Online Seizure Prediction Using An Adaptive Learning Approach

Shouyi Wang, Wanpracha Art Chaovalitwongse, Stephen Wong

Abstract—Epilepsy is one of the most common neurological disorders, characterized by recurrent seizures. Being able to predict impending seizures could greatly improve the life of patients with epilepsy. In this study, we propose a new adaptive learning approach for online seizure prediction based on analysis of electroencephalogram (EEG) recordings. For each individual patient, we construct baseline patterns of normal and pre-seizure EEG samples, continuously monitor sliding windows of EEG recordings, and classify each window to normal or pre-seizure using a K-nearest-neighbor (KNN) method. A new reinforcement learning algorithm is proposed to continuously update both normal and pre-seizure baseline patterns based on the feedback from prediction result of each window. The proposed approach was evaluated on EEG data from 10 patients with epilepsy. For each one of the 10 patients, the adaptive approach was trained using the recordings containing the first half of seizure occurrences, and tested *prospectively* on the subsequent recordings. Using a 150-minute prediction horizon, our approach achieved 73% sensitivity and 67% specificity on average over 10 patients. This result is shown to be far better than those of a non-update prediction scheme and two native prediction schemes.

Index Terms—adaptive online seizure prediction, reinforcement learning, time series pattern recognition

1 Introduction

Epilepsy is one of the most common neurological disorders, affecting approximately 1% of the world's population [15]. Epileptic seizure onset is often considered as an abrupt, unpredictable phenomenon. The unpredictability of seizures represents a significant source of morbidity in patients with epilepsy. Patients with epilepsy frequently suffer from seizure-related injuries due to loss of motor control, loss of consciousness or delayed reactivity during seizures [37]. Current technology has yet to reach a point where epileptic patients can be warned by an automated system to predict seizure onsets. One crucial question in seizure prediction is whether an identifiable, specific, pre-seizure state exists. Over the recent years, there has been accumulating evidence indicating that a transitional pre-seizure state does exist prior to seizure onsets [20], [29], [41], [35], [30], [5], [7]. The majority of the quantitative evidence supporting the existence of a pre-seizure state is derived from EEG analyses. In the literature, seizure prediction algorithms are generally designed to capture some specific EEG features to analyze precursors of imminent epileptic seizures. Examples of published features include dynamical entrainment [23], [17], correlation dimension [28], dynamic similarity index [41], accumulated energy [32], phase synchronization [34], wavelet and median filtering [36]. Iasemidis et al. [20] noted premonitory pre-seizure changes based on the analysis of dynamical entrainment. Lehnertz and Elger [29] showed that the correlation dimension decreases prior to seizures. Le van Quyen et al. [41] reported a reduction in the dynamical similarity index before seizure occurrence. Mormann et al. [35] observed that there was a relative decrease of signal power in the delta band of the EEG up to hours prior to seizure onsets. They also demonstrated statistically significant discrimination between pre-seizure and normal brain states. In our previous study, Chaovalitwongse et al. [5] investigated the EEG characteristics of pre-seizure transition and found that the probability of detecting pre-seizure transition was as high as 83% using the optimized critical EEG channels. In a later study, we built a network-based approach to study the evolution of epileptic seizures by investigating the EEG synchronization among different brain areas. The evolutional changes of the network structure hours prior to seizure onsets indicated that the seizures may slowly develop by an evolutional epileptogenic process instead of an abrupt change [7]. Recently, Feldwisch-Drentrup et al. [16] investigated the possibility of combining different seizure prediction algorithms and different EEG features to improve prediction accuracy. Using Boolean operations, they showed the different prediction methods with different EEG features can be combined and can generate significant better performance than each individual method. In particular, they found that sensitivity can be markedly improved by combining dynamic similarity index [41] and phase synchronization [34], given a fixed maximum false

Shouyi Wang is with the Department of Industrial & Manufacutring Systems Engineering, University of Texas at Arlington, Arlington, TX 76019, USA (E-mail: shouyiw@uta.edu);
Wanpracha Chaovalitwongse is with the Department of Industrial & Systems Engineering and the Department of Radiology, University of Washington, Seattle, WA 98104, USA. (E-mail: artchao@uw.edu);
Stephen Wong is with Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, New Brunswick, NJ 08901, USA (E-mail: wongst@umdnj.edu).

prediction rate (FPR).

Although there have been extensive studies that develop the state-of-the-art seizure prediction algorithms to support the existence of the pre-seizure state [23], [28], [41], [32], [10], [34], [36], prospective seizure prediction from EEG remains a challenging problem. Computational analyses of most studies in the literature are focused on retrospective analyses of EEG recordings, which in turn only address the predictability of epileptic seizures rather than the prediction. A significant challenge of seizure prediction is the high intra-individual variability of epileptic seizures with a variable degree of success [22]. Although many nonadaptive methods for retrospective analyses have achieved promising results, this variability makes it difficult to develop a universal robust predictor to accurately predict seizures for a wide range of patients with different seizures. This variability also highlights the emerging need for an automated adaptive approach for epileptic seizure prediction. A number of adaptive seizure prediction algorithms have been proposed to account for the high inter- and intra-individual variability of epileptic seizures [22], [23], [45], [42], [8]. Iasemidis et al. [22], [23] and Sackellares et al. [45] developed optimization-based prediction algorithms which, based on dynamical synchronization in the human epileptic brain, adaptively selects a group of critical EEG electrodes to predict impending seizures. More recently, Iasemidis's group published similar results, with high sensitivity and specificity, and long warning times prior to seizures on prospective seizure prediction in rodents with chronic epilepsy [17]. Rajdev et al. [42] also proposed an adaptive prediction algorithm based on a Wiener implementation of autoregressive (AR) modeling, which was tested on rats. A warning was issued if the prediction errors over a moving window exceeded a threshold. The threshold was continuously updated online, and it was optimized to maximize the sensitivity and latency, while minimizing the FPR.

The above-mentioned adaptive seizure prediction approaches are generally based on an adaptively-optimized set of EEG channels [22], [23], [45] or an adaptive threshold [42]. In principle, these approaches employed the prediction settings optimized by one or several recently occurred seizures to predict the next seizure. Due to the high intra-individual variability of epileptic seizures, the characteristics of the EEG patterns of the next seizure may become quite different from those of its preceding ones. The current adaptive approaches actually do not make full use of the whole monitored EEG recordings, and thus have problems to deal with the challenging problems of high intra-individual variability of seizures in prediction. Therefore, it is extremely important for a prediction system to accumulate more and more knowledge of predictive patterns over time instead of only holding 'short-term memories'. In this study, we develop an automated adaptive learning approach for online seizure prediction. The approach is based on quantitative EEG analysis, time series classification, and reinforcement learning. For each individual patient, after the first seizure in the EEG recording, the approach will construct baseline EEG patterns from normal period and pre-seizure period. Then our approach continuously monitors and classifies sliding windows of EEG recordings as normal or pre-seizure based on a K-nearest-neighbor (KNN) rule to classify each EEG window to the most similar baseline patterns. Our approach in turn uses a gradient-based reinforcement learning algorithm to continuously update both normal and pre-seizure baseline patterns based on the feedback of true or false prediction. This study is among the first to investigate the use of adaptive learning in seizure prediction [22], [45], [18], [42]. Its framework can be applied to other online monitoring problems such as network intrusion detection and production process control.

The rest of this paper is organized as follows. In section 2, the background and previous related work are discussed. The data collection, feature extraction, the adaptive seizure prediction approach, and the evaluation metrics of prediction performance are presented in section 3. The experimental results are provided and discussed in Section 4, and we conclude the paper in Section 5.

2 Background and Related Works

2.1 Overview of Machine Learning Techniques

With the explosion of computing power in the past decade, machine learning and pattern recognition techniques have become important tools in the analysis of various biological problems, such as cancer research [31], cognitive neuroscience [12], and genomics and proteomics [9]. Machine learning best depicts the computational methods that allow a system to evolve behaviors through an automated process of knowledge acquisition from empirical data. Machine learning techniques generally fall into three broad categories: supervised learning, reinforcement learning and unsupervised learning. A supervised learning technique usually first finds a mapping between inputs and outputs of a training dataset, and then makes predictions for inputs that it has never seen. A large number of supervised learning algorithms have been developed that can be categorized into several major groups, including neural networks, support vector machines, locally weighted learning, decision trees, and Bayesian inference [26]. Reinforcement learning is another learning paradigm in which an agent is able to learn a decision policy by 'trial and error'. A reinforcement learner receives feedback of its actions and makes adjustments to its actions accordingly [50]. Reinforcement learning is a natural framework for building models to accumulate knowledge from

previously learned tasks to new tasks with increasing complexity and variability. Reinforcement learning techniques have been applied to many complex learning tasks, such as robot control [14] and traffic network control [44]. Unsupervised learning is inspired by the brain's ability to recognize complex patterns of visual scenes, sounds or odors. It takes root in neuroscience/psychology and is established on the basis of information theory and statistics. An unsupervised learner usually performs clustering or associative rule learning to extract the implicit structure of a given dataset. The established clusters, categories or associative networks are then used for decision making, prediction, or efficient communication [13].

2.2 EEG Analysis for Epileptic Seizures

Most seizure prediction methods are based on quantitative analysis of the EEG, and can be broadly categorized into univariate and multivariate analysis, respectively.

Univariate analyses focus on the features of each single channel of EEG. Based on the morphological characteristics of EEG, Lange et al. [27] reported that there were consistent changes in EEG spike activity prior to seizures. With the help of advanced signal processing methods, more complex univariate EEG feature extraction techniques have been developed for seizure prediction. Litt et al. [32] introduced signal energy variations to seizure prediction, and reported EEG changes hours before seizure onsets. Autoregressive (AR) and autoregressive moving average (ARMA) models have also been utilized for seizure prediction. Characteristic changes of AR/ARMA coefficients before seizure onsets were reported in [46], [8]. Nonlinear measures based on chaos theory have drawn considerable attention in EEG studies of brain activity. The two well-known nonlinear chaotic measures that have been applied in seizure prediction are the Lyapunov exponent and correlation dimension. Iasemidis et al. [22] monitored the evolution of Lyapunov exponents extracted from EEG data. They designed an adaptive prediction scheme that attempted to select the most informative channels to predict an impending seizure with optimization techniques. Channel selection was adjusted after every seizure since it was assumed that the pre-seizure dynamics may change from seizure to seizure over time. Lehnertz et al. [28] investigated the feasibility of seizure prediction based on transitions of correlation dimension, a feature that is considered as an index of neuronal complexity.

Multivariate analyses take more than one channel of EEG into account simultaneously rather than only looking at each channel individually. The most influential multivariate analysis methods in seizure prediction are phase synchronization and dynamical entrainment. Le Van Quyen et al. [40] used phase synchronization to distinguish pre-seizure features from normal state. They compared the normal synchronization patterns taken from 3-10 hours before seizures with the pre-seizure patterns taken from 30 minutes before seizures. The variables that achieved best discriminating performance were chosen for each individual patient. Mormann et al. [34] designed a seizure prediction scheme based on their finding that the degree of synchronization may decrease up to hours prior to seizure onsets. Iasemidis et al. [23] explored the effectiveness of a method called dynamical entrainment, which estimated the difference of the largest Lyapunov exponents from any two observed time series of EEG. A progressive convergence of the dynamical entrainment was considered as sign of transition from normal to pre-seizure states.

Our group has made extensive EEG studies to investigate the classifiability of the brain's pre-seizure and normal states [6], [2], [3], [4]. Our classification model achieved a a testing accuracy of over 70% on average. The experimental results indicate that it may be possible to design and develop seizure warning algorithms for diagnostic and therapeutic purposes.

2.3 Related Work in Seizure Prediction and Challenges

In the 1970s, accumulating evidence from clinical practice suggested that epileptic seizures might be predicable. Viglione and Walsh started a project to investigate the predictability of seizures based on EEG data [52]. Iasemidis et al. pioneering work started in the 1980s [24], [25], [21]. Since then, many studies have been carried out aiming to predict epileptic seizures.

Most current seizure prediction methods involve two steps. First, univariate or multivariate EEG features are extracted from a sliding window. Then each EEG epoch in the moving window is classified as either pre-seizure or normal based on an optimized threshold level. Whenever a windowed EEG epoch is classified as pre-seizure, a warning alarm is triggered indicating that an impending seizure may occur within a pre-defined prediction horizon. Although some methods have shown promising results for selected patients, the reliability and repeatability of the results have been questioned when tested on other EEG datasets. Many of the earlier optimistic findings were irreproducible or achieved poor performance in extended EEG datasets [1]. This is not surprising since the optimal threshold obtained from a limited number of patients may not be generalizable. Manually tuning a threshold level for each individual patient is a subjective procedure and would pose a significant burden on physicians and patients. The inability to apply these techniques to a wide spectrum of epileptic patients with a variety of types of epileptic seizures may represent the greatest limitation of current seizure prediction methods.

Given our accumulated knowledge regarding seizure prediction, we conjecture that a promising

approach may be the one that processes adaptive learning ability and is capable of achieving personalized seizure prediction autonomously. The flowchart of a prospective adaptive seizure prediction system is illustrated in Figure 1. In this study, we attempted to construct an adaptive prediction system using machine learning algorithms. We developed a novel adaptive learning approach, which combines reinforcement learning, online monitoring, and feedback control theory into an online seizure prediction system. The proposed adaptive seizure prediction approach can be readily integrated to any clinical EEG system. With the attractive adaptive learning ability, the proposed approach is capable of achieving a personalized seizure prediction through baseline-updating as it monitors more and more EEG recordings from a patient.

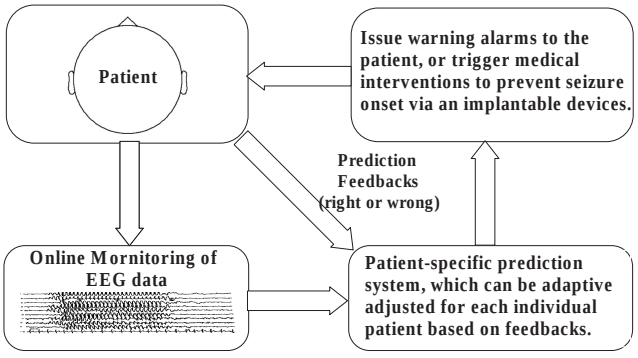


Fig. 1. A prospective adaptive seizure prediction system, which can be adjusted to each individual patient automatically based on feedbacks.

3 Materials and Methods

3.1 Data Collection

In this study, we used a dataset containing long-term continuous intracranial EEG recordings from 10 epileptic patients with temporal lobe epilepsy. The placement of the EEG electrodes is shown in Figure 2, which is a modified image of the inferior transverse view of the brain from Potter [38]. The EEG recordings consist of 26 standard channels. Recording durations ranged from 3 to 13 days. Expert epileptologists annotated the EEG recordings to determine the number of seizures, their onset, and their offset points. The characteristics of the 10 patients and the EEG data statistics are outlined in Table 1.

3.2 Data Preprocessing & Feature Extraction

Since EEG signals are highly nonstationary and seemingly chaotic, there has been an increasing interest in analyzing EEG signals in the context of chaos theory [43]. Several commonly used chaotic measures in many recent studies include largest Lyapunov exponent [22], correlation dimension [48], Hurst exponent [11] and entropy [39]. Among these EEG measures, the Lyapunov exponent has been shown to be useful in characterizing a chaotic system [51]. Lyapunov

TABLE 1
Characteristics of EEG data

Patient	Gender /Age	Number of Seizures	EEG Length (hour)	Average Inter-seizure Interval (hour)	Seizure Type
1	F/45	7	85.18	12.17	CP, SC
2	M/60	7	280.86	40.12	CP, GTC, SC
3	F/41	24	212.28	8.85	CP
4	M/19	17	315.23	18.54	CP, SC
5	M/33	17	286.76	16.87	CP, SC
6	M/38	9	74.60	8.29	CP, SC
7	M/44	23	146.15	6.35	CP, SC
8	M/29	19	142.32	7.49	CP, SC
9	F/37	20	276.65	13.83	CP, SC
10	M/37	12	231.61	19.30	CP, GTC
Total		155	2051.63		

Seizure types: CP, complex partial; SC, subclinical; GTC, generalized tonic/clonic.

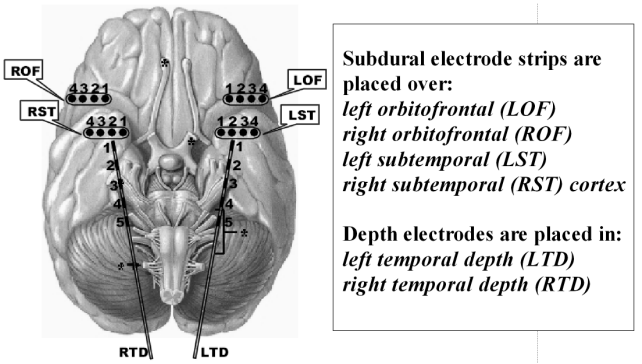


Fig. 2. The interior transverse view of the brain and the placement of the 26 EEG electrodes.

exponents measure the degree of sensitivity to initial conditions for a dynamical system. For an n-dimensional dynamical system, there will be n corresponding Lyapunov exponents that measure the exponential rate of divergence of the different trajectories in the phase space. If an exponent is positive, it indicates that the corresponding orbits locally defined by that exponent diverge exponentially. The magnitude of the exponents indicates the degree of divergence. The largest Lyapunov exponent in a chaotic system is usually more reliable and reproducible than the estimation of all the exponents [51], and is an important indicator to characterize a chaotic system. In our previous studies, we used an estimation algorithm called the short-term largest Lyapunov exponent (STL_{max}) to quantify EEG dynamics [22]. We employed this measure in the current study. A detailed calculation of STL_{max} as well as parameter selection and variation of STL_{max} has been explained by Iasemidis in [19].

3.3 Adaptive Seizure Prediction Approach

The schematic structure of the proposed adaptive seizure prediction system is illustrated in Figure 3. A sliding window was applied to monitor continuous multichannel EEG data. The window size is 10 minutes with 50% overlap between two successive windows. Two baselines of normal and pre-seizure states were constructed and initialized by the beginning part of the EEG recordings for each patient. The two baselines were used to classify the monitored EEG epochs

of the sliding moving window using a K-nearest-neighbor (KNN) method. All the baseline samples and windowed EEG epochs were represented in terms of the multichannel time profile of STL_{max} values. The two baselines were updated by a reinforcement learning algorithm based on feedbacks of prediction actions (true or false). The adaptive seizure prediction system is discussed in detail in the following.

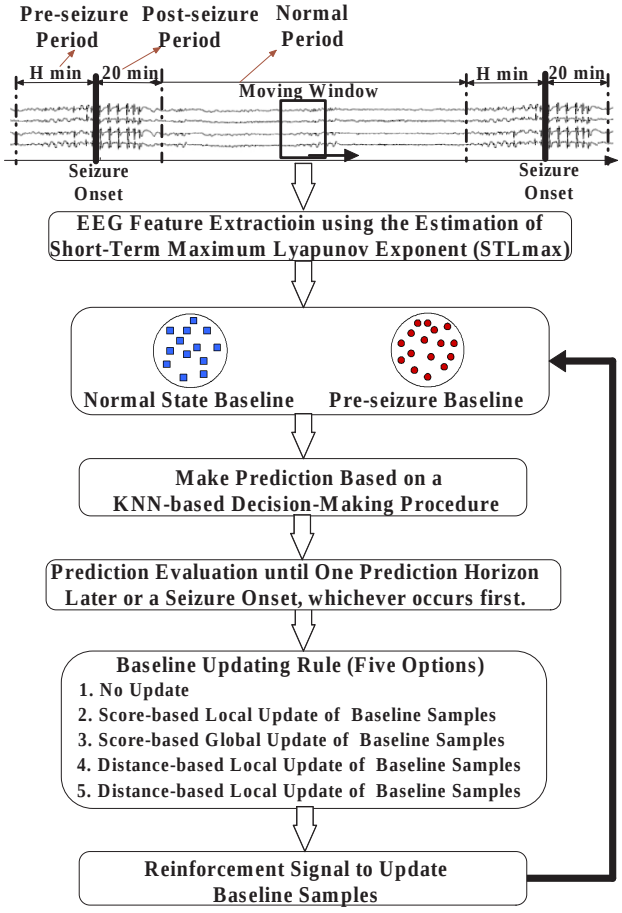


Fig. 3. Schematic structure of the adaptive prediction system.

3.3.1 Baseline Construction & Initialization

To start our prediction system, we first initialize the pre-seizure and normal baseline samples. The selection of baseline samples depends on the presumed time length of pre-seizure period, which is often considered the prediction horizon in the seizure prediction literature. The pre-seizure duration has been reported to be between a few minutes and several hours prior to seizure onset, and remains an open question in epilepsy research. In this study, we tried three prediction horizons (30, 90, and 150 minutes). For convenience, we denote the length of the prediction horizon as H minutes, then the EEG recordings can be divided into the following three periods:

- Pre-seizure period: 0-H minutes preceding a seizure onset.
- Post-seizure period: 0-20 minutes after a seizure onset.
- Normal period: between pre- and post-seizure periods.

The initial samples of the two baselines were randomly chosen from the normal and pre-seizure period preceding the first seizure onset. The length of the baseline samples is equal to that of the moving window. Since there are no guidelines available to determine the number of samples in each baseline, we tentatively stored a fixed number of 50 samples in each baseline.

3.3.2 KNN Prediction Procedure

With baselines for normal and pre-seizure states, it is intuitive to classify a windowed EEG epoch based on its degree of similarity to the two baselines. For this purpose, KNN is a reasonable choice because it classifies a new unlabeled sample by comparing the sample with all the samples in the two baseline sets. For each EEG epoch in the moving window, the KNN method finds its K nearest (best matching) samples in each baseline, and compares the its averaged distances to the two groups of K-nearest neighbors. The epoch is classified to a baseline that is 'closer' to it. The KNN prediction procedure is described in the following.

KNN methods use similarity measures to quantify the closeness between a moving-window EEG and baseline samples. We employed three frequently-used time-series similarity measures. If we denote two time-series of STL_{max} as X and Y with equal length of n, then the three types of distances are briefly described as follows.

- Euclidean distance (EU): measures the degree of similarity in terms of amplitude of the data. The EU between X and Y is defined as $ED_{xy} = \sqrt{\sum_{p=1}^{n}(x_p - y_p)^2}$.
- T-statistical distance (TS): a statistical distance measure between two time series derived from the t-test. It is frequently used to determine if the mean values of two time series differ from each other in a significant way under the assumptions that the paired differences are independent and identically normally distributed. The TS between X and Y is calculated by $TS_{xy} = \sum_{p=1}^{n}|x_p - y_p|/\sqrt{n}\tau_{|X-Y|}$, where $\tau_{|X-Y|}$ is the sample standard deviation of the absolute difference between the time series X and Y.
- Dynamic time warping (DTW): DTW measures similarity based on the best possible alignment or the minimum mapping distance between two time series. The two time series are 'warped' in the time domain to find the optimal pattern matching between them. DTW is particularly suited to matching time series patterns independent of time variations. A detailed calculation of DTW can be found in [47].

Once a similarity measure is chosen, we can obtain the distance between a baseline sample and an EEG epoch in the moving window. For a multichannel EEG epoch, the window-sample distance is calculated as follows:

$$d_{pre,i} = \sum_{j=1}^{M} distance(S^j_{pre,i}, S^j_{mw}) \quad (1)$$

$$d_{int,i} = \sum_{j=1}^{M} distance(S^j_{int,i}, S^j_{mw}) \quad (2)$$

where M=26 is the number of EEG channels. $S^j_{pre,i}$ and $S^j_{int,i}$ is the jth channel of EEG time series in the ith pre-seizure and normal baseline sample, respectively; $S^j_{mv,i}$ is the jth channel of EEG in the windowed EEG epoch; $d_{pre,i}$ and $d_{int,i}$ are the distances between the windowed EEG and the ith sample in the pre-seizure and normal baseline, respectively. The term $distance$ in the above formula represents a time series distance measure, which denotes EU, TS, or DTW in this paper.

We used four choices of K. They were three, seven, half, and all of the baseline samples, respectively. Once K is fixed, the weighted summation of K smallest window-sample distances in a baseline was considered as the distance between the windowed EEG epoch and that baseline. Therefore, we call the two distances as window-normal distance D^K_{int} and window-preseizure distance D^K_{pre}, respectively. For each windowed EEG epoch, its distances to the two baselines can be calculated by $D^K_{pre} = \sum_{k=1}^{K} \alpha_k d_{pre,k}$ and $D^K_{int} = \sum_{k=1}^{K} \beta_k d_{int,k}$. The α_k and β_k are the weights of the kth pre-seizure and normal baseline, respectively. The $d_{pre,k}$ and $d_{int,k}$ are the distances between the windowed EEG epoch and its kth nearest neighbor in the pre-seizure and normal baseline, respectively. Once the two baseline-window distances are obtained, the prediction decision can be made by:

$$predictor = \begin{cases} 1, & \text{if } D^K_{pre}/D^K_{int} \le R^* \text{ (issue an alarm)} \\ 0, & \text{otherwise (no warning);} \end{cases}$$

where the threshold R^* can be used to control the sensitivity of the prediction system. In this study, we employed $R^* = 0.99$ to make the prediction less sensitive to noises which would lead to many false predictions. Note that the impact of this threshold is also investigated in this study.

3.3.3 Evaluation of a Prediction Result

Baseline updating depends on prediction evaluation feedback. We define the evaluation metrics of each prediction outcome by the following. If the predefined prediction horizon is H minutes, then we can categorize each prediction outcome into one of the following four subsets:

- True positive (TP): if $predictor = 1$ and a seizure occurs within H minutes after the prediction.

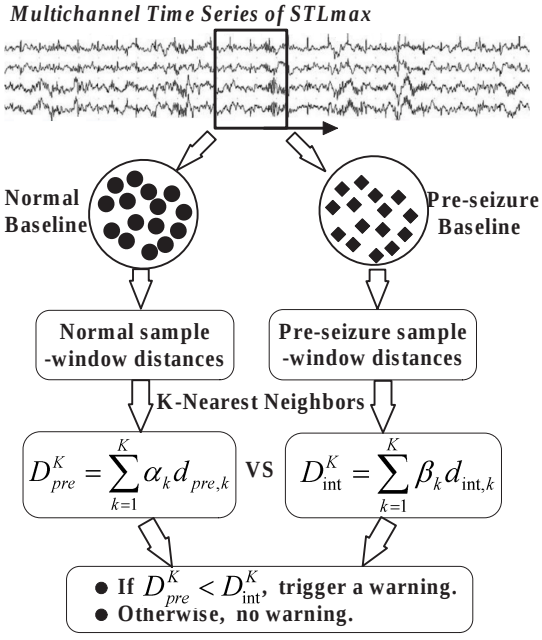


Fig. 4. Schematic structure of the KNN-based prediction rule.

- False positive (FP): if $predictor = 1$ and no seizure occurs within H minutes after the prediction.
- True negative (TN): if $predictor = 0$ and no seizure occurs within H minutes after the prediction.
- False negative (FN): if $predictor = 0$ and a seizure occurs within H minutes after the prediction.

		Prediction Outcome	
		pre-seizure	*normal*
Actual	*pre-seizure*	*TP*	*FN*
	normal	*FP*	*TN*

Fig. 5. The categorization of prediction outcomes. Each prediction outcome can always be classified into one of the four subsets (TP, FP, TN, and FN).

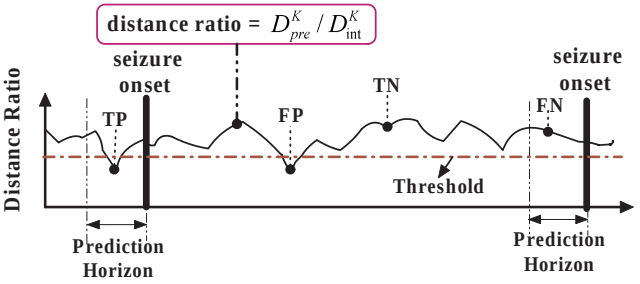


Fig. 6. A demonstration of the evaluation metrics: TP, FP, TN, and FN.

3.3.4 Baseline Updating Mechanism

The flowchart of the baseline update framework from delayed prediction feedback is shown in Figure 7. In medical practice, a physician mentally compares the

EEG patterns from an individual with the patterns from a database of many other patients and healthy people. The search of the best matching patterns can be global within the whole database, and can also be local within a sub-group of the database. We designed both local and global update rules inspired by this consideration. In particular, we designed four update rules including score-based local update (SL), score-based global update (SG), distance-based local update (DL), and distance-based global update (DG).

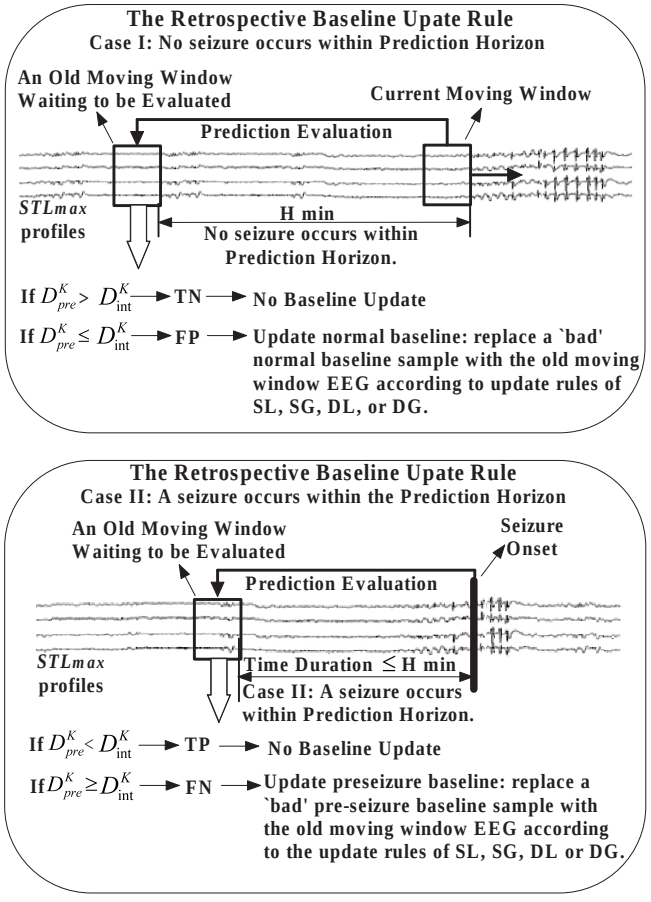


Fig. 7. Flowchart of the retrospective baseline-updating framework.

Score-Based Update: In this prediction scheme, we assume that different baseline samples have different power in decision making. We assigned a score to each baseline sample to indicate its 'importance'. The basic idea of score updating is to reinforce the scores of the 'good' baseline samples when correct predictions are made, and decrease the scores of 'bad' baseline samples when false predictions are made. The score of a baseline sample is determined by its window-sample distances. For example, if a windowed EEG epoch is mis-classified as pre-seizure via the KNN evaluation, then the pre-seizure baseline samples that are closest to, and the normal baseline samples that are furthest from, this windowed epoch will see their scores penalized according to their window-sample distances. The closest pre-seizure baseline sample and the furthest normal baseline sample receive the highest penalties. The mathematical formulations of the score updating rules are stated in the following.

At the beginning, the initial scores of the baseline sample are all equal, and are given by:

$$\alpha_i = \beta_i = \frac{1}{N},\ i = 1, \ldots, N, \tag{3}$$

where α_i and β_i are the scores of the ith sample in the pre-seizure and normal baseline, respectively. $N = 50$ is the number of samples in each baseline. Let $r \in (0, 1)$ denote the learning rate to control the update size for the scores, then the score update rule is represented as follows:

- For feedback of TP or FN (the windowed EEG is in pre-seizure period), the scores are updated by:

$$\alpha_i = \alpha_i(1 - \frac{d_{pre,i} - \overline{d}_{pre}}{\overline{d}_{pre}}) \times r, \tag{4}$$

$$\beta_i = \beta_i(1 + \frac{d_{int,i} - \overline{d}_{int}}{\overline{d}_{int}}) \times r. \tag{5}$$

- For feedback of FP or TN (the windowed EEG is in normal period), the scores are updated by:

$$\alpha_i = \alpha_i(1 + \frac{d_{pre,i} - \overline{d}_{pre}}{\overline{d}_{pre}}) \times r, \tag{6}$$

$$\beta_i = \beta_i(1 - \frac{d_{int,i} - \overline{d}_{int}}{\overline{d}_{int}}) \times r, \tag{7}$$

where $\forall i = 1, 2, \ldots, N$, $\overline{d}_{pre} = \sum_{i=1}^{N} d_{pre,i}/N$, and $\overline{d}_{int} = \sum_{i=1}^{N} d_{int,i}/N$.

For a windowed EEG epoch, the system makes a prediction by the KNN method. The feedback of this prediction is available until either of the following occurs: 1) the prediction horizon passes, or 2) a seizure occurs. Once the feedback of this prediction is given, the score-based retrospective baseline update rules are as follows:

- For case of FP: replace the lowest-scored sample in the normal K-nearest neighbors with the moving-window EEG epoch.
- For case of FN: replace the lowest-scored sample in the pre-seizure K-nearest neighbors with the moving-window EEG epoch..
- For cases of TP and TN: keep the current baseline samples unchanged.

When K equals to N, the above update is a global update rule that replaces the global lowest-scored baseline sample. When K is smaller than N, it is a local update rule which only considers the local K-nearest neighbors of a windowed EEG epoch. The score-based local and global update rules are denoted as 'SL' and 'SG', respectively, in the remaining part of this paper.

Distance-based Update: The distance between two EEG epochs indicates the degree of similarity. Intuitively,

a shorter distance means a better match, and a larger distance indicates a worse match. For a windowed EEG epoch, the goodness of a baseline sample depends on its window-sample distances. For example, suppose a normal state windowed EEG epoch, via KNN evaluation, is falsely classified as pre-seizure. We consider the furthest normal baseline sample as the 'bad' baseline sample, which may be the primary cause of the false prediction, and we replace it with the windowed EEG epoch. In summary, for a windowed EEG epoch, the retrospective distance-based baseline update rules are as follows:

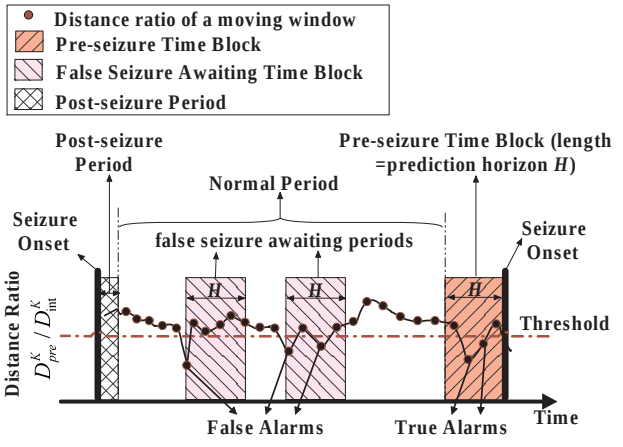
- For feedback of FP: replace the furthest sample in its K-nearest neighbors of the normal baseline with the corresponding windowed EEG epoch.
- For feedback of FN: replace the furthest sample in its K-nearest neighbors of the pre-seizure baseline with the corresponding windowed EEG epoch.
- For feedback of TP or TN: keep the current baseline samples unchanged.

Similar to 'SL' and 'SG', the distance-based update can also be local and global depending on the value of K. The distance-based local and global update rules are denoted as 'DL' and 'DG', respectively.

The overall computational complexity of our algorithm can be analyzed as follows. For the complexity of KNN, suppose each sliding window epoch has n points and each baseline has N samples. The Euclidean distance calculation takes $O(2nN)$. Finding the K nearest neighbors involves sorting, which takes additional $O(2N^2)$ steps. Finally, the summation of K distances of each baseline and the KNN decision-making process take $O(2K)$ steps. In summary, the KNN-based classification takes $O(2nN + 2N^2 + 2K)$. For the score-based updating rule, according to Equations (4)-(7), the score update of a single baseline sample runs $O(1)$ steps, thus the score update of $2N$ baseline samples takes $O(2N)$ steps. The scores of baseline samples are updated at every sliding window step. If a wrong prediction is made, a baseline with the lowest score will be replaced, which takes $O(N)$ steps. For the distance-based updating rule, two baselines are only updated when a wrong prediction is made. It takes $O(1)$ steps to find the baseline sample to be replaced, since the distances have already been computed and sorted in the KNN step.

3.4 Evaluation of Prediction Performance

To evaluate a prediction model, the most commonly used performance measures are specificity and sensitivity. In seizure prediction studies, sensitivity is usually defined as the number of correctly predicted seizures divided by the total number of seizures. A seizure is considered to be correctly predicted if there is at least one warning within its preceding prediction horizon. In this study, we also employed this definition of sensitivity, denoted as sen_{blk}. To estimate the prediction specificity, most studies calculated a false prediction rate, which is defined by the number of false predictions per hour (or unit time). However, false prediction rate does not provide enough information to infer the effect of prediction horizon on the prediction performance. For example, a patient has to wait until the end of prediction horizon to determine if a warning is false. Given the same false prediction rate, an algorithm with a 3-hour prediction horizon will give a patient much longer false awaiting time than the one with a 10-minute prediction horizon. To overcome this bias, Mormann et al. [33] suggested that a prediction specificity can be estimated by quantifying the portion of time during the normal period that is not considered to be false awaiting time. We herein employed this specificity measure, denoted as spe_{blk}. A demonstration of the sen_{blk} and spe_{blk} quantification is shown in Figure 8. In turn, we also define the overall prediction performance (OPP) as an average of sen_{blk} and spe_{blk}, i.e., $OPP = (sen_{blk} + spe_{blk})/2$. The OPP values can range from [0.0, 1.0]. An accurate prediction model should have an OPP close to 1, and a random model should have an OPP around 0.5. The closer the OPP value to one, the better the prediction performance.



- **Prediction Rule: If the distance ratio falls below the threshold, a warning is triggered.**
- **If the time length of the falsely seizure awaiting periods = 2 x *H* = 6 hours, the time length of the normal period = 15 hours, then**
  spe_{blk}= 1- 6/15 = 60% (*The portion of the normal period during which the patient does not falsely await a seizure.*)
  sen_{blk} = 1/1 = 100% (*There is at least one warning within the prediction horizon.*)

Fig. 8. A demonstration of the prediction procedure based on the distance ratio D^K_{pre}/D^K_{int}. The definition of sensitivity (sen_{blk}), specificity (spe_{blk}), false alarms, and false seizure awaiting periods are also illustrated.

Receiver Operating Characteristic (ROC) Analysis:

In any prediction algorithm, one can always make a trade-off between sensitivity and specificity, such as increasing sensitivity at the expense of a lower specificity. A common way to compare different pre-

diction models is to construct a ROC curve that plots sensitivity versus (1-specificity) whereas the decision boundary of the prediction model is varied throughout its range. The *area under the ROC curve* (AUC) is commonly used to access the overall prediction power of a prediction model. AUC values are usually between 0.5 and 1. A perfect prediction model has an AUC value of 1 while a random chance model has an AUC of around 0.5.

4 RESULTS

4.1 Computational Settings

The proposed prediction approach was tested on EEG recordings of 10 patients with epilepsy using three prediction horizons, four baseline-update rules, four settings of KNN, and three types of similarity measures. The summary of the parameter settings of the prediction system is shown in Table 2.

TABLE 2
Summary of the settings of the prediction system.

Parameter Setting	Values or Choices
Moving Window	10 minutes length with 50% overlap each step
Prediction Horizon	30 minutes, 90minutes, 150minutes
Similarity Measure	EU, TS, DTW
The value of K	3, 7, half, all
Update Scheme	1. Non-update (No update to the initial baselines) 2. SL (score-based local update) 3. SG (score-based global update) 4. DL (distance-based local update) 5. DG (distance-based global update)

4.2 Random Predication Models

There has been no definite conclusion whether prospective algorithms can predict seizures based on EEG analysis. Before applying it to any clinical application, it is necessary to evaluate if the designed prediction model is indeed able to perform better than a chance model. Therefore, we compared the performance of the proposed adaptive prediction model with two random prediction schemes: periodic prediction scheme and Poisson prediction scheme. The periodic prediction scheme gives warnings at a fixed time interval T. The Poisson prediction scheme issues a warning according to an exponential distributed random time interval with a fixed mean λ. We performed the periodic prediction scheme and the Poisson prediction scheme for each patient. The values of λ and T were determined according to the average length of inter-seizure intervals for each patient as shown in Table 1. For example, for patient 1, the averaged inter-seizure interval is 12.17 hours, we set $\lambda = T = 12.17$ hours. This is the best value setting of T and λ the one could obtain.

4.3 Prediction Performance of sen_{blk} and spe_{blk}

For each patient, the EEG recordings were divided into training and testing dataset. The training dataset is the EEG recordings that contain the first half of seizure occurrences. It is used to train our approach to find the best parameter setting. The testing dataset is the EEG recordings that contain the second half of seizure occurrences. It is used to test our prediction approach ***prospectively*** using the best parameter setting found from the training data. The best parameter setting is defined as one with the highest OPP value. In addition, to find the most appropriate trade-off between sensitivity and specificity, we also added a constraint that the sen_{blk} must be greater than 0.6, and the spe_{blk} must be greater than 0.4. If none of the settings meet this constraint, we simply selected the one with the highest OPP value.

Table 3 summarizes the performance characteristics of the adaptive learning prediction scheme in the training and testing dataset. To determine the importance and effectiveness of the proposed baseline-update rule, we also summarizes the performance characteristics of the non-update prediction scheme in Table 3. The non-update prediction scheme employed the same initial baselines as the adaptive ones for each patient, and kept the baseline unchanged throughout the prediction process. Table 3 clearly shows that the training and testing OPP values of the adaptive learning approach are considerably higher than those of non-update prediction scheme in all the three prediction horizons. To compare with random predictions, the prediction results of the periodic and Poisson prediction schemes are also shown in Table 3. The adaptive learning approach performed much better than the two random prediction schemes in terms of the overall OPP values.

The adaptive prediction approach achieved the best overall performance using the prediction horizon of 150 minutes. An example of the prediction outcomes of the adaptive prediction system is also shown in Figure 9. In general, the averaged testing OPP over the 10 patients of the adaptive prediction approach is 0.70, which is 14%, 25%, and 27% higher than that of non-update prediction scheme, the Poisson prediction scheme, and the periodic prediction scheme, respectively. Starting from the initial (less representative) baseline samples, the adaptive system increased the prediction performance considerably by baseline-updating for each individual patient. The experimental results confirmed our goal that it is possible to achieve personalized prediction through adaptive learning approaches. In addition, one can observe an increasing trend of the averaged OPP values for both adaptive and non-update prediction schemes when the prediction horizon increases from 30 minutes to 150 minutes. This may indicate that the prediction horizon of 150 minutes is a better estimate of the real length of pre-seizure periods. The length of prediction horizon is very crucial since a better estimate of pre-seizure periods will give better reinforcement feedbacks to the adaptive learning system, and thus will

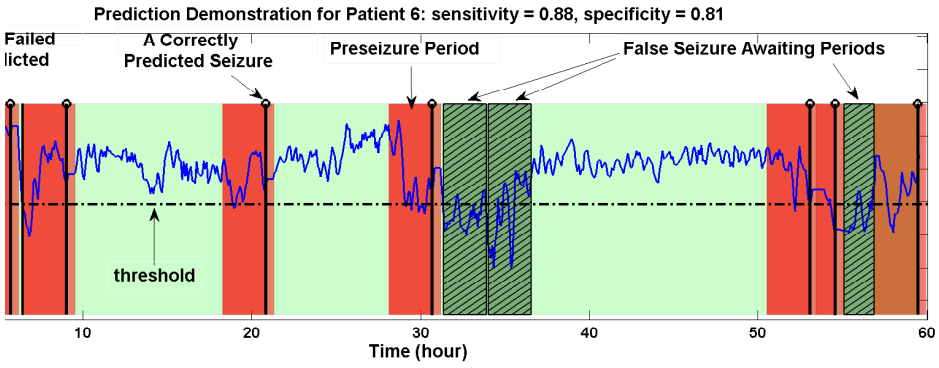


Fig. 9. An example of the prediction outcomes of the adaptive prediction system for patient 6 using the prediction horizon of 150 minutes. Other experimental settings are SG, K =all, and DTW. The vertical black lines are the recorded seizures in this patient, and the dashed horizontal line is the threshold of distance ratio. A warning is issued if the distance ratio falls below the threshold.

TABLE 3
The training and testing performance characteristics of the adaptive prediction approach and the non-update prediction scheme. The performance characteristics of the two random prediction schemes (periodic and Poisson) are also reported using $T = \lambda =$ averaged length of inter-seizure intervals for each patient.

		Adaptive Scheme					Non-Update Scheme					Possion Predictor		Periodic Predictor	
		Training		Testing			Training		Testing						
Horizon	Patient	Setting	sen_{blk}	spe_{blk}	sen_{blk}	spe_{blk}	Setting	sen_{blk}	spe_{blk}	sen_{blk}	spe_{blk}	sen_{blk}	spe_{blk}	sen_{blk}	spe_{blk}
30minutes	1	SG-3-DTW	1.00	0.77	0.33	0.55	3-EU	0.67	0.82	0.33	0.48	0.00	0.53	0.00	0.53
	2	SG-all-TS	0.67	0.87	1.00	0.62	all-DTW	0.67	0.65	1.00	0.24	0.00	0.35	0.00	0.35
	3	DL-half-DTW	0.54	0.79	0.50	0.69	half-TS	0.69	0.43	0.20	0.75	0.17	0.52	0.17	0.52
	4	SG-3-TS	0.38	0.95	0.00	0.97	3-TS	0.13	0.93	0.14	0.91	0.00	0.72	0.07	0.72
	5	SG-3-DTW	0.63	0.60	0.25	0.74	3-DTW	0.88	0.25	1.00	0.31	0.00	0.39	0.00	0.39
	6	DG-half-DTW	1.00	0.73	0.75	0.86	all-EU	1.00	0.48	0.50	0.41	0.13	0.34	0.00	0.34
	7	SL-7-DTW	0.70	0.71	0.44	0.73	all-TS	0.70	0.63	0.56	0.55	0.05	0.52	0.00	0.52
	8	SL-7-EU	0.80	0.91	0.00	0.88	half-EU	0.20	0.90	0.00	1.00	0.06	0.77	0.12	0.77
	9	SG-3-DTW	0.22	0.97	0.30	0.94	3-TS	0.67	0.46	0.80	0.36	0.05	0.95	0.05	0.95
	10	SL-half-DTW	0.80	0.62	0.00	1.00	7-DTW	1.00	0.43	1.00	0.62	0.00	0.70	0.09	0.70
		Ave.	**0.62**	**0.79**	**0.31**	**0.80**		**0.62**	**0.59**	**0.54**	**0.50**	**0.06**	**0.96**	**0.06**	**0.96**
		PA	**0.71**		**0.56**			**0.61**		**0.52**		**0.51**		**0.51**	
90minutes	1	DG-3-DTW	1.00	0.71	1.00	0.30	3-DTW	1.00	0.60	1.00	0.37	0.00	0.47	0.33	0.47
	2	DG-3-TS	0.67	0.46	1.00	0.75	3-EU	0.00	1.00	0.67	0.57	0.00	0.17	0.00	0.17
	3	SG-7-EU	0.77	0.67	0.60	0.35	7-EU	0.85	0.47	0.30	0.73	0.30	0.15	0.39	0.15
	4	SL-7-TS	0.63	0.69	0.71	0.71	half-EU	0.75	0.45	0.43	0.32	0.07	0.39	0.07	0.39
	5	SL-7-EU	0.63	0.96	0.00	0.86	3-DTW	1.00	0.12	1.00	0.12	0.00	0.15	0.00	0.15
	6	DG-3-TS	0.50	0.89	0.75	0.76	3-DTW	1.00	0.23	0.50	0.29	0.25	0.00	0.25	0.00
	7	DL-3-DTW	0.80	0.40	0.89	0.34	3-DTW	1.00	0.14	1.00	0.27	0.21	0.22	0.21	0.22
	8	DG-7-EU	0.90	0.96	0.71	0.64	3-DTW	0.40	0.82	0.00	0.92	0.18	0.57	0.18	0.57
	9	DL-7-TS	0.78	0.57	0.20	0.85	all-DTW	0.67	0.45	0.90	0.36	0.11	0.79	0.16	0.79
	10	SL-7-DTW	0.80	0.43	0.67	0.71	half-EU	1.00	0.28	1.00	0.74	0.18	0.24	0.18	0.24
		Ave.	**0.75**	**0.68**	**0.58**	**0.71**		**0.78**	**0.38**	**0.67**	**0.43**	**0.15**	**0.88**	**0.19**	**0.88**
		PA	**0.72**		**0.65**			**0.58**		**0.55**		**0.52**		**0.54**	
150minutes	1	DL-half-DTW	1.00	0.81	1.00	0.40	all-TS	1.00	0.54	1.00	0.40	0.33	0.56	0.33	0.56
	2	DG-7-EU	0.67	0.53	0.67	0.84	3-TS	0.00	1.00	0.67	0.87	0.00	0.10	0.00	0.10
	3	DL-half-3	0.92	0.74	0.80	0.45	7-DTW	0.85	0.74	0.60	0.64	0.70	0.25	0.65	0.25
	4	DL-3-TS	0.63	0.66	0.43	0.82	3-DTW	1.00	0.18	0.71	0.18	0.13	0.25	0.13	0.25
	5	DG-7-EU	0.63	0.86	0.63	0.59	3-TS	1.00	0.15	1.00	0.12	0.13	0.09	0.00	0.09
	6	SG-all-DTW	0.75	1.00	1.00	0.75	7-DTW	0.75	0.72	1.00	0.84	0.13	0.17	0.13	0.17
	7	DG-3-EU	0.60	0.65	0.89	0.58	7-TS	0.60	0.47	0.89	0.25	0.37	0.19	0.37	0.19
	8	DL-3-EU	0.90	0.92	0.57	0.65	all-DTW	0.50	0.64	0.00	0.93	0.24	0.47	0.24	0.47
	9	DL-3-DTW	0.78	0.54	0.60	0.56	3-DTW	1.00	0.22	1.00	0.13	0.32	0.49	0.26	0.49
	10	DL-7-DTW	0.60	0.55	1.00	0.41	3-DTW	1.00	0.14	1.00	0.25	0.09	0.25	0.27	0.25
		Ave.	**0.75**	**0.69**	**0.73**	**0.67**		**0.79**	**0.28**	**0.78**	**0.45**	**0.29**	**0.82**	**0.28**	**0.82**
		PA	**0.72**		**0.70**			**0.54**		**0.62**		**0.56**		**0.55**	

lead to a better prediction performance.

4.4 Receiver Operating Characteristic Analysis

The effectiveness of the proposed four adaptive prediction schemes was also evaluated by the ROC analysis. Table 4 summarizes the AUC values of the four adaptive schemes (SL, SD, DL, and DG), the non-update scheme, and the two random schemes (periodic and the Poisson). The four adaptive schemes and the non-update scheme employed the best parameter settings obtained from the training data of each patient. For each prediction scheme with a selected setting, the sensitivity and specificity of the entire EEG recordings of a patient were used to generate ROC curves. The parameter used to generate ROC curves is the threshold of the distance ratio R^*, which was tuned from 0.1 to 10 to make a broad spectrum of tradeoff between sensitivity and specificity. For the periodic and Poisson schemes, the sensitivity and specificity tradeoff is controlled by the parameters T and λ, respectively. The ROC curves were obtained by tuning T and λ from 0.1 to 20 hours. We performed 300

TABLE 4
AUC Comparison of the four adaptive prediction schemes with the non-update and the two random prediction schemes. The four adaptive prediction schemes and the non-update prediction scheme employed the best parameter settings using the training data set. Their ROC curves were obtained by tuning the threshold of distance ratio R^* from 0.1 to 10 to make a broad spectrum of tradeoff between sensitivity and specificity. For the periodic and Poisson prediction schemes, the ROC curves were obtained by tuning λ and T from 0.1 to 20 hours for each patient. We performed 300 Monte Carlo simulations for both random prediction schemes, a set of λ and T were randomly, uniformly selected from [0.1, 20] hours at each experiment. The averaged AUC values over the 300 experiments are reported in this table.

		SL		SG		DL		DG		None		Poisson	Periodic
Horizon	Patient	Setting	AUC	Setting	AUC	Setting	AUC	Setting	AUC	Setting	AUC	AUC	AUC
30 minutes	1	all-DTW	0.73	3-DTW	0.8	3-DTW	0.73	all-TS	0.77	3-EU	0.8	0.52	0.51
	2	all-TS	0.73	all-TS	0.73	half-DTW	0.71	half-DTW	0.79	all-DTW	0.35	0.51	0.51
	3	half-TS	0.66	3-TS	0.59	half-DTW	0.64	3-EU	0.7	half-TS	0.52	0.56	0.54
	4	3-TS	0.69	7-DTW	0.61	7-TS	0.61	7-EU	0.64	3-TS	0.59	0.51	0.50
	5	7-DTW	0.57	3-DTW	0.57	all-EU	0.59	3-DTW	0.54	3-DTW	0.64	0.52	0.51
	6	3-DTW	0.62	half-DTW	0.74	half-DTW	0.68	half-DTW	0.85	all-EU	0.53	0.49	0.52
	7	7-DTW	0.7	half-DTW	0.68	half-DTW	0.66	7-EU	0.67	all-TS	0.65	0.53	0.51
	8	7-EU	0.78	7-TS	0.75	3-TS	0.81	half-EU	0.82	half-EU	0.61	0.49	0.50
	9	all-DTW	0.69	3-DTW	0.73	half-TS	0.73	3-EU	0.63	3-TS	0.6	0.51	0.51
	10	half-DTW	0.67	3-DTW	0.63	half-EU	0.7	3-EU	0.74	7-DTW	0.81	0.50	0.52
	Ave.		**0.68**		**0.68**		**0.69**		**0.72**		**0.61**	**0.51**	**0.51**
90 minutes	1	half-TS	0.76	half-EU	0.79	half-TS	0.76	3-DTW	0.84	3-DTW	0.75	0.51	0.54
	2	half-DTW	0.37	3-TS	0.66	7-TS	0.52	3-TS	0.7	half-DTW	0.33	0.49	0.48
	3	half-DTW	0.6	7-EU	0.66	3-TS	0.66	3-TS	0.64	7-EU	0.61	0.62	0.60
	4	7-TS	0.7	3-TS	0.59	3-DTW	0.62	7-DTW	0.53	half-EU	0.59	0.52	0.52
	5	7-EU	0.69	3-EU	0.63	half-EU	0.67	3-DTW	0.56	3-DTW	0.55	0.52	0.52
	6	half-DTW	0.67	all-TS	0.58	all-TS	0.6	3-TS	0.75	3-DTW	0.65	0.54	0.57
	7	7-EU	0.61	3-DTW	0.47	3-DTW	0.6	half-DTW	0.57	3-DTW	0.5	0.57	0.55
	8	3-EU	0.59	3-EU	0.64	half-TS	0.78	7-EU	0.89	3-DTW	0.49	0.51	0.50
	9	half-DTW	0.57	7-DTW	0.67	7-TS	0.57	7-TS	0.65	all-DTW	0.6	0.50	0.53
	10	7-DTW	0.64	all-EU	0.78	7-DTW	0.6	7-DTW	0.58	half-EU	0.52	0.52	0.52
	Ave.		**0.62**		**0.65**		**0.64**		**0.67**		**0.56**	**0.53**	**0.53**
150 minutes	1	half-TS	0.76	7-EU	0.9	half-DTW	0.93	half-TS	0.77	all-TS	0.79	0.56	0.55
	2	3-EU	0.77	7-EU	0.59	3-TS	0.66	7-EU	0.77	all-DTW	0.3	0.46	0.45
	3	half-EU	0.62	7-DTW	0.73	half-DTW	0.73	all-DTW	0.74	7-DTW	0.79	0.62	0.62
	4	7-TS	0.65	7-EU	0.5	3-TS	0.64	3-EU	0.65	3-DTW	0.54	0.54	0.54
	5	7-EU	0.62	3-TS	0.62	half-EU	0.69	7-EU	0.71	3-TS	0.56	0.50	0.51
	6	all-DTW	0.75	all-DTW	0.75	7-DTW	0.76	half-DTW	0.75	3-DTW	0.84	0.54	0.51
	7	7-DTW	0.47	7-DTW	0.49	half-DTW	0.67	3-EU	0.65	7-TS	0.5	0.60	0.59
	8	7-TS	0.86	3-EU	0.82	3-EU	0.84	3-EU	0.83	all-DTW	0.49	0.50	0.50
	9	7-TS	0.68	3-DTW	0.62	3-DTW	0.61	3-DTW	0.6	3-DTW	0.6	0.54	0.54
	10	3-DTW	0.53	3-EU	0.82	7-DTW	0.69	3-EU	0.62	3-DTW	0.52	0.54	0.54
	Ave.		**0.67**		**0.68**		**0.72**		**0.71**		**0.59**	**0.54**	**0.54**

Monte Carlo simulations for both random schemes, a set of λ and T were randomly, uniformly selected from [0.1, 20] hours at each experiment. The averaged AUC values over 300 experiments are reported in Table 4.

One can clearly observe that the four adaptive schemes generally have higher AUC values than the non-update and the two random schemes. When using the prediction horizons of 150 minutes, the averaged AUC values of the four adaptive schemes (SL, SG, DL, and DG) are 0.67, 0.68, 0.72, 0.71, respectively. The averaged AUC values of SL, SG, DL, and DG are 14%, 15%, 22%, and 20% higher than the averaged AUC value of the non-update scheme. This indicates that all the proposed four adaptive prediction schemes increased the overall prediction performance of the system through adaptive baseline-updating. When compared to the random schemes, the averaged AUC values of SL, SG, DL, and DG are 24%, 26%, 33%, and 31% higher than the averaged AUC values of the Periodic and Poisson scheme (both are 0.54). The significant higher AUC values strongly indicate that the adaptive prediction schemes has a much higher prediction power than random predictions. Similar results can also be obtained when using the prediction horizons of 30 minutes and 90 minutes.

To make a solid statistical comparison, it is also interesting to investigate the performance of the four adaptive schemes as well as the non-update scheme on all the parameter settings over the 10 patients. For each scheme (adaptive and non-update), there are 36 settings including four choices of K, three choices of distance measures, and three choices of prediction horizons. Figure 10 shows the boxplots of the averaged AUC values over 10 patients for the entire 36 settings of each scheme. The AUC values of the two random schemes obtained from 300 Monte Carlo simulations are shown in Figure 10 for comparison. The boxplot clearly shows that the AUC values of the four proposed adaptive prediction schemes have significantly different distributions with those of the non-update and random schemes. We used the AUC values of the non-update scheme as the baseline group, and performed paired t-test for the AUC values of the four adaptive schemes and the two random schemes. As shown in the Figure 10, the p-value of each paired t-test is smaller than 0.001. This outcome indicates that the four adaptive

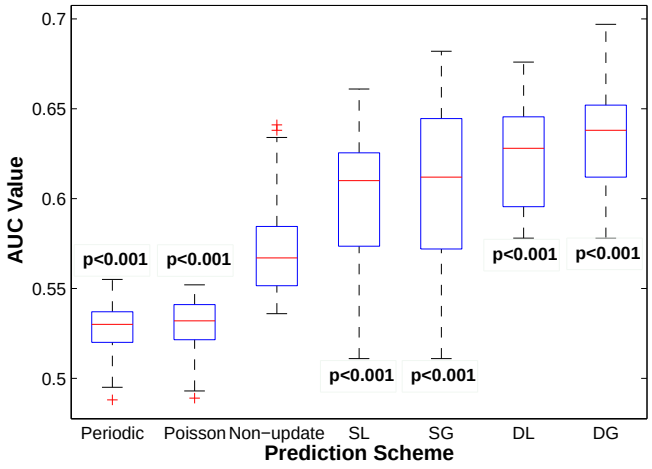


Fig. 10. Box-plot of the AUC values of the four adaptive schemes, the non-update scheme, and the two random schemes. The AUC values of the adaptive and non-update schemes are the averaged AUC values over 10 patients for all possible parameter settings (=36) of each scheme. The AUC values of the two random schemes are obtained from 300 Monte Carlo simulations, in each of which a set of values of $lambda$ and T are randomly and uniformly varied from 0.1 to 20 hours. Each box shows the median, interquartile range, minimum and maximum of the AUC values of each prediction scheme. Using AUC values of the nondicated-update scheme as the baseline group, the p-values of the paired t-tests for the AUC values of other prediction scheme are indicated in the plot. The four adaptive schemes performed significantly better than the non-update scheme with all p-values smaller than 0.001. While the non-update scheme performed significantly better than the two random schemes with both p-values smaller than 0.001.

prediction schemes all performed significantly better than the non-update scheme. While the non-update scheme performed significantly better than the two random schemes. This is not unexpected, since the initial baseline samples employed by the non-update scheme already contained some useful information of the preseizure and normal EEG patterns. It thus worked better than random predictions.

When we compare among the four proposed adaptive schemes, we found that the two distance-based update schemes (DL and DG) performed better than the two score-based update schemes with p-values smaller than 0.001. This outcome implies that the distance-based update rule did a better job in the online baseline-updating than the score-based rule. In addition, the AUC values of the two score-based update schemes SL and SG are comparable with a p-value of 0.15; and DG worked a little better than DL with a p-value of 0.02.

4.5 Comparisons to Other Seizure Prediction Methods

Although over the past decade there have been several studies in seizure prediction, almost all of them are focused on retrospective analyses of prediction, which is to show that there are detectable changes in EEG signals prior to a seizure. Those studies often used short-term EEG recordings sampled 45 to 90 minutes before a seizure. Very few studies investigate online seizure prediction algorithms using prospective analysis of continuous long-term EEG recordings [45], [49]. Because the nature of prospective and retrospective analyses, it is extremely hard to compare the real prediction performances between these algorithms. According to a seminal review paper in seizure prediction [33], more transformable and unambiguous performance measures such as the portion of false awaiting time are suggested. Here we compare our prediction results with the studies by Sackellares et al. (2006) and Snyder et al. (2008) [45], [49] that report the sensitivity and the portion of false awaiting time of their prospective analyses. Sackellares et al. [45] evaluated an adaptive seizure prediction approach on 10 patients. Given a prediction horizon of 150 minutes and a sensitivity of 80%, the portion of false awaiting time is 37% (corresponding to our specificity of 63%) on average over the 10 patients. Snyder et al. [49] performed a prospective seizure prediction on 4 patients using a prediction horizon of 120 minutes. The averaged sensitivity is 82.3% and the portion of false awaiting time is 30.5% (corresponding to our specificity of 69.5%). The OPP values of the two studies are 0.72 and 0.76, respectively. It was not clearly indicated in both papers if the reported OPP values were based on the training set, the testing set or the entire dataset. Thus, we base our comparison on the performance on an entire EEG recording of each patient using the best parameter setting obtained from our previous analysis. With a prediction horizon of 150 minutes, our adaptive learning approach achieved a sensitivity of 77% and a specificity of 73% on average over the 10 patients. The OPP value of our approach is 0.75, which is comparable to those in the two seminal studies. We must note that the comparison might not provide a conclusive result as it is based on different datasets with different population sizes, and the data characteristics vary greatly. A large public EEG database for seizure prediction is on the way and a proper comparison remains to be further investigated in our future study.

5 Conclusions and Discussion

This study investigated the challenging problem of epileptic seizure prediction. We introduced an adaptive learning approach, which combine reinforcement

learning, online monitoring and adaptive control theory to achieve a personalized seizure prediction. Using EEG recordings from 10 patients with epilepsy, we demonstrated that the adaptive learning algorithm was effective in increasing prediction performance of the system through adaptive baseline-updating. The best prediction performance was achieved using the prediction horizon of 150 minutes, in which the averaged sensitivity was 73% and the averaged specificity was 67%. The ROC analysis demonstrated that the adaptive prediction schemes indeed performed much better than the non-update scheme and the two chance models.

The experimental results of this study are quite encouraging, and they suggest that the proposed adaptive approach performed better than random predictors [33]. An autonomous learning framework like the one proposed here was shown capable of self-adjusting the baseline samples for each individual patient without a tedious parameter tuning process. With this attractive online learning ability, the proposed adaptive learning prediction system is expected to be able to further improve the prediction performance when more EEG recordings are available for each patient. It is important to remark that in this study the online prediction algorithm was evaluated based on a perfect seizure detection (i.e., actual seizure timing is provided after our algorithm makes prediction). This study did not, however, investigate the integration of our online prediction algorithm with any existing seizure detection algorithms. The reason is that although there exist a number of seizure detection algorithms embedded in clinical EEG systems, most detection algorithms still suffer from an extremely high false detection rate. Thus, the impact of false detections from seizure detection systems on the performance of our seizure prediction algorithm is beyond the scope of this study and remains to be further investigated in our future work. If one wants to test a fully automated framework, our seizure prediction algorithm can be readily integrated to any clinical EEG system, and it can be fully automated by relying on existing automated seizure detection algorithms in the EEG system. In practice, a prospective seizure prediction system must have both high sensitivity and specificity for clinical use. If such a seizure-warning device is to become a reality, we envision that adaptive learning techniques will definitely play an important role in handling the great variety of brain-wave patterns among different patients.

REFERENCES

[1] R. Aschenbrenner-Scheibe, T. Maiwald, M. Winterhalder, H. Voss, J. Timmer, and A. Schulze-Bonhage. How well can epileptic seizures be predicted? An evaluation of a nonlinear method. *Brain*, 126:2616–2626, 2003.

[2] W. Chaovalitwongse, Y. Fan, and R. Sachdeo. On the time series K-nearest neighbor classification of abnormal brain activity. *IEEE Transactions on Systems, Man, and Cybernetics, Part A:Systems and Humans*, 37(6):1005–1016, 2007.

[3] W. Chaovalitwongse, Y. Fan, and R. Sachdeo. Novel optimization models for abnormal brain activity classification. *Operations Research*, 56(6):1450–1460, 2008.

[4] W. Chaovalitwongse and P. Pardalos. On the time series support vector machine using dynamic time warping kernel for brain activity classification. *Cybernetics and Systems Analysis*, 44(1):125–138, 2008.

[5] W. Chaovalitwongse, P. Pardalos, L. Iasemidis, D. Shiau, and J. Sackellares. Dynamical approaches and multi-quadratic integer programming for seizure prediction. *Optimization Methods and Software*, 20(2-3):383–394, 2005.

[6] W. Chaovalitwongse, P. Pardalos, and O. Prokopyev. Electroencephalogram (EEG) time series classification: applications in epilepsy. *Annals of Operations Research*, 148(1):227–250, 2006.

[7] W. Chaovalitwongse, W. Suharitdamrong, C. Liu, and M. Anderson. Brain network analysis of seizure evolution. *Annales Zoologici Fennici*, 45(5):402–414, 2008.

[8] L. Chisci, A. Mavino, G. Perferi, M. Sciandrone, C. Anile, G. Colicchio, and F. Fuggetta. Real-time epileptic seizure prediction using AR models and support vector machines. *IEEE Transactions on Biomedical Engineering*, 57(5):1124–1132, 2010.

[9] B. Cox, T. Kislinger, and A. Emili. Integrating gene and protein expression data: pattern analysis and profile mining. *Methods*, 35(3):303–314, 2005.

[10] M. D'Alessandro, R. Esteller, G. Vachtsevanos, A. Hinson, J. Echauz, and B. Litt. Epileptic seizure prediction using hybrid feature selection over multiple intracranial EEG electrode contacts: a report of four patients. *Biomedical Engineering, IEEE Transactions on*, 50(5):603 –615, 2003.

[11] S. Dangel, P. Meier, H. Moser, S. Plibersek, and Y. Shen. Time series analysis of sleep EEG. *Computer assisted Physics*, pages 93–95, 1999.

[12] C. Davatzikos, K. Ruparel, Y. Fan, D. Shen, M. Acharyya, J. Loughead, R. Gur, and D. Langleben. Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. *NeuroImage*, 28(3):663–668, 2005.

[13] R. Duda, P. Hart, and D. Stork. *Unsupervised Learning and Clustering (2nd edition)*. Wiley, New York, 2001.

[14] G. Endo, J. Morimoto, T. Matsubara, J. Nakanishi, and G. Cheng. Learning CPG-based biped locomotion with a policy gradient method: application to a humanoid robot. *The International Journal of Robotics Research*, 27(2):213–228, 2008.

[15] J. Engel and T. Pedley. *Epilepsy: A Comprehensive Textbook*. Lippincott Williams & Wilkins, Philadelphia, PA, 1997.

[16] H. Feldwisch-Drentrup, B. Schelter, M. Jachan, J. Nawrath, J. Timmer, and A. Schulze-Bonhage. Joining the benefits: combining epileptic seizure prediction methods. *Epilepsia*, 51(8):1598–1606, 2010.

[17] L. Good, S. Sabesan, S. Marsh, K. Tsakalis, D. Treiman, and L. Iasemidis. Nonlinear dynamics of seizure prediction in a rodent model of epilepsy. *Nonlinear Dynamics Psychol Life Science*, 14(4):411–434, 2010.

[18] S. Haas, M. Frei, and I. Osorio. Strategies for adapting automated seizure detection algorithms. *Medical Engineering & Physics*, 29(8):895–909, 2007.

[19] L. Iasemidis. *On the dynamics of the human brain in temporal lobe epilepsy*. PhD thesis, University of Michigan, Ann Arbor, 1991.

[20] L. Iasemidis, J. Principe, J. Czaplewski, R. Gilman, S. Roper, and J. Sackellares. Spatiotemporal transition to epileptic seizures: A nonlinear dynamical analysis of scalp and intracranial EEG recordings. In *Spatiotemporal Models in Biological and Artificial Systems*, pages 81–88. Amsterdam: IOS Press, 1997.

[21] L. Iasemidis and J. Sackellares. Long time scale spatio-temporal patterns of entrainment in preictal ECoG data in human temporal lobe epilepsy. *Epilepsia*, 31:621, 1990.

[22] L. Iasemidis, D. Shiau, W. Chaovalitwongse, J. Sackellares, P. Pardalos, J. Principe, P. Carney, A. Prasad, B. Veeramani, and K. Tsakalis. Adaptive epileptic seizure prediction system. *IEEE Transactions on Biomedical Engineering*, 50(5):616–627, 2003.

[23] L. Iasemidis, D. Shiau, P. Pardalos, W. Chaovalitwongse, K. Narayanana, A. Prasada, K. Tsakalis, P. Carney, and J. Sackellares. Long-term prospective online real-time seizure prediction. *Clinical Neurophysiology*, 116:532–544, 2005.

[24] L. Iasemidis, H. Zaveri, J. Sackellares, and W. Williams. Linear and nonlinear modeling of ECoG in temporal lobe epilepsy.

25th Annual Rocky Mountain Bioengineering Symposium, 24:187–193, 1988.

[25] L. Iasemidis, H. Zaveri, J. Sackellares, W. Williams, and T. Hood. Nonlinear dynamics of electrocorticographic data. *Journal of Clinical Neurophysiology*, 5:339, 1988.

[26] S. Kotsiantis. Supervised machine learning: A review of classification techniques. *Informatica Journal*, 31:249–268, 2007.

[27] H. Lange, J. Lieb, J. Engel, and P. Crandall. Temporo-spatial patterns of preictal spike activity in human temporal lobe epilepsy. *Electroencephalography and Clinical Neurophysiology*, 56:543–555, 1983.

[28] K. Lehnertz, R. Andrzejak, J. Arnhold, T. Kreuz, F. Morman, C. Rieke, G. Widman, and C. E. Elger. Nonlinear EEG analysis in epilepsy: its possible use for interictal focus localization, seizure anticipation and prevention. *Journal of Clinical Neurophysiology*, 18:209–222, 2001.

[29] K. Lehnertz and C. Elger. Can epileptic seizures be predicted? evidence from nonlinear time series analysis of brain electrical activity. *Physics Review Letters*, 80:5019–5022, 1998.

[30] K. Lehnertz and B. Litt. The first international collaborative workshop on seizure prediction: summary and data description. *Clinical Neurophysiology*, 116(3):493–505, 2005.

[31] L. Li, H. Tang, Z. Wu, J. Gong, M. Gruidl, J. Zou, M. Tockman, and R. Clark. Data mining techniques for cancer detection using serum proteomic profiling. *Artificial Intelligence in Medicine*, 32:71–83, 2004.

[32] B. Litt, R. Esteller, J. Echauz, M. D'Alessandro, R. Shor, T. Henry, P. Pennell, C. Epstein, R. Bakay, and M. Dichter. Epileptic seizures may begin hours in advance of clinical onset: a report of five patients. *Neuron*, 30:51–64, 2001.

[33] F. Mormann, R. Andrzejak, C. Elger, and K. Lehnertz. Seizure prediction: The long and winding road. *Brain*, 130(2):314–333, 2007.

[34] F. Mormann, R. Andrzejak, T. Kreuz, C. Rieke, P. David, C. Elger, and K. Lehnertz. Automated detection of a preseizure state based on a decrease in synchronization in intracranial electroencephalogram recordings from epilepsy patients. *Physical Review E*, 67:021912, 2003.

[35] F. Mormann, T. Kreuz, C. Rieke, R. Andrzejak, A. Kraskov, P. David, C. Elger, and K. Lehnertz. On the predictability of epileptic seizures. *Journal of Clinical Neurophysiology*, 116(3):569–587, 2006.

[36] I. Osorio, M. G. Frei, and S. B. Wilkinson. Real-time automated detection and quantitative analysis of seizures and short-term prediction of clinical onset. *Epilepsia*, 39(6):615–627, 1998.

[37] H. Persson, K. Alberts, B. Farahmand, and T. Tomson. Risk of extremity fractures in adult outpatients with epilepsy. *Epilepsia*, 43(7):768–772, 2002.

[38] H. Potter. Anatomy of the brain. http://faculty.ucc.edu/biology-potter/TheBrain/, 2006.

[39] R. Quiroga, J. Arnhold, K. Lehnertz, and P. Grassberger. Kulback-leibler and renormalized entropies: applications to electroencephalograms of epilepsy patients. *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics*, 62:8380–8386, 2000.

[40] M. Quyen, J. Soss, V. Navarro, R. Robertson, M. Chavez, M. Baulac, and J. Martinerie. Preictal state identification by synchronization changes in long-term intracranial EEG recordings. *Clinical Neurophysiology*, 116:559–568, 2005.

[41] M. L. V. Quyen, V. Navarro, M. Baulac, B. Renault, and J. Martinerie. Anticipation of epileptic seizures from standard EEG recordings. *The Lancet*, 361(9361):970–971, 2003.

[42] P. Rajdev, M. Ward, J. Rickus, R. Worth, and P. Irazoqui. Real-time seizure prediction from local field potentials using an adaptive wiener algorithm. *Computers in biology and medicine*, 40(1):97–108, 2010.

[43] P. Rapp, T. Bashore, J. Martinerie, A. Albano, I. Zimmerman, and A. Mess. Dynamics of brain electrical activity. *Brain Topography*, 2:99–118, 1989.

[44] S. Richter, D. Aberdeen, and J. Yu. Natural actor-critic for road traffic optimisation. In *Advances in neural information processing systems*, pages 1169–1176, Cambridge, MA, 2007. MIT Press.

[45] J. Sackellares, D. Shiau, J. Principe, M. Yang, L. Dance, W. Suharitdamrong, W. Chaovalitwongse, P. Pardalos, and L. Iasemidis. Predictibility analysis for an automated seizure prediction algorithm. *Journal of Clinical Neurophysiology*, 23(6):509–520, 2006.

[46] Y. Salant, I. Gath, and O. Henriksen. Prediction of epileptic seizures from two-channel EEG. *Medical and Biological Engineering and Computing*, 36:549–556, 1998.

[47] P. Senin. Dynamic time warping algorithm review. Technical report, Information and Computer Science Departament University of Hawaii, Honolulu, 2008.

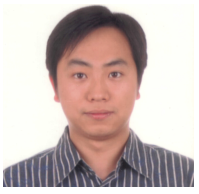
[48] C. Silva, I. Pimentel, A. Andrade, J. Foreid, and E. Ducla-Soares. Correlation dimension maps of EEG from epileptic absences. *Brain Topography*, 11:201–209, 1999.

[49] D. Snyder, J.E., D. Grimes, and B. Litt. The statistics of a practical seizure warning system. *Journal of Neural Engineering*, 5(4):392, 2008.

[50] R. Sutton and A. Barto. *Reinforcement learning: An Introduction*. MIT Press, 1998.

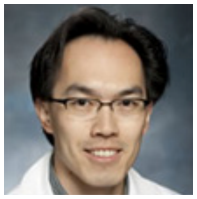
[51] J. Vastano and E. Kostelich. Comparison of algorithms for determining Lyapunov exponents from experimental data. In *International conference on dimensions and entropies in chaotic systems*, pages 100–107, Pecos River, NM, USA, 1985.

[52] S. Viglione and G. Walsh. Epileptic seizure prediction. *Electroencephalography and Clinical Neurophysiology*, 39:435–436, 1975.



Shouyi Wang received a B.S. degree in Control Science and Engineering from Harbin Institute of Technology, Harbin, China, in 2003, and a MS degree in Systems and Control Engineering from Delft University of Technology, Netherlands in 2005, and a Ph.D. degree in Industrial & Systems Engineering from Rutgers University, Piscataway, USA in 2012. Currently, he is working as a research scientist at the Department of Industrial & Systems Engineering and the Integrated Brain Imaging Center at University of Washington, Seattle, USA. His research interests include data mining and machine learning, pattern recognition, intelligent decision-making systems, and multivariate time series modeling and prediction.

Wanpracha Chaovalitwongse received a B.S. degree in Telecommunication Engineering from King Mongkut Institute of Technology Ladkrabang, Thailand, in 1999 and M.S. and Ph.D. degrees in Industrial & Systems Engineering from University of Florida in 2000 and 2003. He worked as a Post-Doctoral Associate in the NIH-funded Brain Dynamics Laboratory, Brain Institute and in the departments of Neuroscience and Industrial & Systems Engineering at University of Florida. He also worked for one year at the Corporate Strategic Research, ExxonMobil Research & Engineering, where he managed research in developing efficient mathematical models and novel statistical data analysis for upstream and downstream business operations. He was an Assistant Professor from 2005 to 2010 and an Associate Professor from 2010 to 2011 in the Department of Industrial & Systems Engineering, Rutgers University. Currently, he is an Associate Professor in the Department of Industrial & Systems Engineering and the Department of Radiology at University of Washington.



Stephen Wong received a B.S. degree in 1996 at the California Institute of Technology in Pasadena, CA. He obtained his M.D. at Duke University in 2000 in Durham, NC, and received subsequent medical training in neurology and epilepsy at the University of Pennsylvania Health System in Philadelphia, PA. He was the recipient of scholarship for a Masters Degree in Translational Research, which he received in 2010, during which he studied signal processing and machine learning methods applied to clinical neurophysiology and electroencephalography. He is currently an assistant professor of neurology at UMDNJ-Robert Wood Johnson Medical School in New Brunswick, NJ, where he provides clinical care to patients with epilepsy, teaches medical students and residents, and performs computational research related to event detection in neurophysiology.