A Novel Approach for Seizure Classification Using Patient Specific Triggers: Pilot Study

Jamie Pordoy(✉), Ying Zhang, and Nasser Matoorian

School of Computing and Engineering, University of West London, London W5 5RF, UK
jamiepordoy@hotmail.com, {ying.zhang,nasser.matoorian}@uwl.ac.uk

Abstract. With advancements in personalised medicine, healthcare delivery systems have moved away from the one-size-fits-all approach towards tailored treatments that meet the needs of individuals and specific subgroups. As nearly one-third of those diagnosed with epilepsy are classed as refractory and are resistant to antiepileptic medication, there is a need for a personalised method of detecting epileptic seizures. Epidemiological studies show that up to 91% of those diagnosed identify one or more triggers as the causation of their seizure onset. These triggers are patient-specific and can affect those diagnosed in different ways dependent on each person's idiosyncratic tolerance and threshold levels. Whilst these triggers are known to induce seizure onset, only a few studies have even considered their use as a preventive component. Therefore, this pilot study investigates the use of patient-specific triggers (PST) in diagnosed epileptics, and whether they can be used as an additional modality when detecting seizures. This study used a precision medicine approach with artificial intelligence (AI), to train and test several patient-specific algorithms that classified epileptic seizures based on the PST of each participant. Experimental results show accuracy, sensitivity, and specificity scores of 94.73%, 96.90% and 93.33% for participant 1 and 96.87%, 96.96% and 96.77% for participant 2, respectively.

Keywords: Multi-modal · Machine learning · Epilepsy · Seizure · Patient-specific

1 Introduction

Epilepsy is a prevalent neurological condition that effects an estimated 70 million people worldwide [1]. An overload of electrical activity between communicating neurons causes a temporal imbalance of neurological activity, culminating in the occurrence of an unprovoked seizure, often leaving an individual with a loss of anatomical motor functions and clarity of memory [2]. An estimated 30% of those diagnosed are classed as refractory and are resistant to anti-epileptic drugs (AEDs) [3]. Those who are resistant have no form of defence and are at a higher risk of triggering a convulsive seizure which can lead to an acute cardiac and respiratory dysfunction [4].

A sudden unexpected death in epilepsy (SUDEP) is the most frequent direct cause of epilepsy-related deaths, predominately affecting those who are resistant or have poorly

H. Gao et al. (Eds.): CollaborateCom 2020, LNICST 350, pp. 455–468, 2021.
https://doi.org/10.1007/978-3-030-67540-0_29

controlled chronic epilepsy. A study by Lambert *et al.* [5] identified 58% of SUDEP cases are nocturnal and occur once an individual has been asleep and experienced a generalised tonic-clonic (GTC) seizure. As the underlying cause of SUDEP remains unknown and without treatment at a therapeutic level, recent case studies have suggested that onset, and in turn SUDEP, could be triggered by several predisposed risk and trigger factors [6].

As observed by Hesdorffer *et al.* [7], the most significant risk factor is an increase in the frequency of GTC seizures, as this can lead to a cardiac and respiratory dysfunction. Patients with epilepsy (PWE) who experience $\geq$3 GTC seizures per year are 15 times more likely to have a fatal epilepsy-related event such as SUDEP. Other frequent risk factors include partial seizures, missing doses of AEDs and an intelligence quotient (IQ) < 70 [8].

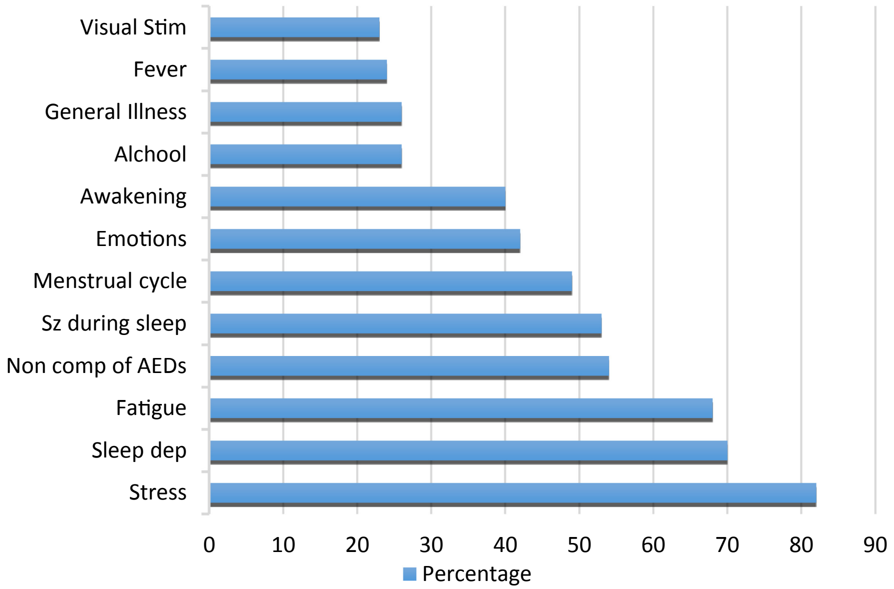

Fig. 1. Percentage of trigger factors observed in *n*-participants ($n = 104$ PWE)

In addition, it is important to ask whether there are specific triggers (precipitants) that increase the probability of onset and whether these could be used in conjunction with the aforementioned risk factors to improve on current methods of seizure detection [9]. Although seizures are sporadic and seemingly random, studies show there are patient-specific triggers (PST) that increase the likelihood of onset [10]. As defined by Aird and Gordon [11], PST can be categorised as seizure inducing or seizure triggering events. Those classed as seizure inducing (lights, noises, and patterns), are caused by environmental or endogenous events and cause a transient lowering of the seizure threshold level [12]. Seizure triggering events (sleep deprivation, stress, and fatigue) are risk factors that vary based on each person's specific threshold and tolerance levels. A study by Ferlisi and Shorvon [13] interviewed 104 PWE to identify the frequency of seizure precipitants (triggering factors).

Results show that seizure triggering events are more frequent, with an estimated 91% of participants experiencing one or more triggers prior to a seizure [13]. The distribution of these triggers is illustrated in Fig. 1, with stress, sleep deprivation, fatigue, and a non-compliance of AEDs as the four most common causes of seizure onset with percentage scores of 82%, 70%, 68% and 54% respectively.

These results reflect the findings of Nakken *et al.* [14], who also identified stress, sleep deprivation and fatigue as the most frequent triggers, with 592 (51%) PWE listing at least one trigger as the causation of their onset [14]. Furthermore, a study by Balamurugan *et al.* [15], analysed 405 PWE, observing that 86.9% experience at least one trigger prior to onset. These results show a noncompliance of AEDs as the most frequent trigger (40.98%), followed by stress (31.35%), sleep deprivation (19.75%) and fatigue (15.30%).

1.1 Patient-Specific Triggers

Precision medicine is a newly adapted paradigm of healthcare that allows medical treatments to be idiosyncratically tailored towards the needs of individuals or specific subgroups [16]. Whilst precision medicine in epilepsy is still relatively unexplored, a recent study by Porumb *et al.* [17], combined precision medicine and machine learning for hypoglycemic event detection from ECG wavelets. By applying a patient-specific approach, a classification model was trained using data from a single participant, which was then tested on unseen data from the same participant. The results demonstrate the potential application of patient-specific classification, with models attaining an accuracy measure of 84.8%, 88.5%, 89.9% and 78.3% for participants 1–4 respectively.

Similarly, a study by Ince *et al.* [18], has explored patient-specific classification of cardiac cycles to detect ventricular ectopic beats (VEBs) and supra-VEBs (SVEBs). Using fully connected feed-forward neural networks that were optimally designed for each persons' idiosyncrasies, this studies patient-specific classification models have surpassed many state of the art algorithms with accuracy and sensitivity scores of 98.3% – 84.6% and 97.4% – 63.5% for VEBs and SVEBS respectively.

Given the successful detection of hypoglycemic anomalies and the use of patient-specific classification in other medical fields, it is our feeling that this method of detection can be used to classify the PST observed in diagnosed epileptics prior to seizure onset.

2 Methodology

This paper presents a preliminary pilot study that investigates the practical application of PST when classifying epileptic seizures. We believe this investigation is supported by the notion that PST are preceding events that are responsible for initiating or precipitating a seizure [14]. Due to the impact of the current global crisis (Covid-19), we cannot conduct a full-scale clinical trial, which in turn has reduced the size of our dataset. However, as this is a pilot study that focuses on patient-specific classification, we decided to proceed using the participants we had available.

Classification models were developed using the Python programming language and coded in Jupyter notebook. The Python libraries used for the development include TensorFlow, Keras and Scikit-learn.

This pilot study will be used as a preliminary component to test the validity of PST and deem whether further research with a full-scale clinical trial is feasible. As far as we know, this is a novel concept and there is no existing research that has used these specific triggers for patient-specific classification of epileptic seizures.

2.1 Data Acquisition

For this study we collected data from two participants with epilepsy, one female (participant 1) in her late twenties and one male (participant 2) in his mid-thirties. Of the 300 days participant 1 was observed, we documented 17 positive instances (seizures), whilst participant 2 was observed for a shorter duration of 248 days, of which 22 positive instances were recorded.

Based on Ferlisi and Shorvon's findings [13], we decided to record the trigger factors of sleep deprivation, stress, and fatigue.

Whilst it has not been observed as a commonly occurring trigger, we have decided to record exercise as an additional PST, as we can see a clear correlation with stress reduction, which in turn could help to prevent or reduce the frequency of seizure onset [19].

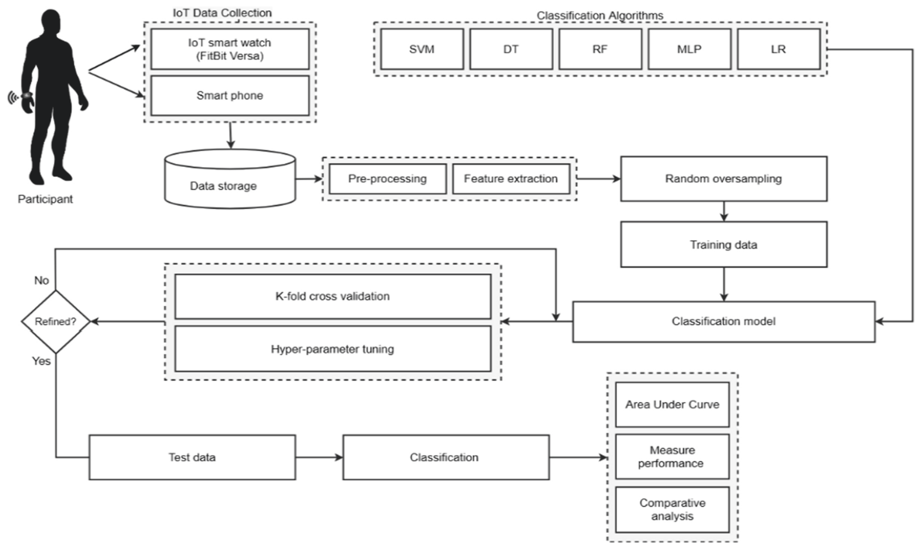

Fig. 2. Pilot study workflow; SVM = support vector machine; DT = decision tree; RF = random forest; MLP = multi-layer perceptron; LR = logistical regression; IoT = Internet of Things.

2.2 Pilot Study Workflow

Figure 2 shows a high-level diagrammatic representation of this study's methodology, which consists of four phases, data collection; data preparation; machine learning and finally a comparative analysis. Both participants were issued with an IoT-enabled smart watch (Fitbit Versa) and smart phone (Samsung SM-N950P) for the duration of the study. We deployed an android based CRUD application with pre-installed quantitative questionnaires on each smart phone, which was then used to record daily measures of stress and fatigue. We used a Fitbit Versa to calculate the metabolic equivalents (METs) of participants, as it provided an accurate measure for the number of minutes that active movement occurred [20]. We also used the embedded accelerometer and heart rate sensor to measure fluctuations in heart rate variability (HRV) and body movement to estimate the quality and duration of a participants sleep cycle [21].

The following sections briefly describe the components that were used to measure the PST and the machine learning models used for patient-specific classification.

2.3 Resampling the Dataset

Accurately classifying rare events can be problematic as the frequency of occurrence can leave a dataset highly imbalanced. For this study, 548 instances were recorded, 509 negative instances where no seizure was documented and 39 positive instances. To account for the varying disparity between positive and negative instances, we used random over-sampling. This sampling technique balanced our datasets distribution as duplicate instances from the minority class were added to our training dataset [22].

2.4 Perceived Stress Scale

Perceived stress scale (PSS) is a widely used stress assessment tool that measures a subject's stress levels [23]. Although the PSS is generally calculated at the end of each month, for this study, we used a modified variant that measures stress daily. The PSS is comprised of 10 questions that require a numerical response from 1 to 5, where $1 :=$ no stress and $5 :=$ maximum stress. To calculate the PSS score, the value responses for questions 4, 5, 7 and 8 are reversed so that question $1 \leftrightarrow 5$, $2 \leftrightarrow 4$, $3 \leftrightarrow 3$, $4 \leftrightarrow 2$, $5 \leftrightarrow 1$ [24].

The values for each question are than combined and divided by 10. The PSS scale also accounts for perception, as two participants who encounter the same set of events can accumulate different scores.

2.5 Rating of Fatigue

The Rating of Fatigue (RoF) is a measurement tool that tracks the intensity of fatigue. RoF uses a linear scale of 11 numerical intervals (0–10) with a response of 0 indicating no fatigue whilst a 10 indicates total fatigue and exhaustion. At the start of each day, the pre-installed RoF app shows a set of diagrammatic and descriptive components that guide participants when measuring fatigue [25].

2.6 Classification Algorithms and Techniques

This section provides a summary of the supervised learning classification models and techniques used for this study's comparative analysis. Classification is an instance of supervised learning, where a classification model (classifier) observes a set of input features and makes a prediction on unseen data that shares the same features. Classification models predict binomial outcomes such as yes/no, true/false and positive/negative to categorise new observations.

K-Fold Cross Validation

• K-fold cross validation segments a dataset into k-subsets of approximately equal size, and in turn each classification model is then trained using k-1 subsets where the remaining subset is used to validate the classifiers performance on unseen data [26]. This process in then repeated for k-iterations, and each iteration uses a different subset for validation, whilst the previous validation subset becomes a training subset of k-1. For this study, we used a k-fold cross validation technique where $k = 10$, to evaluate each classifiers performance on unseen data.

Naive Bayes

• Stemming from Bayes theorem of probability, naive Bayes (NB) classifiers are a family of probabilistic classification algorithms which assume that each predictor is an independent entity that equally contributes to the outcome of the target class [27]. NB classifiers can calculate the posterior probability of event A, given the occurrence of event B [28]. This can be expressed mathematically as

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \tag{1}$$

Where $P(c|x)$ represents the posterior probability, $P(c)$ the prior probability of class c, $P(x)$ the probability of each predictor and $P(x|c)$ the probability of a predictor given the occurrence of class c [29]. As NB assumes that features are independent, only the variances of each training label need to be calculated, instead of the entirety of the covariance matrix. This enables a NB classifier to use a small quantity of training data when predicting the mean and variance for each predictor, which is ideal for this studies patient-specific approach.

Support Vector Machines

• Support vector machines (SVM) are supervised learning models used to increase the predictive accuracy for classification and regression analysis. SVM classifies data by finding the optimum position for a hyperplane in n-dimensional space, where n represents the number of features, and each feature has a specific co-ordinate value. The features closest to the hyperplane act as support vectors and are used to determine the orientation of the hyperplane. This enables the hyperplane to separate n-classes of training data by the maximum distance, leading to maximal generalization and improved performance [30].

Logistical Regression

• Logistic Regression (LR) is a statistical, supervised learning model used to calculate binary and binomial response data. LR computes the probabilistic relationship between a dependent, dichotomous variable (0, 1) and one or more independent predictors (variables).

$$z = b_0 + b_1 \cdot x_1 + \cdots b_n \cdot x_n \tag{2}$$

A sigmoid function transforms a linear equation into a logistic equation that converts any input values between negative and positive infinity to a value between 0 and 1 [31]. The equation for LR is derived from a generalised linear equation of independent predictors where e represents the natural logarithm base, b and x are the classifiers parameters, and P is the probability of 1 [32].

$$p = \frac{1}{1 + e^{-(b_0 + b_1 \cdot x_1 + \cdots b_n \cdot x_n)}} \tag{3}$$

Decision Tree with Gini Index

• Decision Trees (DT) are hierarchical decision analysis structures that use a series of interconnected nodes to classify a decision and its consequences. A generalised tree structure consists of single root node, a series of decision nodes and leaf nodes. Decision nodes represent the consequences of an action and have two or more branches that connect to the leaf nodes, whereas the leaf nodes represent the final classification decision of that action. For this study we used the Gini Index to measure impurity, as it enables the most relevant decision nodes to be closer to the root node. As the tree structure traverses downwards, the level of uncertainty surrounding each decision decreases, ensuring a more accurate method of classification [33]. The Gini index is calculated using the following formula where the sum of the squared probabilities for each class is deducted from 1 [34].

$$Gini\ Index = 1 - \sum\nolimits_{i=1}^{c} (p_i)^2 \tag{4}$$

Random Forest

• Random Forest (RF) is an ensemble classifier that trains multiple decision trees in parallel for increased measures of performance [35]. RF classifiers use a combination of Breiman's "bagging" and random feature selection to form a process called majority voting where each tree classifies input data to identify the most frequently occurring class (prediction). This method of classification exhibits good generalisation, and often outperforms other classification models when measuring accuracy [36]. RF classifiers can be expressed mathematically as

$$\hat{f} = \frac{1}{B} = \sum\nolimits_{b=1}^{B} \hat{f}_b\left(x'\right) \tag{5}$$

Where $\hat{f}$ represents the final prediction, B the number of trees used, b the current tree and x the training sample used to teach the classifier.

Multi-layer Perceptron

Artificial neural networks (ANN) are information processing paradigms that share performance characteristics with the human biological nervous system [37].

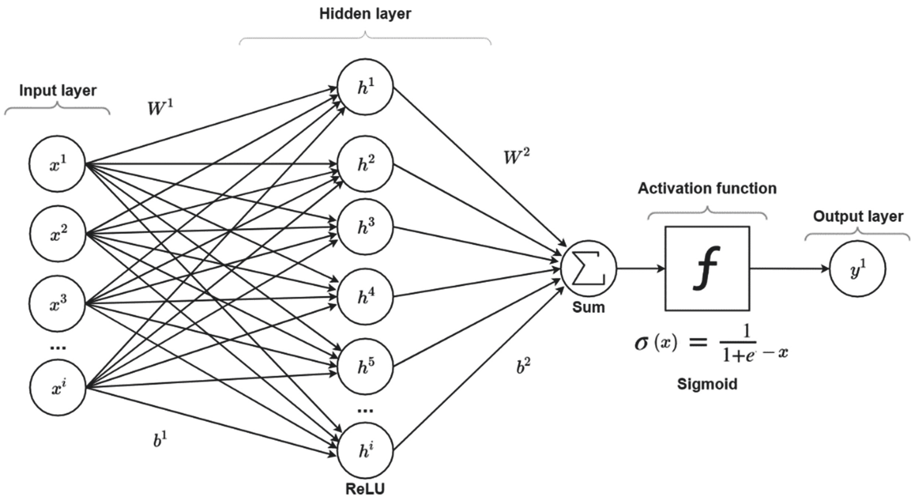

Fig. 3. Multi-layer perceptron architecture; x_i (Input layer) = 6 nodes; h_i (Hidden layer) = 12 nodes; y_i (Output layer) = 1 node; b_i = bias; w_i = weights; ReLU = rectified linear unit;

Multi-layer perceptron's (MLP) are a type of feed-forward neural network that analyses the relationship between a series of independent input variables and a set of dependent output variables. MLP are a modified variation of the standard two layered perceptron that uses three or more layers of neurons with nonlinear activation functions to process complex computations. The MLP used in this study is shown in Fig. 3 and can be expressed mathematically as

$$y_i = \int\left(\sum\nolimits_{i=1}^{n} w_i \cdot w_{ih} + \theta_\mathrm{h}\right) \tag{6}$$

Where y_i stands for the hidden layer, x_i represents node i of the input layer, w_i is the connection weight between node i of the input layer and node h of the hidden layer. The number of input layer nodes is expressed as n, the bias values of node h is θ_h and the network's Sigmoid function is expressed as $\int(u)$ [38].

3 Experiments

As defined by literature [39], the following four metrics are the optimum base measures used to assess the performance of binomial classification models. These base measures are calculated from a series of experiments using a set of positive and negative instances, where TP = true positive, FP = false positive, FN = false negative, TN = true negative. Accuracy (Acc) is used to measure the number of correctly classified instances by the total number of instances; sensitivity (Sen) measures the number of true positive instances by all positive instances; specificity (Spe) measures the number of true negative instances by the number of negative instances and the positive predicted value (Ppv) is the rate of positive instances that are positive. The mathematical formulas for each base measure are expressed below in algorithms 7–10.

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

$$Sen = \frac{TP}{TP + FN} \tag{8}$$

$$Spe = \frac{TN}{TN + FP} \tag{9}$$

$$Ppv = \frac{TP}{TP + FP} \tag{10}$$

To account for the imbalances in our dataset and the use of over-sampling, we constructed a multi-point receiver operating characteristic (ROC) curve of probability. A ROC curve plots a classifiers TP rate against its FP rate at multiple decision thresholds. We then measured the area under curve (AUC) to assess how efficient each classification model was at distinguishing between positive and negative predictions. To calculate AUC, we used the following formula [40]:

$$\hat{A} = \frac{S_0 - n_0(n + 1)/2}{n_0 \cdot n_1} \tag{11}$$

Where n_0 represents the number of positive instances, n_1 the number of negative instances, and $S_0 = \sum r_i$ where r_i is the rank of i th instance [41]. The following classification experiments were conducted on three separate datasets, D1 (participant 1), D2 (participant 2) and D3 (D1 + D2). For D3, we combined the data from both participants to see if it would improve the performance of our patient-specific classifiers (Fig. 4).

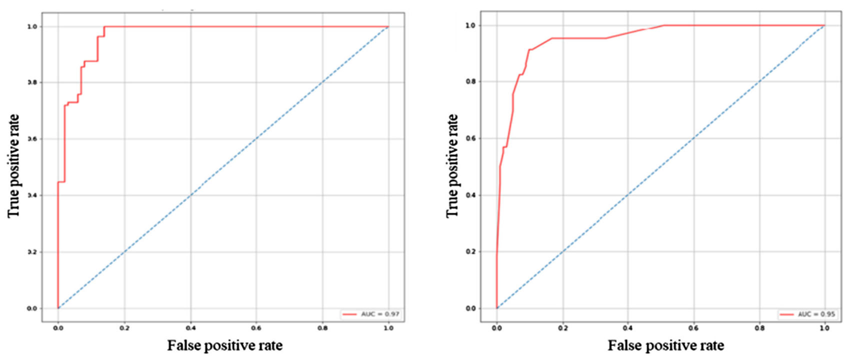

Fig. 4. ROC curve and AUC score for D1 (left) & D3 (Right) using a multi-layer perceptron

4 Experimental Results

The following section illustrates the experimental results recorded for this pilot study. Table 1 summarises the comparative analysis of our classification models for participant 1, with the MLP outperforming the other classifiers with an accuracy measure of 94.73%, sensitivity of 96.29% and an AUC measure of 0.970. Table 2 illustrates the classification results for participant 2, with the MLP once again outperforming the other classification models.

Table 1. Results of each classification model for participant 1

Model	Acc (%)	Sen (%)	Spe (%)	AUC	Ppv (%)
MLP	94.73	96.29	93.33	0.970	96.55
LR	94.70	95.00	94.44	0.947	95.55
NB	94.11	88.76	92.22	0.944	89.01
SVM	93.85	88.75	97.77	0.933	90.72
DT	92.94	91.25	94.44	0.931	92.39
RF	89.44	88.50	90.36	0.942	88.23

Table 2. Results of each classification model for participant 2

Model	Acc (%)	Sen (%)	Spe (%)	AUC	Ppv (%)
MLP	96.87	96.96	96.77	0.987	96.70
SVM	91.20	95.45	87.23	0.913	95.34
DT	91.17	96.87	86.11	0.915	96.87
LR	90.44	90.90	90.00	0.905	91.30
RF	86.75	82.14	91.25	0.961	87.95
NB	85.84	82.45	89.28	0.859	83.33

The following results show how each classifier performed when the datasets for both participants were combined. Once again, the MLP outperformed the remaining classifiers, with accuracy scores of 94.11%, sensitivity of 92.15% and an AUC measure of 0.952 (Table 3).

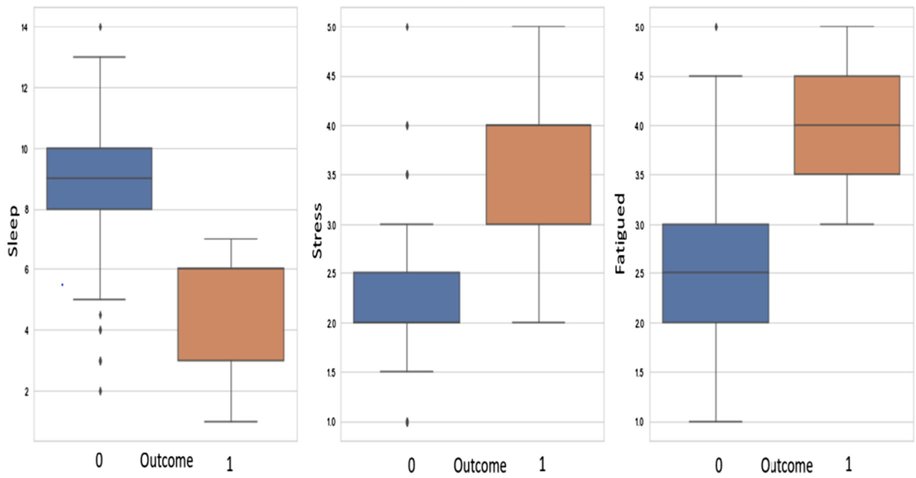

Fig. 5. Participant 1 PST; 0 = no seizure; 1 = seizure.

Table 3. Classification results for dataset D3 (participant 1 + participant 2)

Model	Acc (%)	Sen (%)	Spe (%)	AUC	Ppv (%)
MLP	94.11	92.15	96.07	0.952	92.45
LR	92.48	92.10	92.85	0.925	92.25
NB	92.15	90.12	94.44	0.923	89.47
SVM	91.66	91.17	92.15	0.917	91.26
DT	91.17	88.88	93.75	0.913	88.23
RF	89.54	90.80	88.23	0.958	90.60

5 Discussion

This section summarises the findings and contributions made from this pilot study. Experimental results indicate that PST can successfully train a classification algorithm using a single person's data, and then successfully classify the occurrence of seizure onset using the same persons unseen data. The machine learning approach used for this study successfully classified the fluctuations seen in PST prior to their onset.

For datasets D1, D2 and D3 the MLP outperformed the other classification models regarding Accuracy and AUC. Our findings support the notion that onset is influenced by idiosyncratic triggers as shown in Fig. 5 and Fig. 6 Although there have only been a few studies that assess the correlation between sleep and those diagnosed with epilepsy, our findings indicate seizure onset was more likely to occur when participant 1 had $\leq$ 6 h of sleep and participant 2 $\leq$ 8 h. Stress and fatigue also show a correlation with the frequency of onset, with participant 1 having a stress score of 3.5 or above in 88% of recorded seizures. A lower correlation was observed for participant 2, with 64% of seizures having a stress score of 3.5 or higher. Although participant 2 was less affected by stress, a higher fatigue score was observed throughout, with 77% of the 22 seizures recorded having a fatigue score of 4 or above. It was also observed that participant 2 had a stress score of 3.5 or above in 64% of total observations compared to participant 1's 48%.

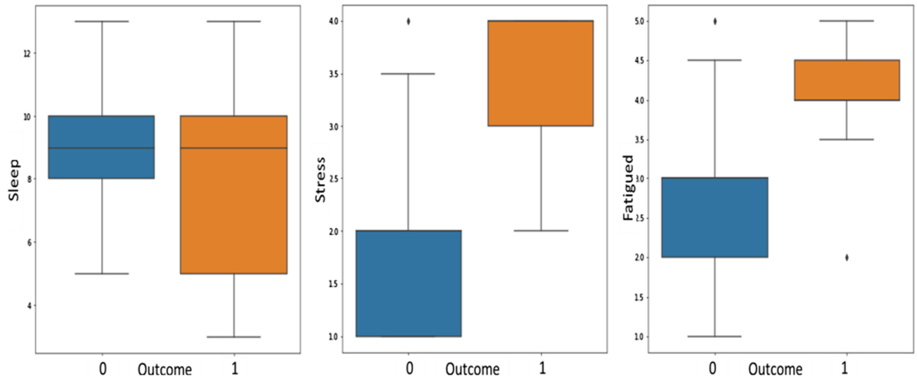

Fig. 6. Participant 2 PST; 0 = no seizure; 1 = seizure.

5.1 Limitations

One concern about the findings of this study was the sample size. Due to Covid-19, our sample size and participant availability was greatly affected, leaving this pilot study with 2 available participants. Whilst this is a preliminary pilot study that used patient-specific models for early detection of seizure onset, a larger sample size would further validate our initial hypothesis and the practical application of PST.

5.2 Future Research

The use of PST should be considered for future research as an additional sensing modality regarding non-EEG detection methods. Current multi-sensor modalities tend to focus on the use of biometric sensors such as electrocardiogram (ECG) to formulate predictions, as they allow for biometric fluctuations to be measured in real time. Using PST in conjunction with standard sensing modalities could account for the varying diversities seen in different types of epilepsy and reduce the frequency of false alarms.

6 Conclusion

This pilot study has undertaken a preliminary investigation into whether PST from the same participant can be used to train and test a classification model. Results indicate that both participants were more likely to experience a seizure if they had $<$ 6–8 h of sleep and/or a stress and fatigue factor $\geq$ 3.5. To our knowledge, this is the first pilot study that has proposed the use of PST for epilepsy detection, and the results presented here warrant further investigation in the form of a full-scale clinical trial.

References

1. Thijs, R.D., et al.: Epilepsy in adults. Lancet **393**(10172), 689–701 (2019)
2. Iasemidis, L.D., et al.: Adaptive epileptic seizure prediction system. IEEE Trans. Biomed. Eng. **50**(5), 616–627 (2003)
3. Laxer, K.D., et al.: The consequences of refractory epilepsy and its treatment. Epilepsy Behav. **37**, 59–70 (2014)
4. Barot, N., Nei, M.: Autonomic aspects of sudden unexpected death in epilepsy (SUDEP). Clin. Auton. Res. **29**(2), 151–160 (2018). https://doi.org/10.1007/s10286-018-0576-1
5. Lamberts, R.J., et al.: Sudden unexpected death in epilepsy: people with nocturnal seizures may be at highest risk. Epilepsia **53**(2), 253–257 (2012)
6. DeGiorgio, C.M., et al.: Ranking the leading risk factors for sudden unexpected death in epilepsy. Front. Neurol. **8**, 473 (2017)
7. Hesdorffer, D.C., et al.: Combined analysis of risk factors for SUDEP. Epilepsia **52**(6), 1150–1159 (2011)
8. Van de Vel, A., et al.: Non-EEG seizure detection systems and potential SUDEP prevention: state of the art: review and update. Seizure **41**, 141–153 (2016)
9. Pack, A.M.: SUDEP: what are the risk factors? Do seizures or antiepileptic drugs contribute to an increased risk? Epilepsy Curr. **12**(4), 131–132 (2012)
10. Manford, M., et al.: An analysis of clinical seizure patterns and their localizing value in frontal and temporal lobe epilepsies. Brain **119**(1), 17–40 (1996)
11. Aird, R.B., Gordon, N.S.: Some excitatory and inhibitory factors involved in the epileptic state. Brain Dev. **15**(4), 299–304 (1993)
12. Smith, S.J.M.: EEG in the diagnosis, classification, and management of patients with epilepsy. J. Neurol. Neurosurg. Psychiatry **76**(2), ii2–ii7 (2005)
13. Ferlisi, M., Shorvon, S.: Seizure precipitants (triggering factors) in patients with epilepsy. Epilepsy Behav. **33**, 101–105 (2014)
14. Nakken, K.O., et al.: Which seizure-precipitating factors do patients with epilepsy most frequently report? Epilepsy Behav. **6**(1), 85–90 (2005)
15. Balamurugan, E., et al.: Perceived trigger factors of seizures in persons with epilepsy. Seizure **22**(9), 743–747 (2013)
16. Ginsburg, G.S., Phillips, K.A.: Precision medicine: from science to value. Health Aff. **37**(5), 694–701 (2018)
17. Porumb, M., et al.: Precision medicine and artificial intelligence: a pilot study on deep learning for hypoglycemic events detection based on ECG. Sci. Rep. **10**(1), 170 (2020)
18. Ince, T., et al.: A generic and robust system for automated patient-specific classification of ECG signals. IEEE Trans. Biomed. Eng. **56**(5), 1415–1426 (2009)
19. Pimentel, J., et al.: Epilepsy and physical exercise. Seizure **25**, 87–94 (2015)
20. Fitbit.: What are active minutes? https://help.fitbit.com/articles/en_US/Help_article/1379. Accessed 2 May 2020
21. Fitbit.: What should I know about sleep stages? https://help.fitbit.com/articles/en_US/Helparticle/2163. Accessed 21 May 2020
22. Xu, X., et al.: Over-sampling algorithm for imbalanced data classification. J. Syst. Eng. Electron. **30**(6), 1182–1191 (2019)
23. Cohen, S., Janicki-Deverts, D.: Who's stressed? Distributions of psychological stress in the United States in probability samples from 1983, 2006, and 2009. J. Appl. Soc. Psychol. **42**(6), 1320–1334 (2012)
24. Cohen, S., et al.: A global measure of perceived stress. J. Health Soc. Behav. **24**(4), 385–396 (1983)

25. Micklewright, D., et al.: Development and validity of the rating-of-fatigue scale. Sports Med. **47**(11), 2375–2393 (2017). https://doi.org/10.1007/s40279-017-0711-5
26. Refaeilzadeh, P., et al.: Cross-validation. In: Liu, L., Özsu, M.T. (eds.) Encyclopedia of Database Systems. Springer, Boston
27. Yusa, M., Utami, E.: Classifiers evaluation: comparison of performance classifiers based on tuples amount. In: 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), pp. 1–7. IEEE, Yogyakarta, Indonesia (2017)
28. Zhang, H., Li, D.: Naïve Bayes text classifier. In: IEEE International Conference on Granular Computing, pp. 708–708. IEEE, Fremont (2007)
29. Jing, N., et al.: Information credibility evaluation in online professional social network using tree augmented naïve Bayes classifier. Electron. Commer. Res. 1–25 (2019). https://doi.org/10.1007/s10660-019-09387-y
30. Mavroforakis, M.E.: Theodoridis, S.: Support vector machine (SVM) classification through geometry. In: 13th European Signal Processing Conference, pp. 1–4. IEEE, Antalya (2005)
31. Lammertyn, J., et al.: Logistic regression analysis of factors influencing core breakdown in 'Conference' pears. Postharvest Biol. Technol. **20**(1), 25–37 (2000)
32. Sun, Y., et al.: Application of logistic regression with fixed memory step gradient descent method in multi-class classification problem. In: 6th International Conference on Systems and Informatics (ICSAI), pp. 516–521. IEEE, Shanghai (2019)
33. Tahsildar, S.: Gini Index for Decision Trees. https://blog.quantinsti.com/gini-index/. Accessed 09 July 2020
34. Decision Tree Flavors: Gini Index and Information Gain. https://www.learnbymarketing.com/481/decision-tree-flavors-gini-info-gain/. Accessed 05 May 2020
35. Misra, S., Li, H., He, J.: Machine Learning for Subsurface Characterization, 1st edn. Elsevier, Cambridge (2020)
36. Sinnott, R.O., Duan, H., Sun, Y.: Chapter 15 - a case study in big data analytics: exploring twitter sentiment analysis and the weather. In: Buyya, R., Calheiros, R.N., Dastjerdi, A.V. (eds.) Big Data: Principles and Paradigms, 1st edn. Elsevier, Burlington (2016)
37. Balaji, S.A., Baskaran, K.: Design and development of artificial neural networking (ANN) system using sigmoid activation function to predict annual rice production in Tamilnadu. Int. J. Comput. Sci. Eng. Inf. Technol. **3**(1), 13–31 (2013)
38. Matarat, K., et al.: Comparison of classification algorithms for movie reviews. In: 16th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), pp. 826–829. IEEE, Pattaya (2019)
39. Canbek, G., et al.: Binary classification performance measures/metrics: a comprehensive visualized roadmap to gain new insights. In: 2nd International Conference on Computer Science and Engineering (UBMK), pp. 821–826. IEEE, Antalya (2017)
40. Hand, D.J., Till, R.J.: A simple generalisation of the area under the ROC curve for multiple class classification problems. Mach. Learn. **45**, 171–186 (2001)
41. Huang, J., Ling, C.X.: Using AUC and accuracy in evaluating learning algorithms. IEEE Trans. Knowl. Data Eng. **17**(3), 299–310 (2005)