Development of a machine learning based algorithm to accurately detect schizophrenia based on one-minute EEG recordings

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

David Beil
Aalen University
david.beil@studmail.htw-aalen.de

Stefanie Scholtz
Aalen University
stefanie.scholtz@studmail.htw-aalen.de

Aadel Djemai
Aalen University
aadel.djemai@studmail.htw-aalen.de

Abstract

While diagnosing schizophrenia by physicians based on patients' history and their overall mental health is inaccurate, we report on promising results using a novel, fast and reliable machine learning approach based on electroencephalography (EEG) recordings. We show that a fine granular division of EEG spectra in combination with the Random Forest classifier allows a distinction to be made between paranoid schizophrenic (ICD-10 F20.0) and non-schizophrenic persons with a very good balanced accuracy of 96.77 percent. We evaluate our approach on EEG data from an open neurological and psychiatric repository containing 499 one-minute recordings of n=28 participants (14 paranoid schizophrenic and 14 healthy controls). Since the fact that neither diagnostic tests nor biomarkers are available yet to diagnose paranoid schizophrenia, our approach paves the way to a quick and reliable diagnosis with a high accuracy. Furthermore, interesting insights about the most predictive subbands were gained by analyzing the electroencephalographic spectrum up to 100 Hz.

Keywords:
Schizophrenia, Electroencephalography, Random Forest, Spectral analysis, Machine Learning

1. Introduction

Mental disorders represent a global growing issue that affects approximately one in two people in their lifetime [1]. Schizophrenia is considered to be one of the most frequent neuropsychiatric disorders with a lifetime risk of one percent in the general population [2, 3]. It can cause a significant burden on both the individuals affected and society [4]. Additionally, it is associated with substantial premature mortality [5, 6] and morbidity [7]. Affected persons suffer heavily under the wide range of symptoms including positive ones (delusions, hallucinations and disorganized speech), which can be treated with antipsychotic drugs effectively [8, 9], and negative ones (social withdrawal, self-neglect, loss of emotional responsiveness and motivation together with mild cognitive impairment) [8-11]. Unemployment levels of schizophrenic persons are extremely high at 80-90 percent and life expectancy is reduced by 10-20 years [4, 12-14]. Schizophrenia can lead people to harm themselves or even commit suicide. Drug abuse and homelessness are known as possible side effects of psychotic disorders [15]. However, little is in fact known about the origin of this particular disease. Schizophrenia is considered a disorder of abnormal brain connectivity [16] caused by genetic or environmental factors, or both [4]. Schizophrenic disorders are categorized by nine different types: paranoid schizophrenia (ICD-10 F20.0), hebephrenic schizophrenia (F20.1), catatonic schizophrenia (F20.2), undifferentiated schizophrenia (F20.3), post-schizophrenic depression (F20.4), Residual schizophrenia (F20.5), simple schizophrenia (F20.6), other schizophrenia (F20.8) and schizophrenia, unspecified (F20.9). Since paranoid schizophrenia is the most common type of schizophrenia in most parts of the world, this work focuses on this type [17].

The major problem in detecting schizophrenia is that so far neither diagnostic tests nor biomarkers are available. Clinical diagnosis is instead made based on a patient's history and their overall mental condition [4].

IT-based healthcare has undergone a dramatic upswing in the past years, largely driven by increases in computational power and the availability of huge new datasets [18, 19]. The field has witnessed spectacular advances in the ability of machines to understand data and this could thus be accompanied by great successes in medicine, in particular for diagnosing diseases [20, 21] or IT-generated recipes [22]. The application of most modern machine learning using big data within the healthcare domain fosters this success [18-22].

Despite this immense progress, only minimal insights have been obtained about schizophrenia as one

HICSS

of the most frequent neuropsychiatric disorders [1-3]. For instance, animal models show that developmental hippocampal lesions may cause abnormal connectivity of the prefrontal cortex [16] and in the human brain irregular recordings (glutamate-mediated neurotransmission) have been detected in neuropsychiatric disorders such as schizophrenia [23]. In general, a lot of studies have shown that schizophrenia patients often show unspecific abnormalities in their EEG recordings [1].

The most prominent theory explaining the schizophrenic disorder is the disturbed functional connectivity theory between small-world brain networks [24-26]. According to this theory, disturbances in functional connectivity are the major pathophysiological mechanism for schizophrenia, and, in particular, for cognitive disorganization. It was found that these disturbances to functional connectivity are reflected in atypical and unspecific EEG recordings of schizophrenic patients, but EEG analysis based on the standard large EEG bandwidths are insufficient to make use of functional connectivity theory to diagnose schizophrenia [24].

However, from a modern machine learning point of view, we make use of this theory and propose a machine learning based classification system analyzing fine-graded EEG spectra to diagnose schizophrenic patients. Such a fast, automatic classification system with a high accuracy for detecting schizophrenia is highly useful in daily clinical practice. That is why in this paper we aim to evaluate the possibility of reliably distinguishing people with schizophrenia from non-schizophrenics based on fine-graded EEG spectra.

Research question: Can we build an artifact to distinguish schizophrenic from non-schizophrenic persons based on fine-graded EEG spectra?

In order to identify schizophrenic persons from non-schizophrenics, we made use of the Random Forest classifier. The classifier uses aperiodic time series data that quote the registration of electrical activity of brain waves at the skull surface.

The most important contributions are:

1) We build a highly effective classifier to distinguish schizophrenic from non-schizophrenic persons based on EEG data with a very good balanced accuracy of 96.77 percent.
2) By using the Random Forest machine learning method, a fast classification using only one-minute of EEG recording is possible.
3) The extension of the analysis spectrum from up to 50 Hz to 100 Hz leads to a further improvement in diagnostic accuracy (96.01 % versus 96.77 %).
4) In addition to the current state of knowledge, the upper gamma frequency band (96.5 Hz to 99.5 Hz) is also very relevant for differentiating schizophrenic from healthy people, something that may

stimulate theory-building work (e.g. by extending functional connectivity theory [24-26]).

The paper is organized as follows: Next we present the procedure of our research methodology by first arguing for the data set used. Furthermore, we set out the data preparation and the Random Forest method. After that, we demonstrate the machine learning results concerning the performance evaluation. Then, we discuss the results and include practical implications, before concluding with limitations and future work.

2. Methodology

In order to show methodological rigor in covering a relevant topic and thus finally contribute an applicable artifact for medicine and research, the work in this paper is done in accordance with the Information Systems design science approach [27]. The procedure covers the following steps (Fig. 1): Reading the dataset, applying Independent Component Analysis (ICA) on the data, calculating mean values over all sensors, performing a spectral analysis with 200 equally ranged frequency bands, and conducting the classification.
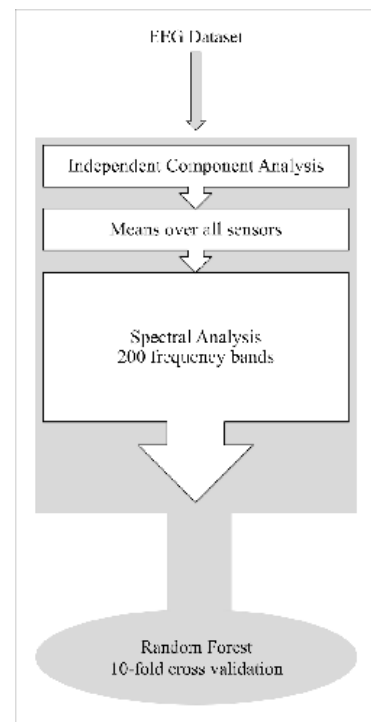

Figure 1: Method overview.

2.1. Dataset

The dataset comprised of 14 patients (7 males: 27.9 ± 3.3 years, 7 females: 28.3 ± 4.1 years) with ICD-10

F20.0 diagnosed paranoid schizophrenia, who were hospitalized at the Institute of Psychiatry and Neurology in Warsaw, Poland, and 14 healthy controls (7 males: 26.8 ± 2.9, 7 females: 28.7 ± 3.4 years) [28]. The patients had to undergo a medication washout of at least seven days. Study protocol was approved by the Ethics Committee of the Institute of Psychiatry and Neurology in Warsaw. All participants received a written description of the protocol and provided written consent to take part in this study. Inclusion criteria were a minimum age of 18, the ICD-10 diagnosis F20.0, as well as medication washout period of a minimum of seven days. Exclusion criteria were pregnancy, organic brain pathology, epilepsy, Alzheimer's, or Parkinson disease, presence of a general medical condition, or very early stage of schizophrenia [28]. The control was matched in gender and age to the 14 patients completing the study [28].

The dataset consists of data from 64 electrodes placed on the scalp, scanned at 250 Hz [28], using the internationally standardized 10-20 EEG montage [29]. This designation relates to the positioning of the electrodes on the scalp as it measures the distance from Nasion to Inion and defines it as 100 percent. The EEG data were recorded for approximately fifteen minutes in all subjects during an eyes-closed resting state condition with 19 EEG channels: Fp1, FP2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2. The full dataset is held in a public repository and is available at https://repod.pon.edu.pl/dataset/eeg-in-schizophrenia [30].

2.2 Preprocessing

This part of the paper covers the data preparation of the collected EEG data. The EEG mirrors many thousands of simultaneous brain processes in specific frequency bands and their associated frequency bandwidth [31]. During the EEG data collection, noise disturbances such as muscle activities, blinking, movements of the eyes and the heartbeat are recorded. The ICA algorithm, which has to be implemented to remove these artifacts, performs well if the following requirements are satisfied:
1) Mixing medium is linear and propagation delays are negligible.
2) The time courses of the sources are independent.
3) The number of sources is the same as the amount of sensors.

For applying ICA on EEG data, the first and second assumptions are clearly met. The recording is linear and instantaneous. Also, the sources of blinking, muscle activities and heartbeat are not generally time locked. Since we do not know the exact amount of statistically independent brain signals, the third

requirement is questionable. Nevertheless, simulations have confirmed that ICA works very well on EEG data [32]. In the case of EEG data analysis, the recorded signals must be the rows of the Input Matrix x. The Output Matrix u = Wx contains the time series of the ICA components. As a result of the ICA, we get the corrected EEG signals x' = W^{-1} u' [33]. The ICA algorithm used for pre-processing is provided with eegkit v.1.0.4 within R x64 3.5.3.

In addition, we conducted a dimensional reduction by calculating the mean values over all 19 EEG channels per recording. The preprocessed recordings of the 28 participants with a total recording time of 499 minutes are separated into one-minute chunks each. This step leads to a total of 499 recordings, which builds the foundation of our subsequent analysis and hence produces a more accurate result.

2.3 Machine Learning Method

The machine learning method comprises the following three steps:

2.3.1. Spectral Analysis and Feature Extractions. In order to transform the cleaned EEG data set from time series data into frequency range data, the EEG spectral analysis is conducted. An appropriate approach is the Fast Fourier Transformation (FFT), which decomposes the EEG signals as a function of frequencies. The Spectral Analysis is provided using eegkit v.1.0.4 within R x64 3.5.3. The simplified process of the FFT is shown in figure 2.
1) First the original signal is broken down into many sinusoidal oscillations.
2) Then the strength of each frequency, within the original signal, is calculated.
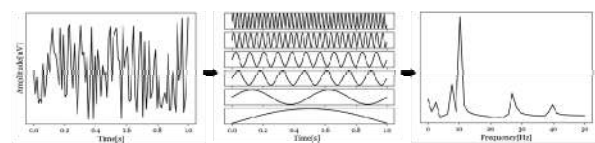

Figure 2: Fourier Transformation [34].

In this work, the commonly used division of frequency bands into alpha, beta, theta, delta and gamma bands, as shown in table 1, was not used as a feature extraction criterion [31]. According to the study of Rieg et al. [34], we decided to apply their method on EEG data of schizophrenic and non-schizophrenic persons. To take the upper gamma frequency bands into account, we additionally decided to extend the frequency range up to 100 Hz. Thus, 200 frequency bands from the same width are considered in this work.

Frequency Band	Frequency range in Hz	Characterization
Delta	0.5–3.5 Hz	Deep sleep
Theta	3.5–7.5 Hz	Sleep and dream.
Alpha	7.5–12.5 Hz	Relaxed awake, closed eyes
Beta	12.5–30 Hz	Inner restlessness, stress, concentration
Gamma	> 30 Hz	Extreme concentration

Table 1: Standard EEG bandwidths [34].

2.3.2. Classification. We used the Random Forest classifier. This classifier shows substantial improvements in classification accuracy that results from growing collections of trees and letting them vote for the most popular class. The process works as illustrated in figure 3, and is described as follows [34-37]:

1) n_{tree} bootstrap training samples are randomly produced from the original data.
2) Each training sample generates the corresponding decision tree. For each leaf node, the m_{try} of the predictors are randomly sampled and the best split among all variables is chosen.
3) Each tree expands unpruned.
4) The corresponding category is determined by using each test sample decision for testing.
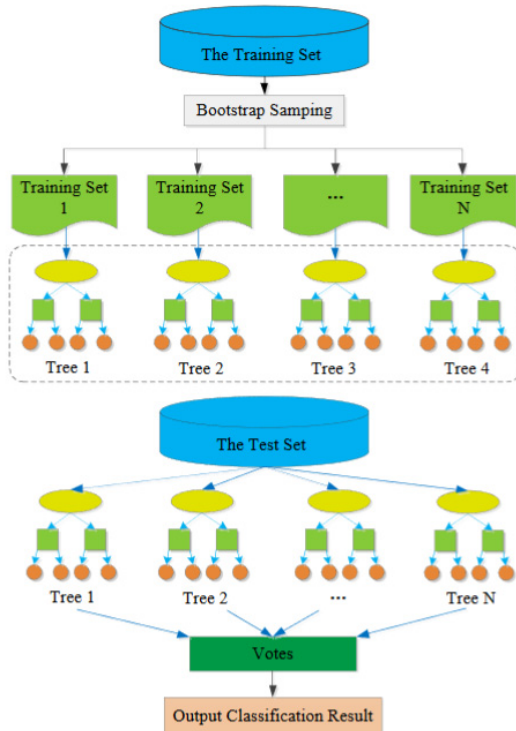5) According to majority voting, the class is picked.

Figure 3: Random Forest method [36].

To classify the dataset of schizophrenic and non-schizophrenic persons, the Random Forest of the caret package was chosen. The dataset was divided in a 75 % training set (375 recordings) and a 25 % test set (124 recordings). The amount of trees is $n_{tree} = 500$.

2.3.3. Validation. To complement and improve the Random Forest classifier we made use of k-fold cross validation with 10 iterations. This method is particularly suitable as it provides information about the ruggedness of a model. Within this operation, the training set is divided into 10 randomly chosen folds. Nine of them are used to train the model, whereby the remaining one is used to test it. The result of the 10-fold cross validation is the mean value of all trials. Correspondingly, we get a confusion matrix with the following fields:

- True positive: The subject suffers from schizophrenia and the model identified correctly.
- False negative: The subject suffers from schizophrenia, but the model has wrongly identified the subject as a non-schizophrenic person.
- False positive: The subject is not suffering from schizophrenia, but the model has wrongly detected the subject as a schizophrenic person.
- True negative: The subject is not suffering from schizophrenia and this has been detected by the model correctly.

The outcome is a valid model, which can be used for prediction on the actual test set.

3. Results

For Random Forest algorithm we applied the caret package v.6.0.82 within a R x64 3.5.3 environment.

To train and evaluate the model, we split the 499 recordings into a training partition (n_T = 375) and an evaluation partition (n_E = 124). The classifier was built using 500 voting trees. Data from 19 sensors in each recording were summarized. Subsequently, we first built 99 power bands with a range of 0.5 Hz each and gained a balanced accuracy of 96.01 percent. By repeating the process, expanding the spectrum up to 100 Hz, we improved our result to a balanced accuracy of 96.77 percent. The validity of our classifier is ensured by the 10-fold cross validation. Only two non-schizophrenic samples were misclassified as schizophrenic and two schizophrenic samples were misclassified as non-schizophrenic using the extended method (Fig. 4). The remaining 120 samples were classified correctly.

We evaluated our classifier in terms of accuracy, positive predictive value, negative predictive value and balanced accuracy. As shown in table 2 the classifier

achieved excellent performance. The results show that the trained classifier has a balanced accuracy of 96.77 percent. Thus, we are able to distinguish a schizophrenic from a non-schizophrenic person with a very high performance and balanced accuracy.

		Reference	
		Non-schizophrenic	schizophrenic
Predicted	Non-schizophrenic	60	2
	schizophrenic	2	60

Figure 4: Confusion Matrix.

Performance indicator	Value
Accuracy	96.77%
Positive predictive value	96.77%
Negative predictive value	96.77%
Balanced accuracy	96.77%
Kappa	0.9355

Table 2: Performance of classifier.

In addition we detected that, beyond the current state of our knowledge, specific frequency subbands in the high gamma range are highly relevant for the diagnosis of schizophrenia. We identified the four most important frequencies for prediction. The most predictive frequency band is 50-50.5 Hz, which is scaled to a value of 100 and serves as a basis for the calculation of the importance of the remaining bands. The following results are obtained for the respective important frequency bands: 96-96.5 Hz has a value of 65.35, 8.5-9 Hz has a value of 58,64 and 96.5-97 Hz has a value of 56.10. Overall 196 frequency bands are below an importance of 50. In comparison to the classical frequency divisions the most important subbands can be assigned to the following bands: three subbands are in the gamma range, and one subband is in the alpha range.

4. Discussion

Our artifact for distinguishing schizophrenic from non-schizophrenic persons performs very well as demonstrated in table 2. While previous studies using other datasets achieved accuracies between 80.5 and 90.48 percent (see table 3), our approach outperforms

this with a balanced accuracy of 96.77 percent in terms of classification [38-41]. Boostani et al. [38] used the AdaBoost and the Boosted version of Direct Linear Discriminant Analysis (BDLDA) method for their classification. Based on a dataset with 13 schizophrenic patients and 18 age-matched control participants they achieved an accuracy of 85.41 percent (AdaBoost) and 87.51 percent (BDLDA). Based on samples of 780 EEG recordings Zhang et al. [39] attained an accuracy of 90 percent by using a high order pattern discovery algorithm. Laton et al. [40] used Naïve Bayes, Decision Tree and Adaboost as classifiers. The dataset is based on 54 patients with schizophrenia and 54 healthy control participants. They reached an accuracy of 81.6 (Naïve Bayes), 80.5 (Decision Tree) and 81.3 percent (AdaBoost). With a combined Linear Discriminant Analysis and Support Vector Machine (LDA/SVM) method Li et al. [41] gained an accuracy of 90.48 percent based on a total amount of 48 participants, including 23 schizophrenic persons and 25 healthy persons.

Author	Year	Method	Accuracy	Sample size
Boostani et al. [38]	2009	AdaBoost BDLDA	85.41 % 87.51 %	31
Zhang et al. [39]	2010	HOPD	90.00 %	22
Laton et al. [40]	2014	Naïve Bayes Decision Tree AdaBoost	81.60 % 80.50 % 81.30 %	108
Li et al. [41]	2019	LDA/SVM	90.48 %	48

Table 3: Related work.

Klimesch et al. [42] as well as Olejarczyk et al. [28] state that within the delta, theta and alpha bands, anomalies have been localized using the EEG data of schizophrenic patients. This insight was supported by Howells et al. [43]. For this reason we first decided to apply the unmodified method from the study of Rieg et al. [34] considering the frequency range up to 50 Hz. Thus, we achieved a balanced accuracy of 96.01 %.

In contrast, Uhlhaas et al. [44] refer to findings of abnormalities in the gamma band by investigating EEG data. Taking this into account, we extended the method and investigated the data up to 100 Hz. This enabled us to further improve our balanced accuracy to 96.77 %.

In order to gain more information about predictive frequency bands, we expanded our analysis from the four most important frequency bands to the most important 10 percent of frequencies. Thus, we found that informative frequency subbands are distributed over the entire spectrum of EEG data (Fig. 5).

Importance																				
Hz	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Band	Delta			Theta				Alpha					Beta							

Importance																				
Hz	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Band	Beta												Gamma							

Importance																				
Hz	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
Band	Gamma																			

Importance																				
Hz	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
Band	Gamma																			

Importance																				
Hz	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
Band	Gamma																			

Figure 5: Most important subbands. Predictive frequencies are distributed over the entire EEG spectrum.

These 20 relevant frequency subbands are distributed as follows: One in the delta band, two in the theta band, three in the beta band, four in the alpha band and ten in the gamma band. The corresponding distribution of the subbands can also be seen in figure 5.
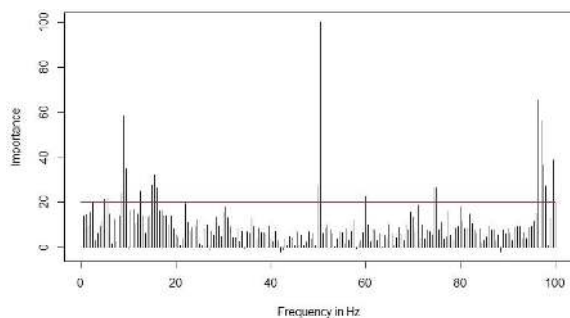

Figure 6: Importance of predictive EEG frequencies.

To provide a clear visualization, the values of importance are normalized by the value of the most important frequency band. The threshold for the 10 percent of the most important subbands is visualized by the red line (see figure 6). Consequently, besides the alpha band, the upper gamma band (96.5 Hz to 99.5 Hz) has a high significance. These findings can lead to an intensification of medical research on EEG data based diagnostics of paranoid schizophrenia.

Due to the very good results we achieved with our classifier, and since so far neither diagnostic tests nor biomarkers are available for detecting schizophrenia [4], our approach can be practically implemented as a diagnostic test into daily medical life. Technical assistance could help to make faster and more detailed decisions. This not only provides relief for physicians but, as diagnosing is based on machine learning algorithms; it also reduces the likelihood of human errors [45] in medical environments. The steps of the diagnosis process are visualized in figure 7 and described as follows:

Step 1: A physician examines a patient and cannot assess whether the patient is suffering from schizophrenia or not.
Step 2: A 10/20 EEG examination is performed. Only one-minute of EEG recording is required.
Step 3: Our fast method works within seconds and can be used to process the EEG data.
Step 4: Our artifact can distinguish with an accuracy of 96.77 percent whether the patient is suffering from schizophrenia or not.
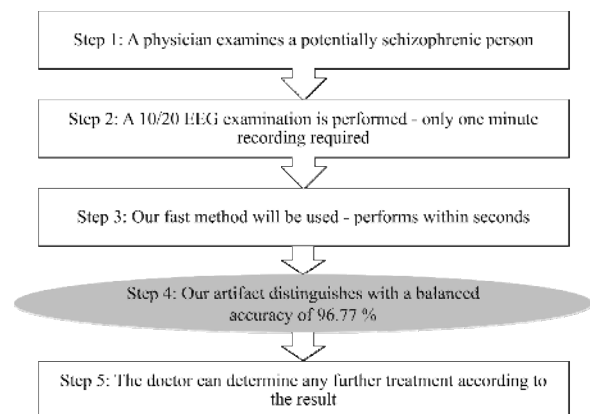**Step 5:** According to these results the doctor can determine any further treatments.

Figure 7: Diagnosis process in daily medical life.

5. Conclusion

In the context of this work we built an efficient Random Forest classifier for identifying whether a person is suffering from schizophrenia or not, based on EEG data. By investigating 499 recordings, with a duration of one-minute respectively, our classifier yields a very good balanced accuracy of 96.77 percent. As demonstrated in figure 4, in total only four misclassifications were made. In comparison to other available works, using different approaches, our

classifier outperforms producing an even better result. Furthermore, we showed that a finer separation of the commonly used frequency bands (alpha, beta, theta, delta and gamma) into 200 frequency bands, with a range of 0.5 Hz each, yields new insights about the most predictive subbands.

5.1 Limitation

While we intensively evaluated other traditional machine learning approaches such as clustering [46] and also most modern convolutional neural networks, which are outstanding in other domains such as image recognition [47-49], we achieved the best results here with our novel tree-based method proposed in [34]. However, the method of choice always limits scientific understanding. Hence our study has these limitations:

While our classification model achieves a good level of accuracy, the classifier is not yet trained and tested on patients suffering from other mental illnesses, which could have one or more symptoms also associated with schizophrenia and thus eventually bias the data. That is why more extensive experimentation is necessary using datasets that contain similar data to schizophrenia behavior. Furthermore, it is sufficiently apparent that medication and personality [50-52] influence the EEG data of schizophrenic patients, and as a result, our classifier. While the internal validity of our model is very high due to the rigorous k-fold-cross-validation, improving external validity by training with additional datasets is also an important step to improve the model. Also, the influence of individual differences in brain activity on EEG e.g. other mental disorders need to be analyzed and considered in future studies.

Another limitation is, that so far the dataset has not been used for this type of classification, which leads to limited grounds for comparability of the performance indicators.

5.2 Future work

In future work we will report common method bias evaluations [53, 54]. In addition, we will triangulate EEG sensor data with other physiological sensor data (i.e., electrocardiogram [55, 56], electro-dermal activity [57, 58], eye fixation [59-61], eye pupil diameter [62-65]). Furthermore, we will experimentally evaluate whether our novel approach is also robust under various conditions of a user's cognitive workload [66-68] and related concepts [69-71], concentration [72], and mindfulness [73. 74]. In addition, we will report results on successfully applying our novel procedure to other schizophrenia data [75], and other diseases such as epilepsy [76, 77] and sleep disorder [78, 79].

In terms of implementing the approach in real clinical environments we will conduct an implementation study to evaluate acceptance [80-82] and trust [83, 84] by physicians and patients and if the automated approach improves the coordination [85, 86] between physicians more efficiently.

Since our analysis is based on professionally recorded data from a public repository future work could not only replicate but also enhance our work. In addition, we want to re-evaluate our classifier on other datasets to increase external validity. In addition, despite achieving good results based on one minute recordings, future work could systematically analyze adjusted time periods.

6. References

[1] OECD, Health at a Glance 2017, Paris, France, 2017.

[2] W. T. Carpenter, and R. W. Buchanan, "Schizophrenia", N Engl J Med, vol. 330, no. 10, pp. 681-690, 1994.

[3] J. L. Ayuso-Mateos, P. Gutierrez-Recacha, J. M. Haro, and D. Chisholm, "Estimating the prevalence of schizophrenia in Spain using a disease model", Schizophr Res, vol. 86, no. 1-3, pp. 194-201, 2006.

[4] M. Owen, A. Sawa, and P. Mortensen, "Schizophrenia", Lancet, vol. 388, no. 10039, pp. 86-97, 2016.

[5] C. Hjorthøj, A. E. Stürup, J. J. McGrath, and M. Nordentoft, "Years of potential life lost and life expectancy in schizophrenia: a systematic review and meta-analysis", Lancet Psychiatry, vol. 4, no. 4, pp. 295-301, 2017.

[6] J. F. Hayes, L. Marston, K. Walters, M. B. King, and D. P. J. Osborn, "Mortality gap for people with bipolar disorder and schizophrenia: UK-based cohort study 2000-2014", Br J Psychiatry, vol. 211, no. 3, pp. 175-181, 2017.

[7] S. L. James et al., "Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017", Lancet, vol. 392, no. 10159, pp. 1789-1858, 2018.

[8] R. Mizutani et al., "Three-dimensional alteration of neurites in schizophrenia", Transl. Psychiatry, vol. 9, no. 85, pp. 1-12, 2019.

[9] C. F. M. G. van Kesteren et al., "Immune involvement in the pathogenesis of schizophrenia: a meta-analysis on

postmortem brain studies", Transl. Psychiatry, vol. 7, no. e1075, pp. 1-11, 2017.

[10] R. Newton, A. Rouleau, A.-G. Nylander, J.-Y. Loze, H. K. Resemann, S. Steeves, and B. Crespo-Facorro, "Diverse definitions of the early course of schizophrenia–a targeted literature review", npj Schizophrenia, vol. 4, No. 21, pp. 1-21, 2018.

[11] B. Deakin et al., "The benefit of minocycline on negative symptoms of schizophrenia in patients with recent-onset psychosis (BeneMin): a randomised, double-blind, placebo-controlled trial", Lancet Psychiatry, vol. 5, no. 11, pp. 885-894, 2018.

[12] E. Chesney, G. Goodwin, and S. Fazel, "Risks of all-cause and suicide mortality in mental disorders: a meta-review", World Psychiatry, vol. 13, no. 2, pp. 153-160, 2014.

[13] I. Kooyman, K. Dean, S. Harvey, and E. Walsh, "Outcomes of public concern in schizophrenia", Br J Psychiatry, vol. 191, no. s50, pp. s29-s36, 2007.

[14] S. Marwaha, and S. Johnson, "Schizophrenia and employment", Soc Psychiatry Psychiatr Epidemiol, vol. 39, no. 5, pp. 337-349, 2004.

[15] J. A. Lieberman, and M. First, "Psychotic Disorders", N Engl J Med, vol. 379, no. 3, pp. 270-280, 2018.

[16] A. Schmitt, A. Hasan, O. Gruber, and P. Falkai, "Schizophrenia as a disorder of disconnectivity", Eue Arch Psy Clin N, vol. 261, no. 2, pp. 150-154, 2011.

[17] World Health Organization, "The ICD-10 Classification of Mental and Behavioural Disorders", Geneva, Switzerland, pp. 78-91, 1993.

[18] K. Tsoi, S. Poon, and P. Hung, Patrick, "Introduction to the Minitrack on Big Data on Healthcare Application", In HICSS-51 Proceedings, p. 2777.

[19] K. Tsoi, S. Poon, and P. Hung, Patrick, "Introduction to the Minitrack on Big Data on Healthcare Application", In HICSS-52 Proceedings, p. 3720.

[20] A. Esteva et al., "A guide to deep learning in healthcare", Nature Medicine, vol. 25, pp. 24-29, 2019.

[21] C. Poon, N. Gorji, M. Latt, K. Tsoi, B. Choi, C. Loy, S. Poon, "Derivation and Analysis of Dynamic Handwriting Features as Clinical Markers of Parkinson's Disease", In HICSS-52 Proceedings, pp. 3721-3730.

[22] S. K. Poon, J. Poon, M. McGrane, X. Zhou, P. W. H. Kwan, R. Zhang, B. Liu, J. Gao, C. Loy, K. Chan, D. Man-yuen Sze, "A novel approach in discovering significant interactions from TCM patient prescription data", Int J Data Min Bioin, vol. 5, no. 4, pp. 353-368, 2011.

[23] B. Moghaddam, and D. Javitt, "From Revolution to Evolution: The Glutamate Hypothesis of Schizophrenia and its Implication for Treatment", Neuropsychopharmacology, vol. 37, pp. 4-15, 2012.

[24] S. Micheloyannis, E. Pachou, C. J. Stam, M. Breakspear, and M. Zervakis, "Small-world networks and disturbed functional connectivity in schizophrenia", Schizophr Res, vol. 87, no. 1-3, pp. 60-66, 2006.

[25] K. J. Friston, "Theoretical neurobiology and schizo-phrenia", Br Med Bull, vol. 52, no. 3, pp. 644-655, 1996.

[26] D. J. A. Smit, C. J. Stam, D. Posthuma, D. I. Boomsma, E. J. C. de Geus, "Heritability of 'small-world' networks in the brain: A graph theoretical analysis of resting-state EEG functional connectivity", Hum Brain Mapp, vol. 29, no. 12, pp. 1368-1378, 2008.

[27] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design Science in Information Systems Research", MIS Qarterly, vol. 28, no. 1, pp. 75-105, 2004.

[28] E. Olejarczyk, and W. Jernajczyk, "Graph-based analysis of brain connectivity in schizophrenia", PLOS ONE, vol. 12, no. 11, pp. 1-28, 2017.

[29] G. Klem, H. Lüders, H. Jasper, and C. Elger, "The ten-twenty electrode system of the International Federation", Electroencephalogr Clin Neurophysiol, vol. 52, pp. 3-6, 1999.

[30] E. Olejarczyk, and W. Jernajczyk, "EEG in schizophrenia", RepOD, 2017, http://dx.doi.org/10.18150/repod.0107441.

[31] G. Müller-Putz, R. Riedl, and S. Wriessnegger, "Elec-troencephalography (EEG) as a Research Tool in the Infor-mation Systems Discipline: Foundations, Measurement, and Applications", CAIS, vol. 37, no. 46, pp. 911-948, 2015.

[32] S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent Component Analysis of Electroencephalo-graphic Data", NIPS, vol. 8, pp. 145-151, 1996.

[33] T.-P. Jung, C. Humphries, T.-W. Lee, S. Makeig, M. J. McKeown, V. Iragui, and T. J. Sejnowski, "Extended ICA Removes Artifacts from Electroencephalographic Recordings", NIPS, vol. 10, pp. 894-900, 1998.

[34] T. Rieg, J. Frick, M. Hitzler, and R. Buettner, "High-performance detection of alcoholism by unfolding the amalgamated EEG spectra using the Random Forests method", In HICSS-52 Proceedings, pp. 3769-3777.

[35] L. Breiman, "Random Forests", Machine Learning, vol. 45, no. 1, pp. 5-32, 2001.

[36] Y. Zhou, J. Guo, L. Fu, and T. Liang, "Research on Aero-engine Maintenance Level Decision Based on Im-

proved Artificial Fish-swarm Optimization Random Forest Algorithm", in SDPC 2018 Proceedings, pp. 606-610.

[37] V. Rodriguez-Galiano, M. Sanchez-Castillo, M. Chica-Olmo, and M. Chica-Rivas, "Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines", Ore Geol Rev, vol. 71, pp. 804-818, 2015.

[38] R. Boostani, K. Sadatnejad, and M. Sabeti, "An efficient classifier to diagnose of schizophrenia based on EEG signals", Expert Syst Appl, vol. 36, no. 3, pp. 6492-6499, 2009.

[39] S. Zhang, Q. Shini, and W. Wang, "Classification of schizophrenia's EEG based on high order pattern discovery", in 2010 IEEE BIC-TA Proc., pp. 1147-1149.

[40] J. Laton, J. Van Schependom, J. Gielen, J. Decoster, T. Moons, J. Keyser, M. Hert, and G. Nagels, "In search of biomarkers using electroencephalography", 2014 PRNI Proceedings, pp. 1-4.

[41] F. Li et al., "Differentiation of Schizophrenia by Combining the Spatial Brain Network Patterns of Rest and Task P300", in IEEE T Neur Sys Reh, vol. 27, no. 4, pp. 594-602, 2019.

[42] W. Klimesch, P. Sauseng, and S. Hanslmayr, "EEG alpha oscillations: The inhibition-timing hypothesis", Brain Res Rev, vol. 53, no.1, pp. 63-88, 2007.

[43] F. M. Howells et al., "Electroencephalographic delta/alpha frequency activity differentiates psychotic disorders: a study of schizophrenia, bipolar disorder and methamphetamine-induced psychotic disorder", Transl Psychiatry, vol. 8, no. 75, pp. 1-11, 2018.

[44] P. J. Uhlhaas and W. Singer, "Abnormal neural oscillations and synchrony in schizophrenia", Nat Rev Neurosci, vol. 11, no. 2, pp. 100-113, 2010.

[45] Y. Donchin et al., "A look into the nature and causes of human errors in the intensive care unit", Crit Care Med, vol. 23, no. 2, pp. 294-300, 1995.

[46] F. Baumann, D. Emmert, H. Baumgartl, and R. Buettner, "Hardcore Gamer Profiling: Results from an unsupervised learning approach to playing behavior on the Steam platform", Procedia Computer Science, vol. 126, pp. 1289-1297, 2018.

[47] R. Buettner, and H. Baumgartl "A highly effective deep learning based escape route recognition module for autonomous robots in crisis and emergency situations", HICSS-52 Proceedings, pp. 659-666, 2019.

[48] H. Baumgartl, and R. Buettner, "Development of a highly precise place recognition module for effective human-robot interactions in changing lighting and viewpoint conditions", HICSS-53 Proceedings, 2020, in press.

[49] H. Baumgartl, J. Tomas, R. Buettner, and M. Merkel, "A novel Deep-Learning Approach for Automated Non-Destructive Testing in Quality Assurance based on Convolutional Neural Networks", ACEX-2019 Proceedings, 2019, in press.

[50] R. Buettner, "Predicting user behavior in electronic markets based on personality-mining in large online social networks: A personality-based product recommender framework," Electronic Markets, vol. 27, no. 3, pp. 247-265, 2017.

[51] R. Buettner, "Innovative Personality-based Digital Services", PACIS 2016 Proceedings, 2016.

[52] R. Buettner, "Personality as a predictor of business social media usage: An empirical investigation of XING usage patterns", PACIS 2016 Proceedings, 2016.

[53] R. Buettner, "Getting a job via career-oriented social networking markets: The weakness of too many ties", Electronic Markets, vol. 27, no. 4, pp. 371-385, 2017.

[54] R. Buettner, "Analyzing the Problem of Employee Internal Social Network Site Avoidance: Are Users Resistant due to their Privacy Concerns?", HICSS-48 Proceedings, pp. 1819-1828, 2015.

[55] R. Buettner, L. Bachus, L. Konzmann, and S. Prohaska, "Asking Both the User's Heart and Its Owner: Empirical Evidence for Substance Dualism", Information Systems and Neuroscience, vol. 29, pp. 251-257, 2018.

[56] R. Buettner, and M. Schunter, "Efficient machine learning based detection of heart disease", Proc. IEEE Healthcom 2019, in press.

[57] R. Buettner, "Asking both the User's Brain and its Owner using Subjective and Objective Psychophysiological NeuroIS Instruments", ICIS 2017 Proceedings, 2017.

[58] A. Eckhardt, C. Maier, and R. Buettner, "The Influence of Pressure to Perform and Experience on Changing Perceptions and User Performance: A Multi-Method Experimental Analysis", ICIS 2012 Proceedings, 2012.

[59] A. Eckhardt et al., "Objective measures of IS usage behavior under conditions of experience and pressure using eye fixation data", ICIS 2013 Proceedings, 2013.

[60] R. Buettner, "A user's cognitive workload perspective in negotiation support systems: An eye-tracking experiment", PACIS 2016 Proceedings, 2016.

[61] R. Buettner, "Cognitive Workload of Humans Using Artificial Intelligence Systems: Towards Objective Measurement Applying Eye-Tracking Technology", LNAI, vol. 8077, pp. 37-48, 2013.

[62] R. Buettner, "Social inclusion in eParticipation and eGovernment solutions: A systematic laboratory-experimental approach using objective psychophysiological measures", LNI Proc., vol. P-221, pp. 260-261, 2013.

[63] R. Buettner, "Investigation of the Relationship Between Visual Website Complexity and Users' Mental Workload: A NeuroIS Perspective", Information Systems and Neuroscience, vol. 10, pp. 123-128, 2015.

[64] R. Buettner, B. Daxenberger, A. Eckhardt, and C. Maier, "Cognitive Workload Induced by Information Systems: Introducing an Objective Way of Measuring based on Pupillary Diameter Responses", SIGHCI 2013 Proc., 2013.

[65] R. Buettner, S. Sauer, C. Maier, and A. Eckhardt, "Towards ex ante Prediction of User Performance: A novel NeuroIS Methodology based on Real-Time Measurement of Mental Effort", HICSS-48 Proc., pp. 533-542, 2015.

[66] R. Buettner, "Analyzing Mental Workload States on the Basis of the Pupillary Hippus", in NeuroIS '14 Proc., p. 52, 2014.

[67] R. Buettner, "Investigation of the Relationship Between Visual Website Complexity and Users' Mental Workload: A NeuroIS Perspective", in Information Systems and Neuro Science, vol. 10 of LNISO, pp. 123-128, 2015.

[68] R. Buettner, S. Sauer, C. Maier, and A. Eckhardt, "Real-time Prediction of User Performance based on Pupillary Assessment via Eye Tracking", AIS Trans Hum.-Comput Interact, vol. 10, no. 1, pp. 26-56, 2018.

[69] R. Buettner, "The relationship between visual website complexity and a user's mental workload: A NeuroIS perspective", in Information Systems and Neuro Science, vol. 16 of LNISO, pp. 107-113, Springer, 2016.

[70] R. Buettner, "A user's cognitive workload perspective in negotiation support systems: An eye-tracking experiment", in PACIS 2016 Proc.

[71] R. Buettner, I. J. Timm, I. F. Scheuermann, C. Koot, and M. Roessle, "Stationarity of a user's pupil size signal as a precondition of pupillary-based mental workload evaluation", in Information Systems and Neuro Science, vol. 25 of LNISO, pp. 195-200, Springer, 2017.

[72] R. Buettner, H. Baumgartl, and D. Sauter, "Microsaccades as a predictor of a user's level of concentration", in Information Systems and Neuroscience, vol. 29 of LNISO, pp. 173-177, Springer, 2018.

[73] S. Sauer, J. Lemke, W. Zinn, R. Buettner, and N. Kohls, "Mindful in a random forest: Assessing the validity of mindfulness items using random forests methods", Pers Individ Differ, vol. 81, pp. 117-123, 2015.

[74] S. Sauer, R. Buettner, T. Heidenreich, J. Lemke, C. Berg, and C. Kurz, "Mindful Machine Learning: Using Machine Learning Algorithms to Predict the Practice of Mindfulness", Eur J Psychol Assess, vol. 34, no. 1, pp. 6-13, 2018.

[75] R. Buettner, M. Hirschmiller, K. Schlosser, M. Roessle, M. Fernandes, and I. J. Timm, "High-performance exclusion of schizophrenia using a novel machine learning method on EEG data", in IEEE Healthcom 2019 Proc., in press.

[76] R. Buettner, T. Rieg, and J. Frick, "High-performance detection of epilepsy in seizure-free EEG recordings: A novel machine learning approach using very specific epileptic EEG sub-bands", in ICIS 2019 Proc., in press.

[77] R. Buettner, T. Rieg, and J. Frick, "Machine Learning based Diagnosis of Diseases Using the Unfolded EEG Spectra: Towards an Intelligent Software Sensor", in Information Systems and Neuroscience, vol. 32 of LNISO, Springer, 2019, in press.

[78] R. Buettner, J. Fuhrmann, and L. Kolb, "Towards high-performance differentiation between Narcolepsy and Idiopathic Hypersomnia in 10 minute EEG recordings using a Novel Machine Learning Approach," in IEEE Healthcom 2019 Proc., in press.

[79] R. Buettner, A. Grimmeisen, and A. Gotschlich, "High-performance Diagnosis of Sleep Disorders: A Novel, Accurate and Fast Machine Learning Approach Using Electroencephalographic Data," in HICSS-53 Proc., in press.

[80] R. Buettner, B. Daxenberger, and C. Woesle, "User acceptance in different electronic negotiation systems - a comparative approach", in ICEBE 2013 Proc., pp. 1-8, 2013.

[81] R. Buettner, "Towards a New Personal Information Technology Acceptance Model: Conceptualization and Empirical Evidence from a Bring Your Own Device Dataset", in AMCIS '15 Proc., 2015.

[82] R. Buettner, "Getting a Job via Career-oriented Social Networking Sites: The Weakness of Ties", in HICSS-49 Proc., pp. 2156-2165, 2016.

[83] F. Meixner and R. Buettner, "Trust as an Integral Part for Success of Cloud Computing", in ICIW 2012 Proc., pp. 207-214, 2012.

[84] R. Buettner, "Cooperation in Hunting and Food-sharing: A Two-Player Bio-inspired Trust Model", in BIONETICS '09 Proc., pp. 1-10, 2009.

[85] R. Buettner, "A Classification Structure for Automated Negotiations," in WI-IAT 2006 Proc., pp. 523-530, 2006.

[86] R. Buettner, "The State of the Art in Automated Negotiation Models of the Behavior and Information Perspective," ITSSA, vol. 1, no. 4, pp. 351-356, 2006.