Comparing Supervised Machine Learning Algorithms for Patient-Specific and Non-Patient Specific Epileptic Seizure Detection Using Multichannel Scalp EEG

Jintian Lyu, Lingchao Mao
Georgia Institute of Technology
ISyE 6740 Fall 2020

1. Problem Statement

Seizure is a paroxysmal alteration of neurologic function caused by the excessive, hypersynchronous discharge of neurons in the brain, which leads to changes in people's behavior, movements or feelings, and in levels of consciousness [1]. *Epileptic seizure* is used to distinguish seizure caused by abnormal neuronal firing from a nonepileptic event, such as a psychogenic seizure. *Epilepsy* is a chronic brain dysfunction syndrome characterized by recurrent, unprovoked seizures. About 1% of the world population suffers from epilepsy and among which, about one-third has refractory epilepsy (seizures retained after two or more medications) [2].

The electroencephalogram (EEG) is one of the oldest non-invasive technologies to measure the neuronal activity of the human brain. EEG records the voltage fluctuations from multiple electrodes placed on the subject's scalp during an extended period of time (Figure 1) [21]. The first recording of the electric field of the human brain was made by the German psychiatrist Hans Berger in 1924 [3]. The detection of EEG signals is one of the most important means to diagnose epilepsy. Attempts to apply scalp EEG data in the detection and prediction of seizures have brought charming outcomes during clinical practice [4].

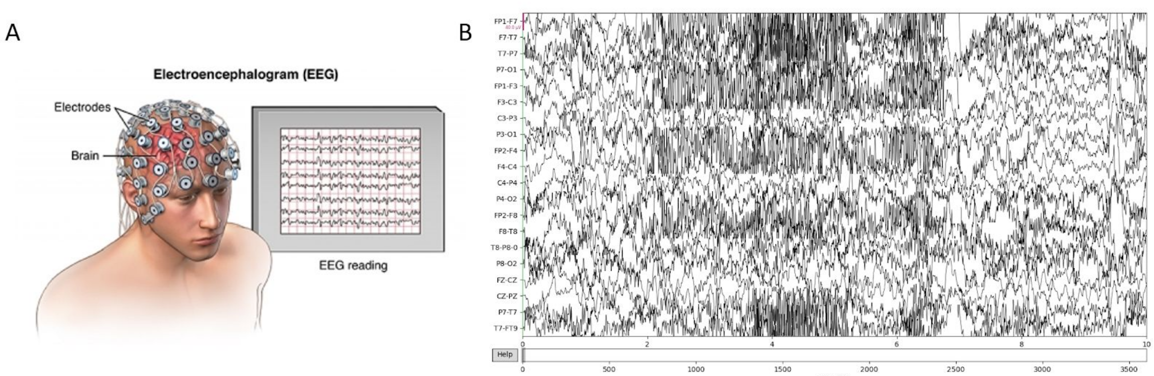


Figure 1. (A) Diagram of electroencephalogram (EEG) [5]; (B) Visualization of 1 hour EEG of chb01

With the development of computer technology, a machine learning algorithm capable of detecting patients' seizure with high accuracy and efficiency shows high clinical significance. However, there are several challenges in the development of such algorithms:

- The EEG of epilepsy patients constantly transitions between regimes within the seizure and non-seizure states; among which seizure signals are very scarce, leading to a highly imbalanced classification problem [6]
- EEG data is contaminated with signals caused by physiological artifacts, such as sleep spindles. Not all abnormal rhythmic activity is associated with seizures [6]

- EEG data are highly dimensional over long-extended periods such that it is infeasible to run algorithms on raw data. Therefore, data preprocessing and feature extraction are key to classification performance and require a tradeoff between efficiency and accuracy
- A patient can exhibit more than one type of seizure, such as partial seizure versus generalized seizure. Each type has stereotypical electrographic patterns for that patient [7]. However, the signal location and rhythmic activity can vary significantly across patients. Cross-patient variability in brian signals causes patient non-specific classifiers to exhibit poor accuracy [8].

Many studies leveraged machine learning algorithms to build seizure detectors. Amongst these, the most popular classification algorithms are Support Vector Machines (SVM) and Neural Networks [9–11]. Other models include Linear Discriminant Analysis (LDA), Quadratic Linear Discriminant Analysis (QLDA), Naive Bayes, and K-Nearest Neighbors (KNN) [11, 12]. The majority of existing studies focus on building patient-specific classifiers due to the challenge of heterogeneous brain signals across patients. Although patient-specific classifiers exhibit compromising accuracies, the disadvantage is that they require model training and parameter adjustments for individual patients [13]. In this study, we build and compare patient-specific and non-patient specific classifiers for EEG epileptic seizure using five nonlinear machine learning algorithms learned in this course.

In terms of feature extraction, a variety of techniques exists in literature such as time-frequency distributions (TFD), fast fourier transform (FFT), eigenvector methods (EM), multi-wavelet decomposition (WT), and time series methods (ARM) [14]. Some studies criticized Fourier transform (FFT) because it assumes signal stationarity, which is not applicable in the case of EEG signals [15]; others have shown that frequency-domain signals have greater potential than time domain signals [16]. Multiple studies found that median frequency, sample entropy, and root mean square have the most potential in seizure classification [17]. Some of the newest studies applied DL models on raw EEG data without feature extraction [18, 19]. In this study, we include a comprehensive set of features from both frequency and time domain collected from the literature.

This study has two specific aims:
Specific Aim 1: Build accurate classifiers to detect epileptic seizure using multichannel EEG data
Specific Aim 2: Compare the performance of five machine learning algorithms for patient-specific and non-patient specific seizure classification

2. Data Source

We use a well known public data source, CHB-MIT Scalp EEG Database [6]. The database consists of EEG recordings from 23 pediatric subjects with intractable seizures collected at the Children's Hospital Boston. EEG signals were sampled at 256 samples per second with 16-bit resolution and recorded simultaneously at 21-28 channels [20]. Seizure events and corresponding start and end times were identified by clinical experts. Due to the limited computation capacity, we select five subjects considering their gender and age distribution. The demographics of the selected subjects are shown in Table 1.

Table 1. Summary of subject demographics

Subject	chb01	chb02	chb03	chbe05	chb08
Gender	Female	Male	Female	Female	Male
Age	11	11	14	7	3.5

3. Methodology

In this section, we describe our methods used to preprocess raw EEG data and extract time and frequency domain features. Then we discuss the selection of algorithms used for patient-specific and non-patient specific classification. Finally, we show the two approaches used to evaluate the models. An overview of this process is visualized in Figure 2. The data and scripts used in this study are publicly available at https://github.com/lingchm/epilepsy_eeg_classification

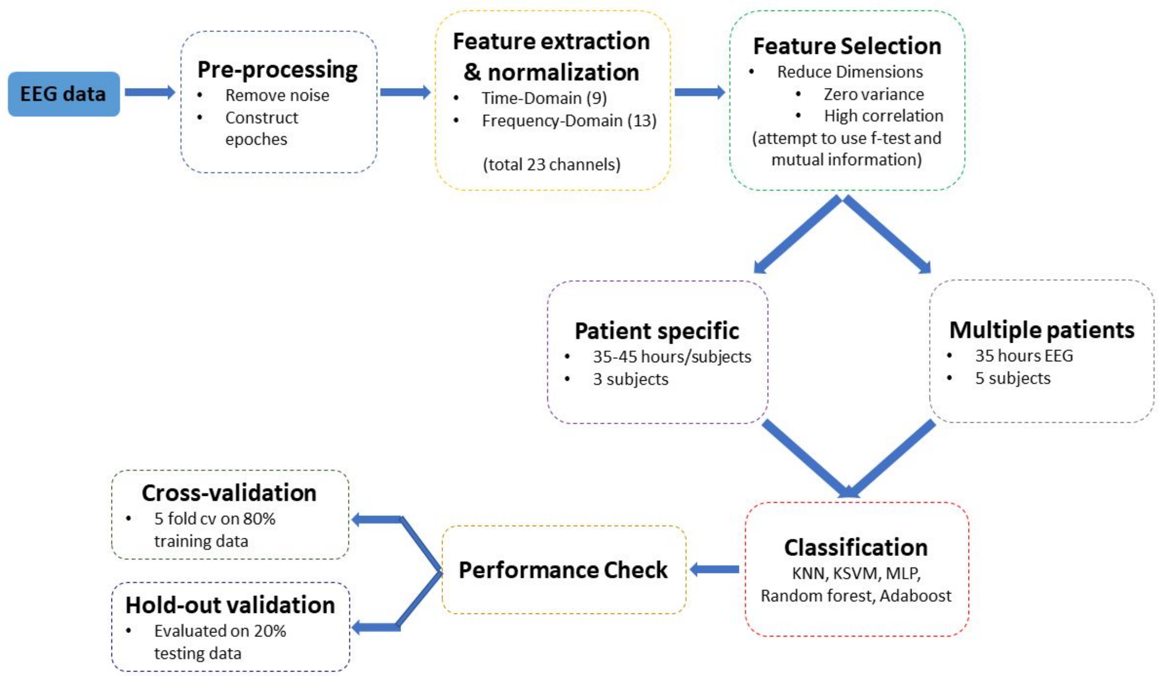


Figure 2. Overview of EEG seizure classification pipeline

3.1. EEG Preprocessing and Feature Extraction

Raw EEG data is recorded in one-hour long edf files. First, we apply a 25 Hz lowpass filter and 0.25 Hz high-pass filter on raw data to remove the noise of brain signals, since it has been shown that most seizure and non-seizure signals fall in this range [6]. Then, raw data is divided into a sequence of 10-second epochs, each epoch being a one-second increment of the previous. This epoch length was chosen because experts require an EEG abnormality to persist and evolve for a minimum of 6-10 seconds before considering the abnormality a seizure [6]. From each epoch, 22 features are extracted for each of the 23 channels, creating a total of 506 features per epoch. We select a list of time and frequency-domain features that are identified to be important from the literature and that are computable given our computational resources [4, 12, 17]. Features are calculated using functions from the *pyeeg* package [21] and using the comprehensive guide of EEG features by Boonyakitanont1 et al. [17].

Time Domain Features (TDFs)

- Common statistics measures including variance, maximum, minimum, kurtosis (tailoring of the distribution), skewness (data asymmetry)
- Root Mean Squared (RMS)
- Number of zero-crossings
- Number of peaks
- Hurst Exponent (degree of time-series tendency)
- Higuchi and Petrosian fractal dimension measures (mathematical index for signal complexity)

- Hjorth mobility and complexity parameters (to characterize the the spectral properties of EEG signals

Frequency Domain Features (FDFs)
- Power ratio of the current and background epochs at 1, 5, 10, 20 Hz sub-bands
- Peak or dominant frequency
- Median frequency
- Spectral entropy (measure of the random process uncertainty from the frequency)
- Total power

From the epoch level dataset obtained from EEG preprocessing, we remove the features with zero variance and features that are highly correlated with each other (>0.9). Since feature values have a wide range of negatives and positives, we apply feature normalization.

The patient-specific datasets are constructed using all EEG recordings available in the database for the first three subjects (35-45 hours). We randomly sample 35 hours of data from five subjects to obtain a patient-mixer data set to train the non-patient specific classifier.

3.2. Data Imbalance
The data used in this study is highly imbalanced data, where the majority of samples are non-seizure (more than 99%). Standard classification algorithms are biased towards the class that has a higher number of instances because they are designed to optimize overall accuracy. Given such a high imbalance, the samples of the minority class may be treated as noise [22]. There are three main approaches to handle imbalanced data, either from 1) data level, such as Synthetic Minority Over-sampling Technique (SMOTE) [23]; from 2) algorithm level, such as specifying cost or weights favoring the minority class; or 3) ensemble techniques, such as bagging or boosting. Although SMOTE mitigates the problem of overfitting caused by other resampling methods, it does not take into consideration neighboring examples from other classes, which results in an increase in the overlapping of classes and can introduce additional noise. SMOTE also has low efficiency in high dimensional data. Therefore, we decide to use algorithm level adjustments and ensemble methods to handle the imbalance in this study.

3.3. Classification models
We apply five nonlinear classification methods: Multi-Layer Perceptron (MLP), K-Nearest Neighbors (KNN), Kernelized Support Vector Machine (KSVM), Random Forest, and Adaboost, implemented by the *scikit-learn* python library. The default parameters were used for the models because patient-specific parameter tuning would not be realistic in a real application scenario.

3.3.1. MLP
Inspired by the biological neural networks, Artificial Neural Networks (ANN) are based on a collection of connected units or nodes which are capable of capturing highly complex, nonlinear patterns. An MLP is a class of feed-forward ANN. It is flexible to use in different problems with its fast speed, good accuracy once trained [24]. The number of inputs and layers can be tuned to achieve the best performance. There are two major drawbacks in applying MLP. First, its “black-box” nature poses interpretability challenges, such as we cannot know how much each independent variable is influencing the dependent variable. Second, MLP can be computationally expensive and time-consuming to train with traditional CPUs. We apply an MLP with ReLU activation function, 10 hidden layers, constant learning rate, and 0.9 momentum.

3.3.2. KNN

KNN is a robust classifier which is often used as a benchmark for more complex classifiers, such as SVM. KNN is intuitive, easy to implement, and does not need assumptions. Another advantage of KNN is that it is a Lazy Learner (instance-based learning) that does not need a training step, therefore new data can be added seamlessly. However, KNN does not work well on high-dimensional data because computing the distance between every point becomes computationally expensive. The implementation of KNN requires careful feature preprocessing to make feature value ranges relatively homogeneous. Another disadvantage of KNN is that it is highly sensitive to noise and outliers [25]. We consider KNN as the benchmark in this study.

3.3.3. KSVM

SVM, a well-known supervised learning model, is one of the most robust classification methods with high efficiency in high dimensional data and economic memory usage. Given the complexity of EEG data, Kernel SVM is predicted to have good performance. However, some studies pointed out SVM might have declined performance with a large dataset [26]. The parameters need to be fine-tuned if the target classes are overlapping and the dataset has a lot of noise [27]. We apply SVM with RBF kernel because EEG data is clearly complex and nonlinear. A class weight of 1:100 is assigned as an algorithmic level approach to data imbalance.

3.3.4. Random Forests

Random Forests are an ensemble learning method by constructing a multitude of decision trees. This method is usually robust to outliers and works well with non-linear data. Though bagging, Random Forest lowers the risk of overfitting and runs efficiently on a large dataset compared to conventional decision trees. However, random forests are found to be biased while dealing with categorical variables and the training process can be very time-consuming [28]. We implement random forests with class weight of 1:100 as an algorithmic level approach to data imbalance.

3.3.5. AdaBoost

Short for adaptive boosting, AdaBoost is adaptive in the sense that subsequent weak learners are emphasized to improve those instances misclassified by previous classifiers. At each iteration, AdaBoost selects the best features known to improve the predictive power of the model as decision stumps, which makes it suitable in this situation with high complexity. AdaBoost has shown compromising performance over traditional algorithms on many datasets. However, it also has a few disadvantages such as it is particularly vulnerable to uniform noise [29]. Sometimes, weak classifiers can lead to low margins and overfitting. We apply AdaBoost with decision trees as base learners.

3.4. Evaluation Criteria

Three metrics are used to evaluate model performance: accuracy, True Positive Rate (TPR), and False Positive Rate (FPR). Accuracy measures the overall percentage of correct classifications; TPR is the percent of true positives correctly classified, and FPR indicates the percent of samples being falsely classified as positive. We choose these metrics considering a future application scenario. For a portable seizure detector, the TPR is of the highest importance since the seizure detector must properly identify seizure to ensure prompt medical interference. The second important readout is FPR of non-seizure, because having high false alarms can significantly worsen user experience. In addition, FPR can indicate how good the model is at correctly separating seizure signals from physiological and pathological EEG activities as well as various artifacts [13]. It is worth mentioning that a high accuracy score can be misleading in highly imbalanced data. A model that classifies all events as non-seizure can still obtain a high accuracy score.

We randomly split data into 80% training and 20% testing. We report the performance measures from the testing data as hold-out validation, which measures the model's effectiveness on new data; and the average performance of a 5-fold cross-validation performed on the training data.

4. Evaluation and Results

From the preprocessing procedure described in Section 3.1, we obtain four data sets as shown in Table 2.

Table 2. Summary of preprocessed EEG data sets

	Subject	EEG hours	Number of seizures	Total epochs	Seizure epocks	% events
Patient specific	chb01	44	5	145610	505	0.35%
	chb02	35	3	126842	207	0.16%
	chb03	38	7	136464	465	0.34%
Multiple patients	01, 02, 03, 05, 08	35	5	125685	430	0.34%

The classification performance of the patient-specific classifiers for the three subjects is shown in Tables 3-5. The majority of TPR of the five algorithms range from 0.7 to 0.9, and FPRs are below 0.0015. Kernel SVM is the only model that correctly classifies more than 95% of the seizures for all three subjects while maintaining a low number of false positives. While KNN achieves a 0.95 TPR for some patients, it does not have a good performance on other patients, which shows that KNN is not robust to changes in data. In addition, KNN seems to face the overfitting problem as seen by having a better cross-validation performance and worse hold-out validation performance on all three subjects. MLP detects over 85% of seizures for chb01 and chb03 with a nearly zero false-positive rate, however, its performance decays to 59% for chb02. One reason for this significant performance decay may be because chb02 has less than half of the proportion of seizure samples than the other subjects, so it is even more challenging for the algorithm to capture the minority class. On the other hand, both ensemble methods, Random Forest and AdaBoost, achieve nearly zero FPR in the sacrifice of a lower TPR compared to other classifiers. Adjusting their hyperparameters may improve the TPR. The performance of our classifiers is similar to the existing patient-specific classifiers from the literature, which range from 70-90% TPR [9, 10, 30–32]. In terms of computation speed, KNN is significantly more time expensive than other models, followed by Adaboost, Kernel SVM, Random Forest, and MLP.

Table 6 shows the classification performance of the five models on the non-patient specific data set. Kernel SVM achieves the best performance by detecting 97% of the seizures and significantly outperforms other classifiers. The comparison of patient-specific versus non-patient specific classifiers can be visualized in Figure 3. In general, patient-specific classifiers achieve higher TPR and lower FPR than non-patient specific classifiers. Kernel SVM, KNN, and Random Forest are the algorithms whose performance is less impacted by patient heterogeneity; amongst which Kernel SVM and Random Forest were implemented with 1:100 class weight assignments to emphasize more on correct predictions for the seizure class. Both MLP and AdaBoost's TPR decreased from >75% to < 65% on the non-patient specific data set. We also compare our best performing classifier with Reveal, a commercially available, patient non-specific detector that uses Neural Networks trained on a large pool of adult patients [8]. When evaluated on the CHB-MIT dataset, Reveal achieved 61% TPR with 33 false positives per 24 hour period [6], which is similar to the performance of our MLP, Random Forest, and AdaBoost classifiers.

Table 3. Classification performance of patient-specific classifiers on chb01 (44h)

		MLP	KNN	Kernel SVM	Random Forest	AdaBoost
cross validation	Accuracy	0.9994	0.9991	0.9978	0.9994	**0.9994**
	TPR	0.8577	0.7467	**0.9924**	0.8166	0.8547
	FPR	0.0001	0.0000	0.0022	**0.0000**	0.0001
hold-out validation	Accuracy	**0.9996**	0.9988	0.9974	0.9993	0.9992
	TPR	0.8899	0.6881	**0.9817**	0.8073	0.8349
	FPR	**0.0000**	0.0001	0.0025	**0.0000**	0.0002

Table 4. Classification performance of patient-specific classifiers on chb02 (35h)

		MLP	KNN	Kernel SVM	Random Forest	AdaBoost
cross validation	Accuracy	0.9993	**0.9999**	0.9988	0.9995	0.9997
	TPR	0.5965	0.9343	**0.9480**	0.7111	0.8287
	FPR	0.0001	**0.0000**	0.0011	**0.0000**	**0.0000**
hold-out validation	Accuracy	0.9995	0.9996	0.9987	0.9993	**0.9997**
	TPR	0.7273	0.8182	**0.9773**	0.6136	0.8409
	FPR	**0.0000**	0.0000	0.0013	**0.0000**	0.0000

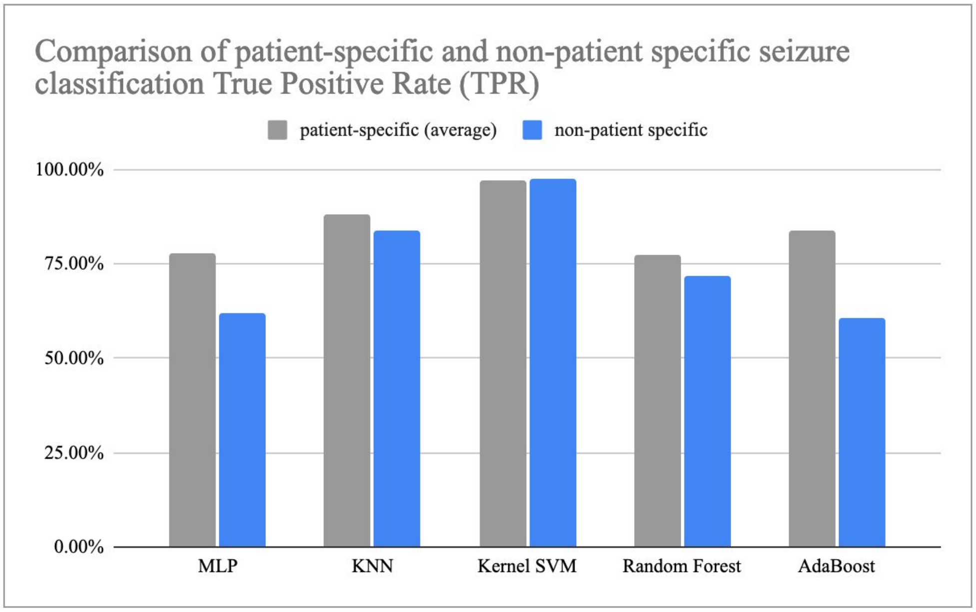
Table 5. Classification performance of patient-specific classifiers on chb03 (38h)

		MLP	KNN	Kernel SVM	Random Forest	AdaBoost
cross validation	Accuracy	0.9995	**0.9998**	0.9983	0.9993	0.9993
	TPR	0.8750	**0.9594**	0.9713	0.7952	0.8366
	FPR	0.0001	0.0000	0.0017	**0.0000**	0.0001
hold-out validation	Accuracy	0.9990	**0.9998**	0.9985	0.9990	0.9991
	TPR	0.7391	**0.9348**	0.9783	0.7065	0.7826
	FPR	0.0001	0.0000	0.0015	**0.0000**	0.0001

Table 6. Classification performance of non-patient specific classifier (35 hours)

		MLP	KNN	Kernel SVM	Random Forest	AdaBoost
cross validation	Accuracy	0.9984	0.9994	0.9965	**0.9990**	0.9983
	TPR	0.6203	0.8400	**0.9736**	0.7194	0.6055
	FPR	0.0003	0.0000	0.0035	**0.0000**	0.0004
hold-out validation	Accuracy	0.9995	0.9995	0.9967	**0.9992**	0.9987
	TPR	0.8684	0.8421	**0.9605**	0.7237	0.6316
	FPR	0.0001	**0.0000**	0.0032	**0.0000**	0.0002

Figure 3. Comparison of TPR of patient-specific and non-patient specific classifiers



We have also attempted to use F-test and Mutual Information to select the best 50, 100, 200 features. However, selecting a subset of features significantly worsened classification performance. More careful feature selection or channel selection methods can be considered in the future since multi-channel EEG signals may be redundant [12], such as an attempt by [33] using back-ward elimination. The EEG channel selection is still an open research question.

In conclusion, we build accurate classifiers to detect epileptic seizures using multichannel EEG data, achieving 70-90% TPR and below 0.2% FPR. We compare the performance of five machine learning algorithms for patient-specific and non-patient specific seizure classification. Kernel SVM outperforms the other models for both patient-specific and non-patient specific seizure detection, achieving >95% TPR in both scenarios; also outperforming an existing commercial non-patient seizure detector. We observe that adjusting class weight assignments can significantly improve the classification performance of the minority class when data is highly imbalanced. For future studies, we recommend to include a larger pool of subjects to better assess the performance of non-patient specific classifiers.

Scalp-EEG-based seizure detection systems can be useful in clinical epilepsy monitoring units as a support tool for trained personnel for continuous supervision of patients [13]. In an outpatient setting, currently, there are no implementations of scalp EEG available because it is inconvenient and uncomfortable for patients to wear electrodes for long-lasted hours. These seizure detection algorithms can be used with future portable EEG devices that can wear in everyday life.

References

[1] M. D. Holmes, “The causes of epilepsy: common and uncommon causes in adults and children,” *Arch. Neurol.*, vol. 69, no. 9, pp. 1212–1213, Sep. 2012, doi: 10.1001/archneurol.2012.1266.

[2] W. J. Nowack, “Book Review: Epilepsy: Frequency, Causes, and Consequences, by W. Allen Hauser and Dale C. Hesdorffer. Published in 1990 by Demos Publications, New York, 378 pages, $74.75,” *Journal of Child Neurology*, vol. 6, no. 4. pp. 374–374, 1991, doi: 10.1177/088307389100600416.

[3] A. Coenen and O. Zayachkivska, “Adolf Beck: A pioneer in electroencephalography in between Richard Caton and Hans Berger,” *Adv. Cogn. Psychol.*, vol. 9, no. 4, pp. 216–221, Dec. 2013, doi: 10.2478/v10053-008-0148-3.

[4] G. R. Minasyan, J. B. Chatten, M. J. Chatten, and R. N. Harner, “Patient-specific early seizure detection from scalp electroencephalogram,” *J. Clin. Neurophysiol.*, vol. 27, no. 3, pp. 163–178, Jun. 2010, doi: 10.1097/WNP.0b013e3181e0a9b6.

[5] S. Siuly, Y. Li, and Y. Zhang, “Electroencephalogram (EEG) and Its Background,” in *EEG Signal Analysis and Classification: Techniques and Applications*, S. Siuly, Y. Li, and Y. Zhang, Eds. Cham: Springer International Publishing, 2016, pp. 3–21.

[6] A. Shoeb, H. Edwards, J. Connolly, B. Bourgeois, T. Treves, and J. Guttag, “Patient-specific seizure onset detection,” *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2006, pp. 419–422, 2004, doi: 10.1109/IEMBS.2004.1403183.

[7] A. H. Shoeb, “Application of machine learning to epileptic seizure onset detection and treatment,” Massachusetts Institute of Technology, 2009.

[8] S. B. Wilson, M. L. Scheuer, R. G. Emerson, and A. J. Gabor, “Seizure detection: evaluation of the Reveal algorithm,” *Clin. Neurophysiol.*, vol. 115, no. 10, pp. 2280–2291, Oct. 2004, doi: 10.1016/j.clinph.2004.05.018.

[9] K. Rasheed *et al.*, “Machine Learning for Predicting Epileptic Seizures Using EEG Signals: A Review,” *IEEE Rev. Biomed. Eng.*, vol. PP, Jul. 2020, doi: 10.1109/RBME.2020.3008792.

[10] B. Karlik and S. Hayta, “Comparison Machine Learning Algorithms for Recognition of Epileptic Seizures in EEG,” 2014, Accessed: Dec. 06, 2020. [Online]. Available: http://iwbbio.ugr.es/2014/papers/IWBBIO_2014_paper_1.pdf.

[11] A. Nandy, M. A. Alahe, S. M. Nasim Uddin, S. Alam, A. Nahid, and M. A. Awal, “Feature Extraction and Classification of EEG Signals for Seizure Detection,” in *2019 International Conference on Robotics,Electrical and Signal Processing Techniques (ICREST)*, Jan. 2019, pp. 480–485, doi: 10.1109/ICREST.2019.8644337.

[12] P. Boonyakitanont, A. Lek-uthai, K. Chomtho, and J. Songsiri, “A review of feature extraction and performance evaluation in epileptic seizure detection using EEG,” *Biomed. Signal Process. Control*, vol. 57, p. 101702, Mar. 2020, doi: 10.1016/j.bspc.2019.101702.

[13] C. Baumgartner and J. P. Koren, “Seizure detection using scalp-EEG,” *Epilepsia*, vol. 59 Suppl 1, pp. 14–22, Jun. 2018, doi: 10.1111/epi.14052.

[14] A. S. Al-Fahoum and A. A. Al-Fraihat, “Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains,” *ISRN Neurosci*, vol. 2014, p. 730218, Feb. 2014, doi: 10.1155/2014/730218.

[15] J. Rasekhi, M. R. K. Mollaei, M. Bandarabadi, C. A. Teixeira, and A. Dourado, “Preprocessing effects of 22 linear univariate features on the performance of seizure prediction methods,” *Journal of Neuroscience Methods*, vol. 217, no. 1–2. pp. 9–16, 2013, doi: 10.1016/j.jneumeth.2013.03.019.

[16] M. Zhou *et al.*, “Epileptic Seizure Detection Based on EEG Signals and CNN,” *Front. Neuroinform.*, vol. 12, p. 95, Dec. 2018, doi: 10.3389/fninf.2018.00095.

[17] P. Fergus, D. Hignett, A. Hussain, D. Al-Jumeily, and K. Abdel-Aziz, “Automatic epileptic seizure detection using scalp EEG and advanced artificial intelligence techniques,” *Biomed Res. Int.*, vol. 2015, p. 986736, Jan. 2015, doi: 10.1155/2015/986736.

[18] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, “Deep convolutional neural network

for the automated detection and diagnosis of seizure using EEG signals," *Comput. Biol. Med.*, vol. 100, pp. 270–278, Sep. 2018, doi: 10.1016/j.compbiomed.2017.09.017.
[19] R. Hussein, H. Palangi, R. Ward, and Z. Jane Wang, "Epileptic Seizure Detection: A Deep Learning Approach," *arXiv [eess.SP]*, Mar. 27, 2018.
[20] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. E215–20, Jun. 2000, doi: 10.1161/01.cir.101.23.e215.
[21] F. S. Bao, X. Liu, and C. Zhang, "PyEEG: an open source Python module for EEG/MEG feature extraction," *Comput. Intell. Neurosci.*, vol. 2011, p. 406391, Mar. 2011, doi: 10.1155/2011/406391.
[22] Y. Sun, A. K. C. Wong, and M. S. Kamel, "CLASSIFICATION OF IMBALANCED DATA: A REVIEW," *Int. J. Pattern Recognit Artif Intell.*, vol. 23, no. 04, pp. 687–719, Jun. 2009, doi: 10.1142/S0218001409007326.
[23] H. Hartono, O. S. Sitompul, T. Tulus, and E. B. Nababan, "Biased support vector machine and weighted-smote in handling class imbalance problem," *International Journal of Advances in Intelligent Informatics*, vol. 4, no. 1, pp. 21–27, Mar. 2018, doi: 10.26555/ijain.v4i1.146.
[24] M. W. Gardner and S. R. Dorling, "Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences," *Atmos. Environ.*, vol. 32, p. 2627, Jan. 1998, doi: 10.1016/S1352-2310(97)00447-0.
[25] S. Dhanabal and S. A. Chandramathi, "A Review of various k-Nearest Neighbor Query Processing Techniques," *Int. J. Comput. Appl.*, vol. 3, no. 7, Jan. 2011, Accessed: Dec. 08, 2020. [Online]. Available: http://dx.doi.org/.
[26] G. Orrù, W. Pettersson-Yeo, A. F. Marquand, G. Sartori, and A. Mechelli, "Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review," *Neurosci. Biobehav. Rev.*, vol. 36, no. 4, pp. 1140–1152, Apr. 2012, doi: 10.1016/j.neubiorev.2012.01.004.
[27] H. P. Bhavsar and M. H. Panchal, "A Review on Support Vector Machine for Data Classification," 2012, Accessed: Dec. 08, 2020. [Online]. Available: https://www.semanticscholar.org/paper/e1b4d814f106f33ea9939a85de8d9edba266ad1c.
[28] "Belgiu M, Drăguț L. Random forest in remote sensing: A review of applications and future directions[J]. ISPRS Journal of Photogrammetry and Remote Sensin - Google Search." https://www.google.com/search?q=Belgiu+M%2C+Dr%C4%83gu%C5%A3+L.+Random+forest+in+remote+sensing%3A+A+review+of+applications+and+future+directions%5BJ%5D.+ISPRS+Journal+of+Photogrammetry+and+Remote+Sensin&oq=Belgiu+M%2C+Dr%C4%83gu%C5%A3+L.+Random+forest+in+remote+sensing%3A+A+review+of+applications+and+future+directions%5BJ%5D.+ISPRS+Journal+of+Photogrammetry+and+Remote+Sensin&aqs=chrome..69i57.180j0j4&sourceid=chrome&ie=UTF-8 (accessed Dec. 08, 2020).
[29] R. Wang, "AdaBoost for Feature Selection, Classification and Its Relation with SVM, A Review," *Phys. Procedia*, vol. 25, pp. 800–807, Jan. 2012, doi: 10.1016/j.phpro.2012.03.160.
[30] K. M. Tsiouris, V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis, "A Long Short-Term Memory deep learning network for the prediction of epileptic seizures using EEG signals," *Computers in Biology and Medicine*, vol. 99. pp. 24–37, 2018, doi: 10.1016/j.compbiomed.2018.05.019.
[31] H. Khan, L. Marcuse, M. Fields, K. Swann, and B. Yener, "Focal Onset Seizure Prediction Using Convolutional Networks," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 9, pp. 2109–2118, Sep. 2018, doi: 10.1109/TBME.2017.2785401.
[32] N. D. Truong, L. Kuhlmann, M. R. Bonyadi, D. Querlioz, L. Zhou, and O. Kavehei, "Epileptic Seizure Forecasting With Generative Adversarial Networks," *IEEE Access*, vol. 7, pp. 143999–144009, 2019, doi: 10.1109/ACCESS.2019.2944691.
[33] L. A. Moctezuma and M. Molinas, "EEG Channel-Selection Method for Epileptic-Seizure Classification Based on Multi-Objective Optimization," *Front. Neurosci.*, vol. 14, p. 593, Jun. 2020, doi: 10.3389/fnins.2020.00593.