy paradox is also spoken of in this context: epilepsy has been known for thousands of years, millions of people are ill because of it and successful treatment has been possible for over 90 years. Nevertheless, it is assumed that 85% of treatable epileptics do not receive treatment (Kale 2002). However, the longer epilepsy is left undiagnosed and not treated, the lower the chance of long-term remission.

That is why a fast but accurate, robust, and cost-efficient diagnosis of epileptics is important.

Diagnosing of Epilepsy

Besides using cost-intensive and time-consuming brain imaging (fMRI, CT, PET, SPECT) and genetic techniques, the state-of-the-art method (device) of choice for diagnosing epilepsy is to analyze EEG recordings (Devinsky et al. 2018). EEG data provide relevant information on the correct diagnosis of the disease and on medication (Devinsky et al. 2018; Pohlmann-Eden and Newton 2008). The recorded EEG is examined for epileptic activity by searching for spikes, sharp waves and spike wave discharge. An accurate diagnosis is then made based on the clinical description of the seizures, the patient's age, comorbidities, and EEG patterns (Duncan et al. 2006).

While EEG abnormalities are present in a timeframe of about 48 hours after the epileptic seizure occurs, EEG abnormalities vanish after that critical timeframe (Schreiner and Pohlmann-Eden 2003). When seizure-loaded activity is visible in EEG recordings, the diagnosis is feasible. Since the likelihood of detecting EEG abnormalities is as high as 71% when the EEG is performed within 48 hours of the first seizure (Schreiner and Pohlmann-Eden 2003), a prompt EEG recording is important. After that timeframe, physicians have to wait till the next seizure activity.

This is a dilemma since a patient is already heavily suffering from epileptic seizures when in the so-called ictal stage (Fig. 1, right side). Before the first-time epileptic seizure occurs – and also between epileptic seizures (Fig. 1, left side) – a proper diagnosis has not to date been possible.

From a reliability point of view, the analysis and interpretation of ictal EEG recordings loaded with seizures (Fig. 1, right side) is already problematic. Even if the physicians are confronted with ictal EEG sessions loaded with seizures the accuracy of correcting diagnosing epilepsy is not high enough.

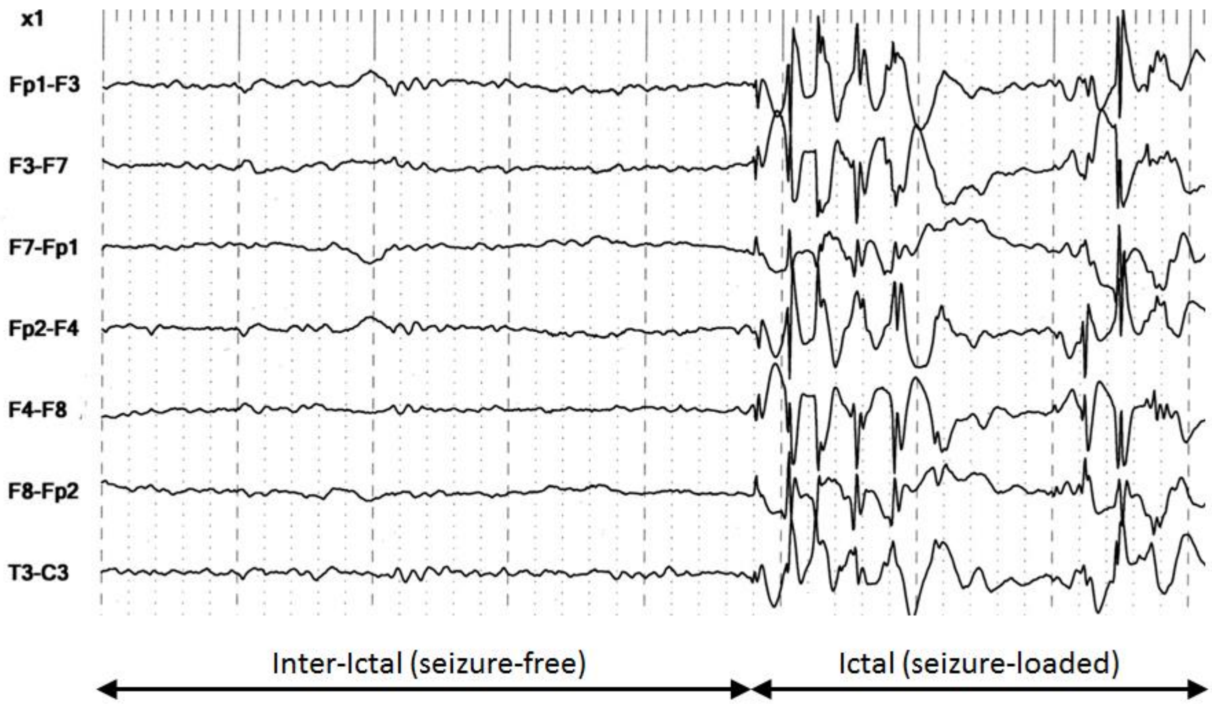

Figure 1. While diagnosing epileptics within right-sided seizure-loaded (so called ictal) stages is feasible by physicians at an accuracy between 70-95%, it is not feasible within inter-ictal (seizure-free) stages (left side). Data from Takahashi et al. (2015, p. 36).

Human based Diagnosis of Epilepsy in Ictal EEG Recordings

Human errors in the diagnosis of epilepsy are still widespread today. Due to different data interpretations in a seizure-loaded EEG recording by physicians, the accuracy of the diagnosis of epilepsy varies from a misdiagnosis rate of 5% in childhood epilepsy – in which the diagnosis was made by a panel of three experienced pediatric neurologists – to at least 23% in adult populations (van Donselaar et al. 2006). The human error rate in epilepsy diagnosis based on larger recordings is even higher – up to 30% (van Donselaar et al. 2006; Zaidi et al. 2000). These errors lead to misdiagnosis and incorrect treatments (Devinsky et al. 2018; van Donselaar et al. 2006). Some epilepsy diagnoses after a seizure, for example, can be explained by a disorder of the heart or metabolism, which is life-threatening for the patient (Zaidi et al. 2000). Further consequences of misdiagnoses can be of a socio-economic or psycho-social nature, e.g. driving bans or even loss of employment can occur. Often the patient is treated with the wrong high-dose medication, which does not contribute to recovery (van Donselaar et al. 2006).

Machine Learning based Diagnosis of Epilepsy in Ictal EEG Recordings

The massive progress in machine learning methods within the last few years has led to the fact that machines overcome weak human accuracy in terms of analyzing *ictal* (seizure-loaded) EEG recordings.

To develop and validate these ictal-EEG related methods, most scholars use the specific ictal-EEG parts from the same benchmarking dataset as we used in our inter-ictal (seizure free) EEG study here (Phys. Rev. E, 64, 061907).

One successful research stream analyzing ictal EEG data to accurately detect seizures follows the concept of using time-frequency representations to extract the relevant features for classification:

Martinez-del-Rincon et al. (2017) for instance, proposed a method that uses non-linear classifiers and a bag-of-words model for feature extraction. To extract the non-linear classifiers, Daubechies wavelet transformation was used to decompose the original signal into six frequency bands. Four statistical values (maximum, minimum, average and standard deviation) were generated for each of the six bands, resulting in a total of 24 features. For the bag-of-words model, the EEG signal serves as a text document where each segment of the signal is quantified by a set of words. The identified features were classified by a support vector machine. By this approach, Martinez-del-Rincon et al. (2017) achieved classification accuracies from 81.64% up to 99.12%.

Compared to Martinez-del-Rincon et al. (2017), Hassan et al. (2016a) achieved an accuracy of 98.4% with fewer features. They used another form of wavelet transformation, the tunable-Q factor wavelet transform, to decompose the incoming EEG signal. Six spectral moment-based features were extracted from the sub-bands of the wavelet transformation. In the next step, these features were classified through bootstrap aggregating by training several weak classifiers with a random subset of the features. Their decisions are then combined and assigned according to the principle of majority voting.

Using six spectral moment-based features selected from the complete ensemble empirical mode decomposition with adaptive noise method (CEEMDAN), Hassan et al. (2016b) achieved an accuracy of 97%. On the basis of higher order spectra features using support vector machines, Abdulhay et al. (2017) achieved an accuracy of 98.5% to classify ictal recordings.

Another research stream investigating ictal EEG data to find seizures focused on further statistical methods to classify epilepsy in ictal recordings:

For example, Tiwari et al. (2017) analyzed key points in the EEG signal. The signal was first smoothed with several Gaussian filters and then a pyramid of difference of Gaussian filters was formed. The maxima and minima of each level of this pyramid form the key points. At each key point a local binary pattern (LBP) was calculated. The decimal equivalent of the LBP creates a histogram which serves as a feature set. For classification, a support vector machine was used which classifies with an accuracy of 98.8%.

Only based on scaled raw data, Acharya et al. (2017) trained a 13-layer deep convolutional neural network and achieved an accuracy of 88.67% in detecting seizures in ictal EEG.

Beyond using the epilepsy benchmarking dataset (Phys. Rev. E, 64, 061907) for evaluation, van Diessen et al. (2013) as well as Li et al. (2017) proposed methods to detect seizures in ictal EEG based on other datasets: van Diessen et al. (2013) used functional EEG networks. For this purpose, a network was formed for each subject based on the statistical interdependencies. The synchronization likelihood was used for the recognition of linear and non-linear dependencies. It expresses the functional connectivity with a value between 0 and 1 and thus allows the creation of a weighted functional network. A Random Forest classifier was used which achieves an area under the receiver operating curve of 0.89.

Li et al. (2017) also applied a discrete wavelet transformation to decompose the EEG signal. Using the Hilbert transformation, an envelope analysis was performed to identify the periodic influences in the signals. Thus, six sub-bands were generated, from each of which five features were extracted. Artificial neural networks were used for training. Thus, Li et al. (2017) achieved an accuracy of 98.78%.

Despite the very good progress within the last years in applying modern machine learning methods to detect epileptic periods in ictal (seizure-loaded) EEG recordings with accuracy far above 80%, the key challenge of *diagnosing epileptics in inter-ictal (seizure-free) EEG records* still remains.

We are the first to build such a highly effective approach to diagnosing epileptics based on inter-ictal (seizure-free) EEG records.

Methodology

Our research methodology is threefold. We (1) make use of a very recent empirical finding from clinical neurophysiology and combine this finding with (2) machine learning based feature subset selection and (3) subsequent algorithm design:

1. We make use of the finding by Pittau et al. (2018, 2016) that epileptics have also very few (hidden) spikes in seizure-free (resting state) EEG recordings compared to healthy people.
2. Since the related spectral sub-bands of these hidden spikes are still unknown, we use a Random Forests based subset feature selection to identify the EEG resting state sub-bands differentiating healthy people from epileptics.
3. Based on the spectral power of the identified fine-graded sub-bands differentiating healthy people from epileptics we train and rigorously evaluated a Random Forests decision tree with an established high-quality epilepsy EEG dataset (Phys. Rev. E, 64, 061907) to quickly and accurately distinguish epileptics from healthy persons.

Dataset

The data used (Phys. Rev. E, 64, 061907) originate from Andrzejak et al's (2001) group from the Department of Epileptology, University of Bonn, Germany. The data consists of five sets (Z, O, N, F, S) each containing 100 single-channel EEG segments extracted from multichannel EEG recordings with a duration of 23.6 seconds. Sets Z and O consist of segments from healthy individuals on which EEG measurements were taken according to the international 10-20 scheme. Set Z was measured with eyes open, Set O with eyes closed. Sets N, F and S come from people suffering from epilepsy. Set N from the epileptogenic zone and set F from the hippocampal formation of the opposite hemisphere of the brain. Set S was derived from recordings made during a seizure. All EEG data were measured with the same 128-channel amplifier system and stored at a sampling rate of 173.61 Hz. In this study only the data sets Z, O, N and F are considered, since they allow a distinction to be made between epileptics and healthy people without having to rely on a seizure condition. Therefore, 400 recordings were examined, with 200 belonging to epileptics and the other 200 belonging to the control group (healthy people).

Data preprocessing

Independent Component Analysis (ICA) was performed to preprocess the data, followed by spectral analysis. Subsequently, a division into 99 fine-graded frequency sub-bands with a step width of 0.5 Hz from 0.5 Hz to 50 Hz was performed. The data had already been visually inspected by Andrzejak et al. (2001) for artifacts caused by eye movements or muscle activity. In addition, the linear decomposition approach by Bell and Sejnowski (1996) was used to clean the data. Independent components within a data recording can be extracted by using their ICA. There are three requirements for performing an ICA: (i) the mixed medium is linear and propagation delays are negligible, (ii) the time course of the sources are independent from each other, and (iii) the number of sources is equal to the number of sensors. These three requirements are given for electroencephalographic data, because the data is linear (i), interfering signals such as heartbeat or blinking are not bound to the source of EEG activity and therefore independent (ii) and condition (iii) can be confirmed by numerical simulation (Makeig et al. 1996). The basic idea of ICA is based on the central limit theorem. This states that a sum of n independent random variables approaches a normal distribution. The logic of ICA works exactly the other way around: if a signal has no normal distribution, it must be an independent component, whereas a normally distributed signal is probably a mixture of several components. For the analysis of EEG data, the rows of the input matrix x correspond to the values of the electrodes, the rows of the output matrix $u = Wx$ correspond to the time curves of the ICA components, and the columns of the inverse matrix W-1 indicate the projection force of the respective electrode component. The topography of the signals on the scalp contains information about their position. For example, blinking can be found in signals from the anterior part of the head. The corrected EEG data equals x' = (W)-1 u', where u' is the matrix of the excitation waves, and the series of noise signals is set to zero (Comon 1994).

In order to be able to further process the EEG signal cleaned by the ICA, it is converted into a frequency signal using spectral analysis. The EEG signal is represented by a function of frequencies using the Fourier transform (Fig. 2). First, the signal is separated into many sinusoidal oscillations with different amplitudes and frequencies. Then each wavelength can be checked for conformity to the EEG signal using correlation analysis. The result of the Fourier transformation is a power spectrum that allows an estimation to be made of the distribution of the frequencies of the EEG signal (Dumermuth and Molinari 1987).

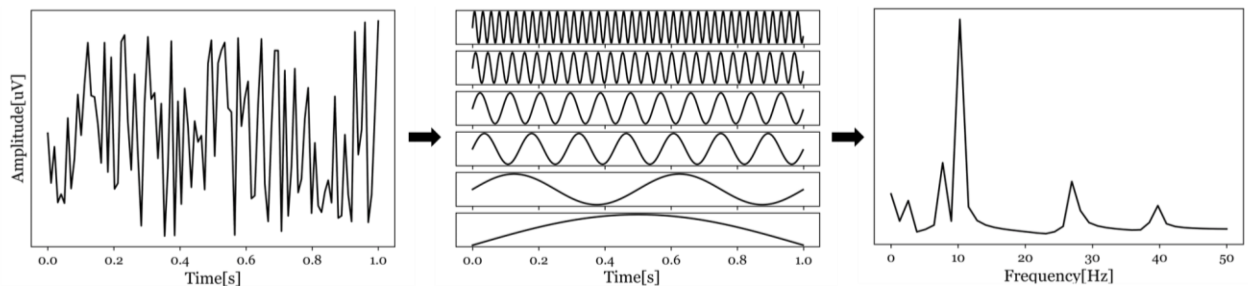

Figure 2. Schematic diagram of Fourier transformation, left: signal (time domain), middle: frequency components (time domain), right: power spectrum (frequency domain).

Feature Subset Selection

The traditional division of EEG frequency bands into alpha, beta, theta, delta and gamma bands was deliberately not used for feature extraction. Instead, based on the power per frequency obtained by the spectral analysis of the inter-ictal EEG, a division into 99 fine-graded frequency sub-bands with a step width of 0.5 Hz from 0.5 Hz to 50 Hz was performed. The hypothesis behind this division is that the information content of finer frequency bands is greater and that in order to obtain the best possible classification result, the information density must be as high as possible. This assumption was already supported in diagnosing schizophrenia (Buettner et al. 2019a, 2020a), sleep disorder (Buettner et al. 2019b, 2020b) and alcoholism (Rieg et al. 2019) using the same method (Buettner et al. 2019c).

Using the Random Forests based variable importance function, which gives a score for the importance of each feature taking into account the mean squared error (Genuer et al. 2010), the most influential variables can be determined. A feature subset was then created by selecting only those bands that had an importance on the scaled variable importance above 45.

Random Forest Training and Evaluation Method

Random Forest algorithm

The Random Forest algorithm was proposed by Breiman (2001). It is a machine learning classifier which is based on an ensemble (a bag) of unpruned decision trees. A Random Forest contains a collection of tree predictors, each tree being based on independently selected vectors. The classification outputs of the individual trees are used to determine the overall classification (Breiman 2001). Ensemble methods are related to the concept that an aggregated decision from various experts is often superior to a decision from a single system. The final classification decision is built on the majority vote.

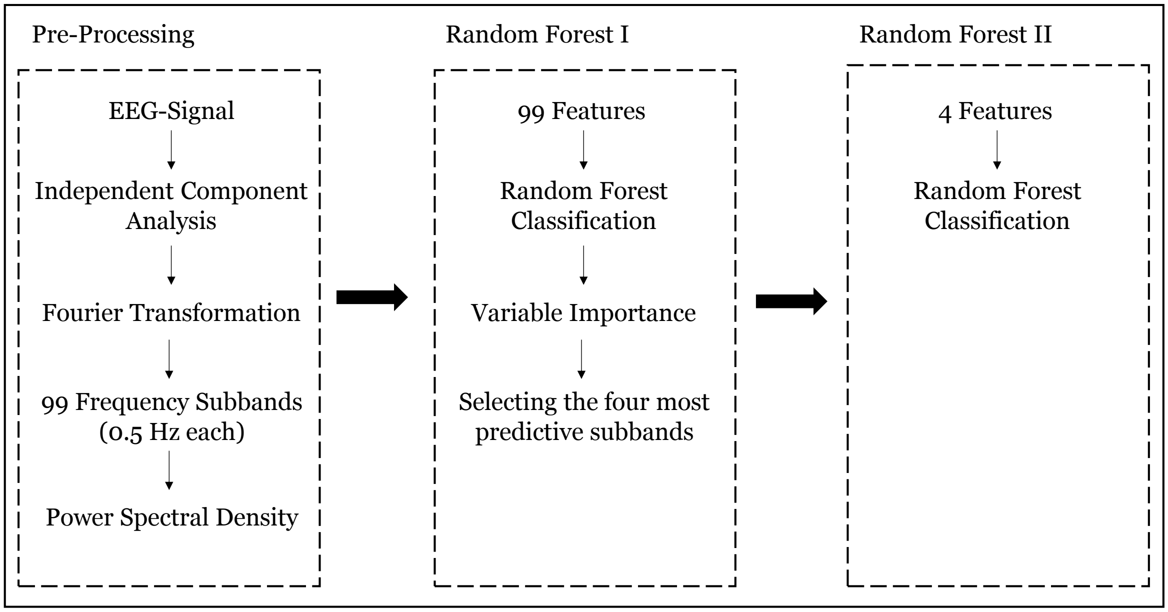

Figure 3. Methodology overview

Random Forests were successfully applied to solve complex decision problems in information systems (Buettner 2017a), human computer interaction (Buettner et al. 2018; Buettner 2018), and economics (Buettner 2016).

The Random Forest approach offers a number of advantages in dealing with data. Large amounts of data can be processed efficiently with Random Forest. In terms of classification accuracy, it is currently one of the best machine learning algorithms. Random Forest indicates which variables are relevant for its decision and it is able to estimate missing data effectively, which is why the classification accuracy remains constant even if large parts of the data are missing. If there is an uneven distribution of the data with respect to its classes, the Random Forest algorithm can compensate for such differences. The Random Forest process is described by Liaw and Wiener (2002) as follows: (i) Draw ntree bootstrap samples for the original data, (ii) For each of the bootstrap samples, an unpruned classification or regression tree grows, with the following modification: at each node, rather than choosing the best split among all predictors, randomly sample mtry of the predictors and choose the best split from among those variables. (Bagging can be thought of as a special case of Random Forest obtained when mtry = p, the number of predictors), (iii) Predict new data by aggregating the predictions of the ntree trees (i.e., majority votes for classification, average for regression).

The Random Forest based classifier was trained using the most predictive fine-graded frequency inter-ictal EEG sub-bands, identified by the feature subset selection procedure described before.

Evaluation Method

To evaluate the model, the dataset (n=400) was randomly split into training data (n_{train}=300) and unseen test data (n_{test}=100). Thus, 75% of the data were used for training the algorithm, the other unseen 25% for testing. In order to evaluate the trained model, the test set was used and the performance of the model was measured. Furthermore, the random forest based classifier was trained with k-fold cross validation. This method randomly divides the existing training data set into k equal parts. The training is executed on k-1 data, the validation is performed on the excluded data, which is used as a test data set. This process is repeated k times so that each part is used once as a test dataset. For this work, a k of 10 was selected. By cross-validation a possible instability of the model can be avoided (Kohavi 1995).

The epilepsy diagnosis algorithm will be rigorously exalted using the most established performance indicators of balanced accuracy, sensitivity (true positive rate), specificity (true negative rate), positive predictive value, negative predictive value, and Cohen's Kappa score.

Results

Identified Inter-Ictal Feature Subset

The variable importance analysis revealed that the frequency sub-bands at 1.5-2 Hz, 11-11.5 Hz, 11.5-12 Hz and 12-12.5 Hz were extremely relevant for the classification (Fig. 4). Therefore, these four inter-ictal EEG sub-bands were used to train the Random Forest based classifier distinguishing epileptics from healthy people.

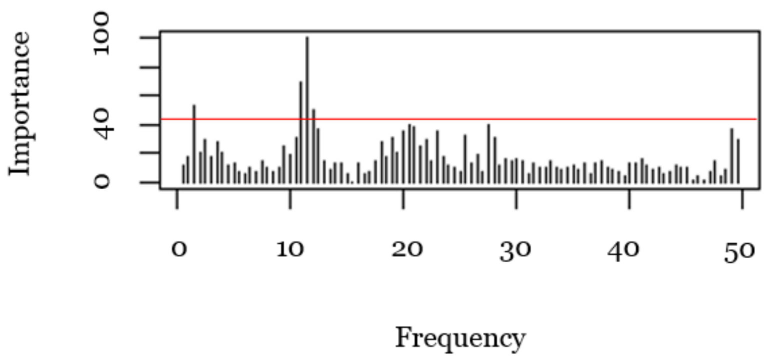

Figure 4. Variable importance of inter-ictal frequency sub-bands.

Performance of the Random Forests Classifier

The Random Forest based classifier which was trained with the four inter-ictal EEG sub-bands (1.5-2 Hz, 11-11.5 Hz, 11.5-12 Hz and 12-12.5 Hz) at ntree = 500 and mtry = 2 achieves an outstanding balanced accuracy of 99% (Table 1). Also all other performance indicators achieve excellent values (Table 1).

Table 1. Performance of our method. Values based on unseen test data (n_{test}=100).

Performance Indicator	Value
Balanced accuracy	99.0 %
Sensitivity (true positive rate)	100 %
Specificity (true negative rate)	98.0 %
Positive predictive value	98.0 %
Negative predictive value	100 %
Kappa	98.0 %

The new model was able to assign all healthy EEG recordings correctly to the non-epileptic group and achieved a sensitivity of 100%. Forty-nine of the 50 epileptic EEG recordings were also correctly classified as "epileptics". Only one error was made in the model in which an epileptic recording was misclassified as healthy. Table 2 summarizes the outstanding classification results.

Table 2. Confusion matrix of our method. Values based on unseen test data (n_{test}=100).

		Reference	
		Non-Epileptic	Epileptic
Predicted	Non-Epileptic	50	1
	Epileptic	0	49

Discussion

As demonstrated in Table 1 our epilepsy diagnosis algorithm performs very well and achieves a very good classification performance. In addition, with an accuracy of 99 percent and a Kappa of 98 percent using unseen data from a well-established epilepsy dataset, our algorithm significantly outperforms the current human benchmark of 70-95 percent accuracy. Put simply, our algorithm correctly classified every epileptic recording, and every healthy recording – with only a single exception. As shown in Table 2 one epileptic recording was misclassified as a healthy.

Post-hoc analysis of spectral power between epileptics and healthy persons in the four identified inter-ictal EEG sub-bands revealed strong differences (Table 3). While epileptics have more spectral power in the Mid-Delta sub-band 1.5-2 Hz (1.358 vs. 0.828, Cohen's d = 2.155, $p<0.01$), they have much less spectral power in the High-Alpha sub-band 11-12.5 Hz (0.270 vs. 0.783, Cohen's d = 2.324, $p<0.01$).

Table 3. Spectral power and statistical characteristics of the 4 bands, healthy vs. epileptic.

Characteristic	1.5-2 Hz	11-11.5 Hz	11.5-12 Hz	12-12.5 Hz
Mean healthy	0.828	0.948	0.775	0.628
Mean epileptic	1.358	0.295	0.274	0.241
SD healthy	0.221	0.593	0.251	0.187
SD epileptic	0.269	0.076	0.083	0.066
Cohen's d	-2.155	1.545	2.677	2.751
p-value	<0.01	<0.01	<0.01	<0.01

A representation of the spectral power differences between epileptics and controls is shown in Fig. 5.

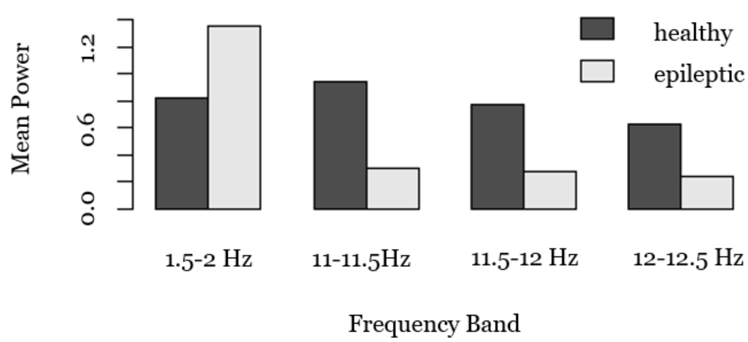

Figure 5. Mean spectral power of the four inter-ictal EEG sub-bands, healthy vs. epileptic.

Our findings of increased mid-delta power and reduced alpha power in epileptics confirm the results of the few references empirically dealing with inter-ictal (seizure-free) EEG recordings of epileptics:

The delta area normally attracts attention mainly in the area of anesthesia, coma or sleep, i.e. the responsible area with a reduced state of consciousness. For years, high spectral power in the delta range was almost exclusively observed in these fields and received little attention in epilepsy research (Lundstrom et al. 2019). However, current research suggests that the power of delta activity is mainly influenced by synaptic thalamocortical input (Lundstrom et al. 2019). Fisher and Velasco (2014) found that electrical brain stimulation, which also stimulated the thalamus, reduces the frequency and intensity of seizures in human volunteers. In addition, recent experiments on rats, in which low-frequency stimulation neutralizes the high spectral power in the delta region, were also able to induce an anti-epileptic reaction (Cheng et al. 2015).

Larsson and Kostov (2005) also found reduced alpha power in epileptics through the analysis of EEG recordings free of seizures and also Pyrzowski et al. (2015) reported a reduced alpha band. Reduced alpha power reflects various changes in thalamo-cortical and cortical network communication (Howells et al. 2018). Pyrzowski et al. (2015) argue that the strong reduction of alpha power is a possible biomarker to improve the diagnosis of epilepsy.

While confirming increased mid-delta power and reduced alpha power in epileptics in inter-ictal EEG recordings, our work may stimulate future empirical research on analyzing the high-alpha band 11-12.5 Hz and the mid-delta sub-band 1.5-2 Hz in more detail.

Conclusion

We built a highly effective epilepsy diagnosis algorithm that achieved very good performance based only on the analysis of seizure-free (resting state) EEG recordings. The algorithm achieved an accuracy of 99% which significantly outperforms the current benchmark of 70% to 95% by up to three experienced neurologists (van Donselaar et al. 2006). This novel algorithm is accurate, robust, fast, and cost-efficient, which substantially contributes to Information Systems research in healthcare (see Romanow et al. 2012).

We extend the work by Pittau et al. (2018, 2016) in terms of identifying the two EEG resting state sub-bands differentiating healthy from epileptic people (1.5-2 Hz and 11-12.5 Hz). Our empirical finding of reduced Alpha power in epileptics confirms the results of the few references empirically dealing with inter-ictal (seizure-free) EEG recordings of epileptics (Pyrzowski et al. 2015; Larsson and Kostov 2005). These results may stimulate future empirical research on analyzing the High-Alpha band 11-12.5 Hz and the Mid-Delta sub-band 1.5-2 Hz in more detail.

The novel algorithm diagnosing epileptics based on seizure-free EEG recordings is also of high practical relevance since epilepsy is one of the most common and serious brain diseases worldwide, causing a high level of individual suffering, discrimination, and social stigmatization (Moshé et al. 2015; Quintas et al. 2012), while the high human error rate in epilepsy diagnosis based on EEG recordings (up to 30%) has so far led to incorrect treatments with heavy consequences for patients (Devinsky et al. 2018; Duncan et al. 2006; van Donselaar et al. 2006; Zaidi et al. 2000). The epilepsy diagnosis algorithm could be implemented in modern EEG analysis devices, especially in intensive care units where an early diagnosis and appropriate treatment are decisive for life and death scenarios and where physician error rate is particularly high (van Donselaar et al. 2006; Zaidi et al. 2000; Donchin et al. 1995; Kirwan 1994). Thus our machine learning algorithm could be used in clinical everyday life in addition to the current EEG evaluation on the basis of the classical five frequency bands. In stress- and error-prone situations an automated control by the algorithm can be very useful, which prevents errors and thus could save lives. Due to the focus on small frequency ranges only a fraction of the data of the classical EEG is generated, so that a continuous control is possible. In addition, the effectiveness of drugs or brain stimulation can be measured and possibly individually adapted to the patient.

The main limitation of our work is rooted in limited external validity. That is why in future work we will re-evaluate the algorithm in a laboratory environment testing new patients (epileptics and healthy people randomly assigned, using a double blind procedure). We will also apply the identified EEG resting state sub-bands differentiating healthy from epileptic people (1.5-2 Hz and 11-12.5 Hz) in a modified deep learning approach using convoluational neural networks (Esteva et al. 2019; Buettner and Baumgartl 2019; Baumgartl and Buettner 2020). In addition, we will triangulate EEG sensor data with other physiological sensor data (Buettner 2017b).

Acknowledgements

We thank the reviewers and editors, who provided very helpful feedback on the refinement of the manuscript. This research is partly funded by the German Federal Ministry of Education and Research (no. 13FH4E03IA, no. 13FH176PX8) and by the German Federal Ministry of Economics and Technology (EXIST program, no. 03EGSBW477).

References

Abdulhay, E., Elamaran, V., Chandrasekar, M., Balaji, V. S., and Narasimhan, K. 2017. "Automated diagnosis of epilepsy from EEG signals using ensemble learning approach," *Pattern Recognition Letters*, in press.

Abouzahra, M., Guenter, D., and Tan, J. 2015. "Integrating Information Systems and Healthcare Research to Understand Physicians' Use of Health Information Systems: A Literature Review," in *Proceedings of the* 36th *International Conference on Information Systems*, T. Carte, A. Heinzl, C. Urquhart (eds.), Fort Worth, TX, pp. 1-19.

Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., and Hojjat, A. 2015. "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," *Computers in Biology and Medicine* (100), pp. 270-278.

Andrzejak, R. G., Lehnertz, K., Mormann, F., Rieke, C., David, P., and Elger, C. E. 2001. "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* (64:6), 061907.

Aoe, J., Fukuma, R., Yanagisawa, T., Harada, T., Tanaka, M., and Kobayashi, M. 2019. "Automatic diagnosis of neurological diseases using MEG signals with a deep neural network," *Scientific Reports* (9), 5057.

Baumgartl, H., and Buettner, R. 2020. "Development of a highly precise place recognition module for effective human-robot interactions in changing lighting and viewpoint conditions," in *Proceedings of the 53rd Hawaii International Conference on System Sciences*, T. Bui (ed.), Maui, HI, in press.

Bell, A., and Sejnowski, T. J. 1996. "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation* (7:6), pp. 1129-1159.

Breiman, L. 2001. "Random Forests," *Machine Learning* (45:1), pp. 5-32.

Buettner, R. 2016. "Innovative Personality-based Digital Services," in *Proceedings of the 20th Pacific Asia Conference on Information Systems*, T.-P. Liang, S.-Y. Hung, P. Y. K. Chau, S. Chang (eds.), Chiayi, Taiwan, 1-13.

Buettner, R. 2017a. "Predicting user behavior in electronic markets based on personality-mining in large online social networks: A personality-based product recommender framework," *Electronic Markets: The International Journal on Networked Business* (27:3), pp. 247-265.

Buettner, R. 2017b. "Asking both the User's Brain and its Owner using Subjective and Objective Psychophysiological NeuroIS Instruments," in *Proceedings of the 38th International Conference on Information Systems*, Y. J. Kim, R. Agarwal, J. K. Lee (eds.), Seoul, South Korea, pp. 1-11.

Buettner, R., Sauer, S., Maier, C., and Eckhardt, A. 2018. "Real-time Prediction of User Performance based on Pupillary Assessment via Eye-Tracking," *AIS Transactions on Human-Computer Interaction* (10:1), pp. 26-56.

Buettner, R. 2018. "Robust user identification based on facial action units unaffected by users' emotions," in *Proceedings of the 51st Hawaii International Conference on System Sciences*, T. Bui (ed.), Big Island, HI, pp. 265-273.

Buettner, R., and Baumgartl, H. 2019. "A highly effective deep learning based escape route recognition module for autonomous robots in crisis and emergency situations," in *Proceedings of the 52nd Hawaii International Conference on System Sciences*, T. Bui (ed.), Maui, HI, pp. 659-666.

Buettner, R., Hirschmiller, M., Schlosser, K., Rössle, M., Fernandes, M., and Timm, I. J. 2019a. "High-performance exclusion of schizophrenia using a novel machine learning method on EEG data," in *Proceedings of the IEEE Healthcom 2019*, C. Lozano-Garzon, J. Rodrigues (ed.), Bogotá, Colombia, in press.

Buettner, R., Fuhrmann, J., and Kolb, L. 2019b. "Towards high-performance differentiation between Narcolepsy and Idiopathic Hypersomnia in 10 minute EEG recordings using a Novel Machine Learning Approach," in *Proceedings of the IEEE Healthcom 2019*, C. Lozano-Garzon, J. Rodrigues (ed.), Bogotá, Colombia, in press.

Buettner, R., Beil, D., Scholz, S., and Djemai, A. 2020a. "Development of a machine learning based algorithm to accurately detect schizophrenia based on one-minute EEG recordings," in *Proceedings of the 53rd Hawaii International Conference on System Sciences*, T. Bui (ed.), Maui, HI, in press.

Buettner, R., Grimmeisen, A., and Gotschlich, A. 2020b. "High-performance Diagnosis of Sleep Disorders: A Novel, Accurate and Fast Machine Learning Approach Using Electroencephalographic Data," in *Proceedings of the 53rd Hawaii International Conference on System Sciences*, T. Bui (ed.), Maui, HI, in press.

Buettner, R., Rieg, T., and Frick, J. 2019c. "Machine Learning based Diagnosis of Diseases Using the Unfolded EEG Spectra: Towards an Intelligent Software Sensor," in Information Systems and Neuroscience, LNISO (32), F. D. Davis, R. Riedl, J. vom Brocke, P.-M. Léger and A. B. Randolph (eds.), New York: Springer, in press.

Cheng, H., Kuang, Y., Liu, Y., Wang, Y., Xu, Z., Geo, F., Zhang, S., Ding, M., Chen, Z. 2015. "Low-frequency stimulation of the external globus palladium produces anti-epileptogenic and anti-ictogenic actions in rats," *Acta Pharmacologica Sinica* (36), pp. 957-965.

Comon, P. 1994. "Independent component analysis, a new concept?," *Signal Processing* (*36*:3), pp. 287-314.

Devinsky, O., Vezzani, A., O'Brien, T., Jette, N., Scheffer, I., Curtis, M., and Perucca, P. 2018. "Epilepsy," *Nature Reviews Disease Primers* (4), 18024.

Donchin, Y., Gopher, D., Olin, M., Badihi, Y., Biesky, M., and Sprung, C. L. 2003. "A look into the nature and causes of human errors in the intensive care unit," *Critical Care Medicine* (12:2), pp. 294-300.

Dumermuth, G., & Molinari, L., 1987. "Spectral analysis of the EEG," *Neuropsychobiology* (*17*:1-2), pp. 85-99.

Duncan, J., Sander, J., Sisodiya, S., Walker, and Matthew C. 2006. "Adult epilepsy," *The Lancet* (367:9516), pp. 1087-1100.

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., and Dean, J. 2019. "A guide to deep learning in healthcare," *Nature Medicine* (25), pp. 24-29.

Fisher, R., van Emde Boas, W., Blume, W., Elger, C., Genton, P., Lee, P., and Engel, J. 2005. "Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilesy (ILAE) and the International Bureau for Epilepsy (IBE)," *Epilepsia* (46:4), pp. 470-472.

Fisher, R., Velasco, A. 2014. "Electrical brain stimulation for epilepsy," *Nature Reviews Neurology* (10), pp. 261-270.

Genuer, R., Poggi, J. M., and Tuleau-Malot, C. 2010. "Variable selection using random forests," *Pattern Recognition Letters* (31:14), pp. 2225-2236.

Hassan, A., Siuly, S., and Zhang, Y. 2016a. "Epileptic seizure detection using tunable-Q factor wavelet transform and bootstrap aggregating," *Computer Programs and Methods in Biomedicine* (137), pp-247-259.

Hassan, A., and Subasi, A. 2016b. "Automatic identification of epileptic seizures from EEG signals using linear programming boosting," *Computer Methods and Programs in Biomedicine* (136), pp. 65-77.

Howells, F., Temmingh, H., Hsieh, J., van Dijen, A., Baldwin, D., Stein, D. 2018. "Electroencephalographic delta/alpha frequency activity differentiates psychotic disorders: a study of schizophrenia, bipolar disorder and methamphetamine-induced psychotic disorder," *Translational Psychiatry* (8:75), pp. 1-11.

Kale, R. 2002. "The Treatment Gap," *Epilepsia* (43:Suppl. 6), pp. 31–33.

Kirwan, B. 1994. *A Guide To Practical Human Reliability Assessment*, Boca Raton, FL: CRC Press.

Kruse, C. S., and Beane, A. 2018. "Health Information Technology Continues to Show Positive Effect on Medical Outcomes: Systematic Review," *Journal of Medical Internet Research* (20:2), e41.

Kohavi, R. 1995. "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, Montreal, Quebec, Canada, pp. 1137-1145.

Larsson, P. G., and Kostov, H. 2005. "Lower frequency variability in the alpha activity in EEG among patients with epilepsy," *Clinical Neurophysiology* (116:11), pp. 2701–2706.

Li, M., Chen, W., and Zhang, T. 2016. "Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble," *Biomedical Signal Processing and Control* (31), pp. 357-365.

Liaw, A., and Wiener, M. 2002. "Classification and Regression by randomForest," *R News* (2:3), pp. 18-22.

Lundstrom, B., Boly, M., Duckrow, R., Zavari, H., Blumenfeld, H. 2019. "Slowing less than 1 Hz is decreased near the seizure onset zone," *Scientific Reports* (9:6218), pp. 1-10.

Makeig, S., Bell, A. J., Jung, T. P., and Sejnowksi, T. J. 1996. "Independent Component Analysis of Electroencephalographic Data," *Advances in Neural Information Processing Systems* (8), *pp*. 145-151.

Martinez-del-Rincon, J., Santofimia, M. J., del Toro, X., Barba, J., Romero, F., Navas, P., and Lopez, J. C. 2017. "Non-linear classifiers applied to EEG analysis for epilepsy seizure detection," *Expert Systems with Applications* (*86)*, pp. 99-112.

Moshé, S. L., Perucca, E., Ryylin, P., and Tomson, T. 2015. "Epilepsy: New Advances," *The Lancet* (385:9971), pp. 884-898.

Öksüz, N., Kowatsch, T., Shcherbatyi, I., and Maass, W. 2018. "A Data-analytical System to Predict Therapy Success for Obese Children," in *Proceedings of the 39th International Conference on Information Systems*, J. Pries-Heje, S. Ram, and M. Rosemann (eds.), San Francisco, CA, pp. 1-17.

Pohlmann-Eden, B., and Newton, M. 2008. "First seizure: EEG and neuroimaging following an epileptic seizure," *Epilepsia* (49:1), pp. 19–25.

Pittau, F., Genetti, M., Birot, G., Tomescu, M., Baldini, S., Vulliémoz, S., Michel, C., and Seeck, M. 2016. "Detection of epileptic activity in absence of EEG interictal epileptic discharges," *Clinical Neurophysiology* (127:3), pp. e44–e45.

Pittau, F., Baldini, S., Tomescu, M. I., Michel, C. M., and Seeck, M. 2018. "EEG resting states are altered in focal epilepsy," *Clinical Neurophysiology* (129:S1), pp. e58.

Pyrzowski, J., Siemiński, M., Sarnowska, A., Jedrzejczak, J., and Nyka, W. M. 2015. "Interval analysis of interictal EEG: pathology of the alpha rhythm in focal epilepsy," *Scientific Reports* (5), 16230.

Romanow, D., Cho, S. and Straub D. 2012. "Riding the Wave: Past Trends and Future Directions for Health IT Research," *MIS Quarterly* (36:3), pp. iii-x.

Quintas, R., Raggi, A., Giovannetti, A., Pagani, M., Sabariego, C., Cieza, A., and Leonardi, M. 2012. "Psychosocial difficulties in people with epilepsy: a systematic review of literature from 2005 until 2010," *Epilepsy & Behavior* (25:1), pp. 60-67.

Rieg, T., Frick, J., and Buettner, R. 2019. "High-performance detection of alcoholism by unfolding the amalgamated EEG spectra using the Random Forests method," in *Proceedings of the 52nd Hawaii International Conference on System Sciences*, T. Bui (ed.), Maui, HI, pp. 3769-3777.

Schreiner, A., and Pohlmann-Eden, B. 2003. "Value of the Early Electroencephalogram after a First Unprovoked Seizure," *Clinical EEG and Neuroscience* (34:3), pp. 140-144.

Takahashi, S., Yamamoto, S., Tanaka, R., Okayama, A., Araki, A., and Azuma, H. 2015. "Focal frontal epileptiform discharges in a patient with eyelid myoclonia and absence seizures," *Epilepsy & Behavior Case Reports* (4), 35–37.

Tiwari, A. K., Pachori, R. B., Kanhangad, V., and Panigrahi, B. K. 2017. „Automated diagnosis of epilepsy using key-point-based local binary pattern of EEG signals," *IEEE Journal of Biomedical and Health Informatics* (21:4), pp. 888-896.

Tjepkema-Cloostermans, M. C., de Carvalho, R. C. V., and van Putten, M. J. A. M. 2018. „ Deep learning for detection of focal epileptiform discharges from scalp EEG recordings," *Clinical Neurophysiology* (129:10), pp. 2191-2196.

van Diessen, E., Otte, W. M., Braun, K. P., Stam, C. J., and Jansen, F. E. 2013. „Improved diagnosis in children with partial epilepsy using a multivariable prediction model based on EEG network characteristics," *PLOS ONE* (8:4), e59764.

van Donselaar, C. A., Stroink, H., and Arts, W.-F. 2006. "How confident are we of the diagnosis of epilepsy?," *Epilepsia* (47:1), pp. 9-13.

Wang, Y., Wang, L., Rastegar-Mojarad, M., Moon, S., Shen, F., and Afzal, N. 2018a. "Clinical information extraction applications: A literature review," *Journal of Biomedical Informatics* (77), pp. 34-49.

Wang, D., Ren, D., Li, K., Feng, Y., Ma, D., Yan, X., and Wang, G. 2018b. "Epileptic Seizure Detection in Long-Term EEG Recordings by Using Wavelet-Based Directed Transfer Function," *IEEE Transactions on Biomedical Engineering* (65:11), pp. 2591-2599.

Zaidi, A., Clough, P., Cooper, P., Scheepers, B., and Fitzpatrick, A. P. 2000. "Misdiagnosis of epilepsy: many seizure-like attacks have a cardiovascular cause," *Journal of the American College of Cardiology* (36:1), pp. 181-184.