IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 18, 2020, accepted September 28, 2020, date of publication October 1, 2020, date of current version October 13, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3028139

Empirical Evaluation on the Impact of Class Overlap for EEG-Based Early Epileptic Seizure Detection

YUBIN QU[1,6], (Member, IEEE), XIANG CHEN[2], (Member, IEEE), FANG LI[3], FAN YANG[4], JUNXIA JI[5], AND LONG LI[1], (Member, IEEE)

[1]Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology, Guilin 541004, China
[2]School of Information Science and Technology, Nantong University, Nantong 226019, China
[3]School of Civil Engineering, Jiangsu College of Engineering and Technology, Nantong 226001, China
[4]Library and Information Center, Jiangsu College of Engineering and Technology, Nantong 226001, China
[5]Department of Geriatrics, Affiliated Hospital of Nantong University, Nantong 226001, China
[6]School of Information Engineering, Jiangsu College of Engineering and Technology, Nantong 226001, China

Corresponding author: Junxia Ji (yangfan@jcet.edu.cn)

This work was supported in part by the Nantong Science and Technology Project under Grant JC2019106, in part by the Research Topics on Education Informationization in Universities under Grant 2019JSETKT064, in part by the Scientific Research Projects of Jiangsu College of Engineering and Technology under Grant GYKY/2019/9, and in part by the Guangxi Key Laboratory of Trusted Software under Grant kx202013.

ABSTRACT Important physiological information is hidden in electroencephalography (EEG), which can reflect the human brain's activity. EEG, which is a kind of complicated signal, can be used for epileptic seizure detection and epilepsy diagnosis via machine learning. A large amount of effort, including raw signal preprocessing and data preprocessing for machine learning, is required for constructing high-quality training datasets because the classification performance highly depends on high-quality data. Feature extraction has been widely used in EEG-based early epileptic seizure detection. Due to the complexity of data collection and labeling, some of the training instances are inevitably mislabeled. That means some similar instances have different labels. This is called the issue of class overlap, which leads to a poor class boundary for classification models and makes constructing a high-quality classification model more difficult. However, the previous studies investigating the impact of the class overlap for EEG data is quite limited. Our goal is to investigate the impact of the class overlap on EEG-based early epileptic seizure detection. We propose a special neighborhood cleaning rule (SNCR) to solve the class overlap issue. To alleviate the class overlap issue, we conduct large-scale experiments on two widely-used EEG datasets and compare our proposed SNCR strategy with a state-of-the-art data clean strategy, i.e., the improved k-means clustering cleaning approach (IKMCCA). The experimental results show that the classification model can achieve significantly better performance in terms of AUC, recall, and F1 metrics when using our proposed SNCR strategy. Therefore, for EEG-based early epileptic seizure detection, we recommend researchers to apply the SNCR strategy to mitigate the class overlap issue and use the SNCR strategy to perform data preprocessing in a future related study.

INDEX TERMS EEG, early epileptic seizure detection, class overlap, class imbalance, empirical evaluation.

I. INTRODUCTION

Small metal discs (electrodes) are attached to the scalp to detect the brain's activity, which is called electroencephalography (EEG). This method has been widely used in clinical domain [1]. It can promote brain science research from the medical perspective by acquiring the mapping relationship between brain information and behavioral information. Due to the rapid development of the Internet of things (IoT) technology, EEG information can reflect the relationship between brain activity information and behavior information. Human brain cells communicate via electrical impulses and are active all the time, even during sleep. This activity shows up as wavy lines on an EEG recording, which can be collected

The associate editor coordinating the review of this manuscript and approving it for publication was Zhaojun Li.

using IoT. There are many related studies based on the EGG, including sleep pattern recognition and epilepsy [2]. An EEG is one of the main diagnostic tests for epilepsy. 1% of the world's population is affected by epilepsy, which can have multiple effects on the human body, such as memory loss, depression, and other psychological symptoms [3]. Epilepsy is associated with brain disorders and involves recurrent, unprovoked epileptic seizures resulting from the abnormal firing of cortical neurons, recruiting neighboring cells into a critical mass [4]. Therefore, it is necessary to detect epilepsy as early as possible and respond to changes in brain waves in advance to be able to provide medical care and assistance to patients on time to prevent malignant results.

The detection of epilepsy can start with brain waves, which are different from regular brain activity. A large number of methods based on pattern recognition can analyze and model the brain waves of patients from a statistical point of view to predicting possible epileptic seizures in advance. Many machine learning methods have been used to characterize the dynamic behavior of EEG signals, like the linear model [5], logistic model [6], Gaussian model [7], and deep learning [8]. Feature extraction has been widely studied in EEG-based early epileptic seizure detection. Some of the training instances are inevitably mislabeled due to the automatic data collecting method and a small sampling interval. There may be some similar instances with different labels, which is called the class overlap issue. High-quality data is required for constructing a high-quality classification model. However, the Class overlap issue has not been investigated in previous studies for EEG-based early epileptic seizure detection.

Similar instances may overlap densely in the space based on different features. The class overlap issue has been investigated in other application domains, such as software defect prediction [9]. That is to say, for EEG-based early epileptic seizure detection, different epilepsy seizures may have the same feature. The instances at the intersection of vector space cause the class overlap issue. These instances resent a serious challenge to the classification model of machine learning.

The class overlap issue can be solved by data preprocessing technology, and after cleaning the data, high-quality training data can be provided for the classification model. In the previous related studies, the class overlap has been investigated in many application areas, including credit card fraud [10], text classification [11], and software defect prediction [12]. Moreover, the class imbalance problem often accompanies the class overlap issue. The current commonly used strategies include the neighborhood cleaning rule learning (NCR), and the improved k-means clustering cleaning approach (IKMCCA) [9]. The NCR method removes the conflicting majority instances to solve the class overlap issue, while the minority instances are not processed to achieve the balance between the majority class and the minority class [12]. The IKMCCA method is based on the standard k-means algorithm. For each cluster, the majority instances and the minority instances are eliminated according to the ratio between the minority instances and the majority instances [9].

In this paper, we propose a novel neighborhood cleaning rule (SNCR) strategy. This strategy is divided into three stages, considering data oversampling and NCR. The motivation for this strategy is that, intuitively, the EEG dataset has a large amount of data, and the problem of the class overlap is inevitable. Therefore, the class overlap is likely to exist for each seizure type.

In our empirical studies, we design the following two research questions (RQs):

RQ1: How is the prediction performance affected by the class overlap problem of epilepsy seizures?

RQ2: Which classification model performs best on epilepsy seizures in terms of different performance measures?

To achieve an objective estimation of the class overlap issue, we conduct the experiments on two widely used EEG datasets and compare our proposed SNCR strategy with a state-of-the-art data clean strategy [9]. Performance evaluation measures (i.e., AUC, recall, and F1) are used to compare the performance of different strategies.

In this study, we aim to identify and remove overlapping instances and find a crosponding effective method for epilepsy seizures. In summary, the contributions of this paper can be summarized as follows:

- To the best of our knowledge, we are the first to investigate the impact of the class overlap problem on epilepsy seizures.
- We are the first to investigate how the class overlap problem influences the prediction performance on epilepsy seizures.
- We are the first to propose our SNCR strategy for the class overlap problem on epilepsy seizures.
- Empirical results on two real-world datasets show the effectiveness of our proposed SNCR strategy.

The rest of this paper is organized as follows. Section II introduces the background of EEG and previous studies on the class overlap problem in machine learning. Section III describes the method in detail, including EEG data preprocessing and data cleaning strategies. Section IV reports our experimental setup, including experimental subjects, performance evaluation measures, strategies for experimental comparison, and experimental design. Section V discusses the results of our experiments. Section VI analyzes the potential threats to validity for our empirical results. Section VII concludes the paper with some future work.

II. BACKGROUND AND RELATED WORK

In this section, we mainly discuss the related studies on EEG-based early epileptic seizure detection and the class overlap issue.

A. ELECTROENCEPHALOGRAPHY

The brain-computer interface is a technology used to obtain information from the user's brain, control external designs, or communicate. The data can reflect the information or command that the user wants to send. The signal processing tool uses electrodes and other methods to identify the information or command and send it to the corresponding output device. The current four common brain-computer interfaces include EEG, electrocorticography, deep electrodes, and functional magnetic resonance imaging. EEG is a micro-current detection technology that detects the activity in the brain through the measurement of micro-currents. This technology is a non-invasive detection technology. Its implementation is equipped with contact electrodes on the scalp of the brain. Multiple electrodes record the patient's brain wave activity overtime on the scalp in many medical fields.

Currently, research on user intentions using EEG technology is still evolving and continuing. The creation of the user intention model contains three challenging issues. The first problem is how to effectively and reasonably map the user's emotional expression to the labeled state; the second problem is to perform signal denoising, transformation, and other preprocessing on the input data; finally, the third is manual data annotation of the EEG state [13]. The effect of preprocessing methods on downstream EEG has been researched. Although the general structure of the results is similar across these preprocessing methods, there are significant differences, particularly in the low-frequency spectral features and in the residuals left by blinks [14].

EEG-based early epileptic seizure always includes three key steps, as shown in Figure 1. Raw EEG data are first collected using IoT technology. In the first step, these instances are preprocessed, like data normalization. The second step is feature engineering so that the distinguishing features are selected. The third step is constructing the classification model using the preprocessed data.

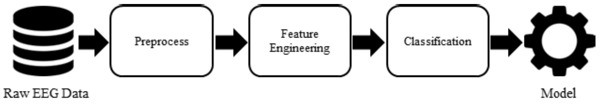


FIGURE 1. EEG-based early epileptic seizure detection framework.

In the original EEG data, due to factors such as errors collected by the device, there may exist data noise and artifacts in the original dataset [15]. Although EEG data is used to record the brain's wave activity, it also records some other weak currents. These noise instances are called artifacts and must be preprocessed using two common techniques, including physiologic and extra physiologic artifacts cleaning technology.

For EEG-based early epileptic seizure classification, two popular methods have been used to preprocess EEG data from TUH EEG Seizure Corpus [16]. The fast Fourier transform (FFT) method has been used in the TUH dataset [17]. The FFT preprocessing technology for the TUH dataset is shown in Figure 2. For non-periodic signals, discrete Fourier transform based on discrete signals can meet the requirements of signal processing. However, only handle discrete and finite-length data can be handled, so here we use FFT in our study.

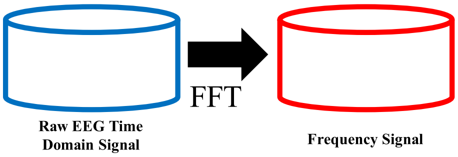


FIGURE 2. FFT preprocessing technology for TUH dataset.

On all electrode channels, we trim the EEG data and sample every s seconds. Then, we use a log function with a base of 10 to process the data at different frequencies. The minimum processing frequency is 1HZ, and the highest is $f(max)$HZ (max means the max sample frequency). Finally, the data is entered into the model as raw data.

B. MACHINE LEARNING-BASED EEG ANALYSIS

With the rapid development of mobile devices, patient information can be collected efficiently and quickly. The status information of these patients can be sensed in real-time through the Internet of things technology and transmitted back to the Internet of things cloud platform. High-quality user data has laid a good foundation for the creation of a patient information management system. Machine learning technology has been successfully applied in many fields, including medical image recognition, cancer diagnosis, and so on. In recent years, there have been reports that the use of optimized machine learning techniques can divide EEG data into normal or abnormal data [17]. Using supervised learning to construct classification models for EEG data has recently played an increasingly important role in EEG diagnosis [18].

The diagnosis and treatment system based on human-machine communication interface technology has been widely used at present. The treatment prediction technology for epilepsy has also been applied [19], and the EEG analysis technology for epilepsy has also been proposed. This new technology is highly innovative and applicable, and it is being accepted by more and more nerve brain scientists [20], [21].

For EEG-based early epileptic seizure detection, the development of miniaturized and standardized equipment has made the monitoring of patients' pre-seizure status more accurate. The automated epilepsy prediction system uses machine learning models to classify EEG data [22]. The classification model uses typical features to distinguish whether there is epilepsy or to predict epilepsy. The feature used in the machine learning model must have a very high degree of discrimination. This feature should be used not only for the status analysis of the same patient in a period but also for different patients' status analysis at different times. Therefore, effective feature engineering technology is essential for EEG-based early epileptic seizure detection.

So far, there has been a series of studies to detect seizures from EEG data. Zandi *et al.* [23] proposed wavelet-transform technology to distinguish seizure or non-seizure states using feature extraction preprocessing technology. Deep learning

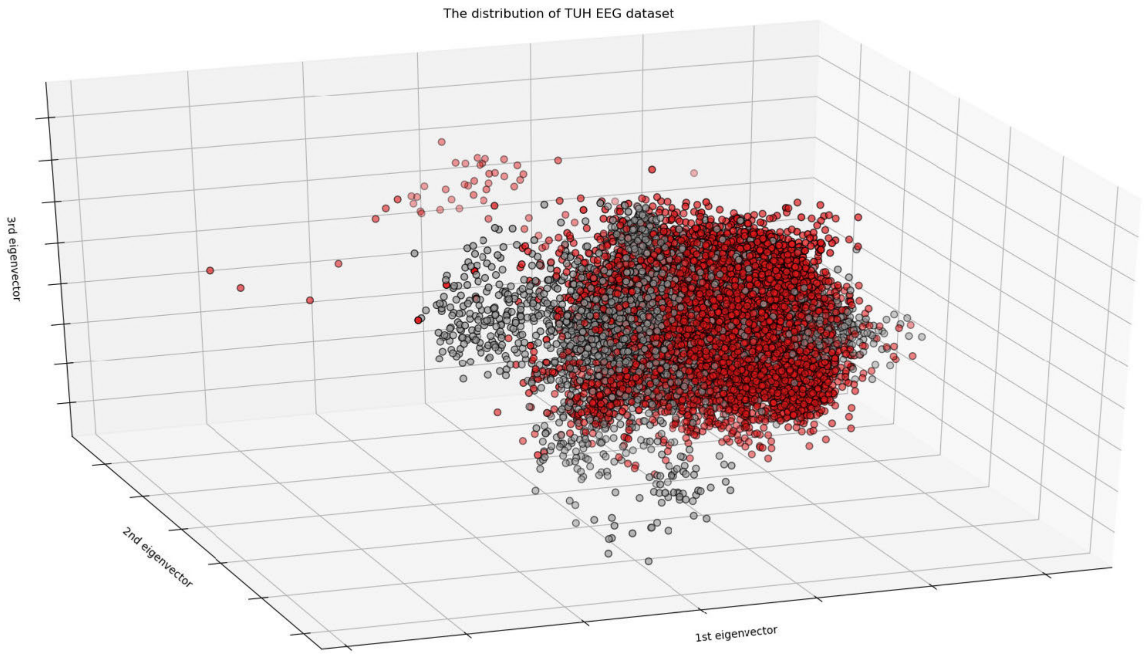


FIGURE 3. The distribution of TUH EEG dataset.

has been used in many fields because of its high performance. Vidyaratne *et al.* [24] used bidirectional recurrent neural networks to extract features for seizure analysis. Unsupervised learning in deep learning, such as autoencoders, has also been introduced to learn features from EEG data for seizure detection [25] automatically.

C. CLASS OVERLAP

The class overlap issue can be described as instances with the same characteristics but with different class labels. The existence of the class overlap issue makes it difficult for the classifier to effectively establish classification boundaries, which significantly affects classification performance, including accuracy, recall rate, and so on. In other fields of machine learning, such as software defect prediction, the class overlap issue is mainly related to the quality of the data or the noise in the samples [12]. Tang and Khoshgoftaar [26] used outlier removal technology to detect potential noise modules and improve data quality, and the experiment revealed that the total error rates decreased with decreasing noise examples. Chen *et al.* [12] proposed a new classification model for software defect prediction that combines class overlap reduction and ensemble imbalance learning. The neighbor cleaning method was first applied to remove the overlapping non-defective samples. The whole dataset was then randomly sampled several times to create an ensemble classification model. Gong *et al.* [9] proposed an improved k-means clustering cleaning approach (IKMCCA) to solve the class overlap issue and the class imbalance problem. The experiment revealed that it is better to consider both the class overlap problem and the class imbalance problem.

To our best knowledge, there is no consideration of class overlap for EEG data. Many instances from the TUH EEG Seizure Dataset are overlapping, as shown in Figure 3, which impact the prediction performance of the constructed models. EEG data comes from an automated data collection system. However, EEG data is subject to current interference from various sources, such as current interference from the collection system itself, abnormal current interference from the body itself, and errors that may occur during data transmission. Therefore, it is essential to perform an overlap analysis of EEG data. At the same time, there is an obvious kind of imbalance in the type of epilepsy. By using noise-cleaning techniques, it is also possible to achieve a balanced sampling of the dataset.

III. OUR PROPOSED METHOD

In this section, we briefly describe the EEG data preprocessing technology and then the whole experimental process, especially data cleaning strategies. Figure 4 provides an overview of the steps in our study. Based on an automated brain wave acquisition system, raw egg data is gathered, and FFT preprocessing is performed on the original dataset to obtain the training dataset. The popular classification models, including random forest (RF), naive Bayesian model (NB), logistic regression (LR), and k-nearest neighbor (KNN), are trained on the training set. Experimental results are gathered based on the test set in terms of AUC, recall, and F1 performance measures.

A. EEG DATA PREPROCESSING

The Fourier transform was firstly used for brain wave analysis in 1932. The successive development of classic analysis

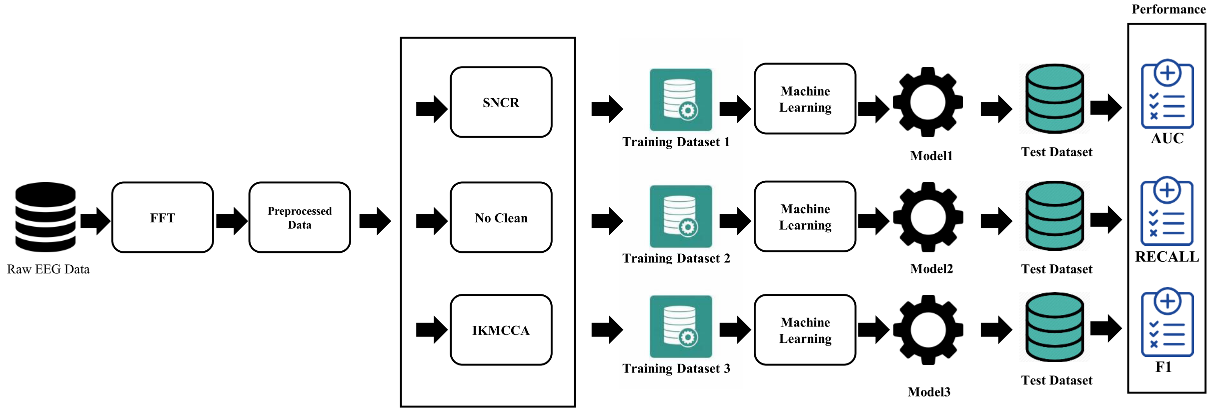


FIGURE 4. An overview of our study.

methods, such as time-domain analysis, frequency-domain analysis, and time-frequency analysis, has effectively promoted the study of brain wave signals [27]. Fast Fourier transform (FFT) can be used to analyze the frequency domain characteristics of the signal, and it is now one of the most popular methods to preprocess EEG data [17], [28].

Fourier transform is derived from the Fourier series by introducing a spectral density function. The calculation process can be defined as follows:

$$F(\omega) = \int_{-\infty}^{+\infty} f(t)e^{i\omega t}dt \tag{1}$$

where t is the time domain, and ω represents the frequency. However, FFT can only be used for the analysis of stationary signals. For non-stationary signals, a short-time Fourier transform (STFT) must be used to perform analysis. For non-stationary signals, the short-time Fourier transform strategy is to add a window to the signal, which is generally a hamming window. Of course, it can also be other types of window functions. The signal after windowing is divided into a set of short-length sequences, and subsequences can be viewed as stationary sequences, which can be analyzed by Fourier transform. The common method of EEG signal analysis using STFT is to use STFT to separate the bands of EEG signals, to obtain the energy of each band as a feature (such as alpha, beta, theta, gamma, and delta).

B. DATA CLEANING STRATEGIES

Since the class overlap issue exists in the EEG dataset, it is essential to preprocess the EEG data. In other fields, like software defect prediction, class overlap issue is often considered as the data quality or noise detection. In our experiment, the special neighborhood cleaning rule (SNCR) and improved k-means clustering cleaning approach (IKMCCA) [9] methods have been used to remove the class overlap instances.

Gong *et al.* [9] improved the k-means clustering cleaning approach for the class overlap issue. This innovative method uses the standard k-means algorithm on the training dataset to cluster the dataset, which is divided into k clusters. For each cluster, they calculate the ratio of the number of defective modules to the number of non-defective modules. If the ratio is higher than the distribution value of the defective modules on the training dataset, they delete all non-defective modules; if the ratio is less than the distribution value of the defective modules on the training dataset, they remove all defective modules on the cluster. Finally, the processed dataset is merged into the final training dataset.

Considering the vast amount of EEG data and the high degree of the class overlap issue found from data visualization analysis, we conjecture that the class overlap problem exists in the current clusters. Therefore, we design a special neighborhood cleaning rule (SNCR). The pseudo-codes for the simulation experiments are provided in Algorithm 1 to evaluate the impact of the class overlap and then answer RQ1 and RQ2. The SNCR algorithm is shown in Algorithm 1.

The motivation for this strategy is that, intuitively, the EEG dataset has a large amount of data, and the problem of the class overlap is inevitable. The class overlap is likely to exist in each seizure type. Therefore, it is unreasonable to solve the problem of class imbalance by undersampling only for most classes. In this study, we think that oversampling should be used to make different types in the datasets to reach the class balance. Moreover, using oversampling can also likely to worsen the class overlap problem.

SMOTE (Synthetic Minority Oversampling technique) [29] algorithm is used to create artificial instances of the minority class. An artificial instance of the minority class x_{i1} is based on the randomly selected x_i, then another neighbor $x_{i(nn)}$ is chosen to calculate the distance between x_i and $x_{i(nn)}$. A randomly selected parameter δ is used to guarantee the randomness.

$$x_{i1} = x_i + \delta \times (x_{i(nn)} - x_i) \tag{2}$$

At this time, the nearest neighbor learning is performed on the current majority class and minority class at the same time, and potential class overlap instances are eliminated. Since the amount of EEG data is relatively large and uses the above nearest neighbor method to find possible class

Algorithm 1 Special Neighborhood Cleaning Rule (SNCR)

```
Input: training set $T = C_{max}, C_{min}$, where $C_{max}$ is the
       majority class, $C_{min}$ is the minority class, and $d$ is
       the ratio r of defective instances to all instances.
Output: a new cleaned training set $T' = \{C'''_{max}, C'''_{min}\}$
for data in $C_{min}$ do
    Choose $k$ neighbors using Euclidean distance;
    Randomly choose a sample $x_{i(nn)}$, and then generate
      a random number $\delta$,$\delta \in \{0,1\}$;
    Generate a new instance $x_{i1} = x_i + \delta \times (x_{i(nn)} - x_i)$;
end
for data in $C'_{min}$ do
    Find the top $N$-nearest neighbors $N_x$ of $x$ according
      to the Euclidean distance;
    if any neighbors $N_x$ in $C_{max}$ then
        $N = N \cup N_x$ ;
        $C'_{max} = C_{max}$-$N$;
    end
end
for data in $C_{max}$ do
    Find the top $N$-nearest neighbors $N_x$ of $x$ according
      to the defined Euclidean distance;
    if any neighbors $N_x$ in $C'_{min}$ then
        $N = N \cup N_x$;
        $C''_{min} = C'_{min} - N$;
    end
end
Define the new input dataset = $\{C''_{min}, C'_{max}\}$ ;
Calculate ratio $\delta = len(C''_{min})/len(C'_{max})$;
Use standard $k$-means algorithm to divide the dataset
  into $k$ cluster;
for cluster in k clusters do
    calculate the new $\delta' = len(C^{(i)''}_{min})/len(C^{(i)'}_{max})$ ;
    if $\delta' > \delta$ then
        delete the minority class instances in current
          cluster;
    end
    if $\delta' \leqslant \delta$ then
        delete the majority class instances in current
          cluster ;
    end
end
Merge the remained instances in each cluster;
```

overlap instances, we can also analyze the current dataset by introducing standard k-means algorithms, then we perform cluster analysis on the dataset and remove the abnormal instances in each cluster. In the k-means algorithm, the distance between each object and the cluster center is calculated using Euclidean distance.

$$d(x, x\prime) = \sqrt{\sum (x - x\prime)^2} \tag{3}$$

For a fair comparison, we set the No Clean strategy as the default data cleaning strategy in our study.

IV. EXPERIMENTAL SETUP

In this section, we first provide the motivation for our research questions. Then, before answering this question, we introduce the experiment setup, including experimental subjects, performance evaluation measures, strategies for experimental comparison, comparative classifications based on machine learning, and experimental design.

A. RESEARCH QUESTIONS

Our study is to evaluate the effect of the overlapping instances on EEG epilepsy seizures. To achieve this research goal, we seek to answer the following two questions:

RQ1: How is the prediction performance affected by the class overlap problem of epilepsy seizures?

RQ2: Which classification model performs best on epilepsy seizures in terms of different performance measures?

RQ1 and RQ2 aim to compare the performance of the existing state-of-the-art learning models by removing overlapping instances in the epilepsy seizure datasets. We studied popular classification models for epilepsy seizures datasets. If the class overlap instance is removed and the classifier's performance is improved, then practitioners can perform corresponding preprocessing on the original EGG data to improve the classifier's performance in future studies on the epilepsy seizures. Besides, by comparing the classifier's performance, it also helps to guide subsequent researchers to choose a classification model suitable for their use.

B. EXPERIMENTAL SUBJECTS

To compare these data clean strategies, we used two publicly available datasets.

The first dataset is the world's largest publicly available dataset of epilepsy seizures, which is published and maintained by Temple University Hospital. We chose the sub-dataset of the TUH EEG Seizure Corpus as our research object. These EEG records are sampled at a frequency of 250 Hz and contain up to 20 electrode channels. The TUH EEG Seizure Corpus contains 2,012 seizure cases, which contain eight different types of epilepsy. Seizure of different patients may be classified into the unified command seizure type. For seizure type classification experiments, we exclude only myoclonic seizures because of the small number of seizures recorded (three seizure events). The seven types of seizure selected for analysis are focal non-specific seizures (FNSZ), generalized non-specific seizures (GNSZ), simple partial seizures (SPSZ), complex partial seizures (CPSZ), absence seizures (ABSZ), tonic seizures (TNSZ), and tonic clonic seizures (TCSZ) [30]. Clinically SPSZ and CPSZ are more specific subclasses of FNSZ, while ABSZ, TNSZ, and TCSZ are more specific subclasses of GNSZ. ABSZ and SPSZ seizure samples are selected respectively to represent one seizure type to test the three strategies.

After preprocessing the EEG dataset for the TUH EEG Seizure Corpus, there are almost 60,000 instances. There are 3,087 instances for the ABSZ seizure type, labeled as "1"; meanwhile, there are 6,028 instances for the SPSZ seizure type, labeled as "0". The ratio of the majority instances to the minority instances is about 1.95. There are 3,087 samples in the original absence seizure (ABSZ) group, and there are 6,028 samples in the original simple partial seizure (SPSZ) group. Thus, there is a clear class imbalance problem. To better test the impact of the class overlap issue on the dataset and cleaning strategy, some noise data was artificially added to the selected data.

The IBM TUSZ preprocessed dataset is inputted into the classification model as the original EEG. For this dataset, the temporal central parasagittal montage preprocessing was performed on 20 electrode channels, and fixed-length windows were used for FFT on all channels. The format of the input dataset is [#data sample, #channels, #frequency bands].

TABLE 1. **Statistics for each seizure type.**

Seizure Type	Seizures	Data Sample
Tonic seizure (TNSZ)	67	4,888
Complex partial seizure (CPSZ)	342	132,200
Generalized non-specific seizure (GNSZ)	415	137,033
Focal non-specific seizure (FNSZ)	992	292,725
Tonic clonic seizure (TCSZ)	50	22,524

The second dataset is another publicly available dataset of epilepsy seizures. The dataset is available on UCI's machine learning repository [31]. The dataset includes 4,097 EEG readings per patient over 23.5 seconds, with 500 patients in total [32].

In the epilepsy seizures dataset on UCI's machine learning repository, there is a total of 11,500 instances, in which there are 2,300 epilepsy seizure instances labeled as "1". The remaining instances are labeled as "0". Therefore, this dataset also contains a class imbalance problem.

C. PERFORMANCE EVALUATION MEASURES

To investigate the impact of the class overlap on the performance of the constructed models, we consider three performance measures: the area under the receiver operating characteristic curve (AUC), recall, and F1-measure (F1).

TABLE 2. **Confusion matrix.**

	Predicted Positive	Predicted Negative
Actually Positive	True Positive	False Negative
Actually Negative	False Positive	True Negative

In Table 1, we can find that there is a significant class imbalance in TUSZ datasets. Due to the imbalance distribution, multiple performance measures are usually adopted to evaluate different aspects of constructed prediction models. We measure the performance with F1-measure and AUC, which have been widely used for the class imbalanced datasets. For a binary classification problem, an unambiguous way to present the prediction results of a constructed classifier is to use a confusion matrix.

$$precision = \frac{TP}{TP + NP} \tag{4}$$

$$recall = \frac{TP}{TP + FN} \tag{5}$$

$$F1\text{-}measure = \frac{2 \times (precision \times recall)}{precision + recall} \tag{6}$$

AUC is defined as the area enclosed by the ROC curve and the coordinate axis. The maximum value cannot exceed 1. The closer the AUC value is to 1, the higher the authenticity of the classifier detection. Conversely, when it is close to the minimum value of 0.5, it represents that there is no application value. F1 is the harmonic mean of precision and recall, and this performance measure can solve the trade-off between precision and recall.

To statistically evaluate the detailed prediction results, we first employ the Friedman test to determine whether there are statistically significant differences among compared methods. If there is a statistically significant difference, the post-hoc Nemenyi test is applied to compare the difference.

When the null hypothesis is rejected, the average rank should be calculated and is compared with the critical distance (CD).

$$CD = q_a \times \sqrt{\frac{k \times (k+1)}{6N}} \tag{7}$$

In our experiment, k represents 12 different algorithms, and N represents all 20 training datasets. q_a is defined as 3.2. Therefore, the result of CD is 2.5799.

In addition, to evaluate the degree of difference among the compared methods in terms of AUC, recall, and F1-measure, we apply Cohen's d to measure the effect size.

$$Cohen's\ d = \frac{M_1 - M_2}{\sqrt{\frac{\sigma_1^2+\sigma_2^2}{2}}} \tag{8}$$

where M_1 and M_2 represent the mean of the statistic, and σ represents the standard deviation of the statistic. If $d \in \{0, 0.2\}$, this indicates the effect size is negligible. If $d \in \{0.2, 0.5\}$, this indicates the effect size is negligible. If $d \in \{0.5, 0.8\}$, this indicates the effect size is medium. If $d \in \{0.8, 1\}$, this indicates the effect size is large.

D. STRATEGIES FOR EXPERIMENTAL COMPARISON

To compare the classification performance of the impact of class overlaps on EEG-based early epileptic seizure detection, the special neighborhood cleaning rule (SNCR) strategy is compared with the improved k-means clustering cleaning approach (IKMCCA) strategy. For the sake of fairness, the two strategies for data preprocessing are compared using the case without data preprocessing. This strategy is named the No Clean Strategy.

E. MODELING METHODS

In our experiments, three preprocessing strategies were evaluated and compared with four state-of-art classification algorithms. The details of the models are introduced as follows:

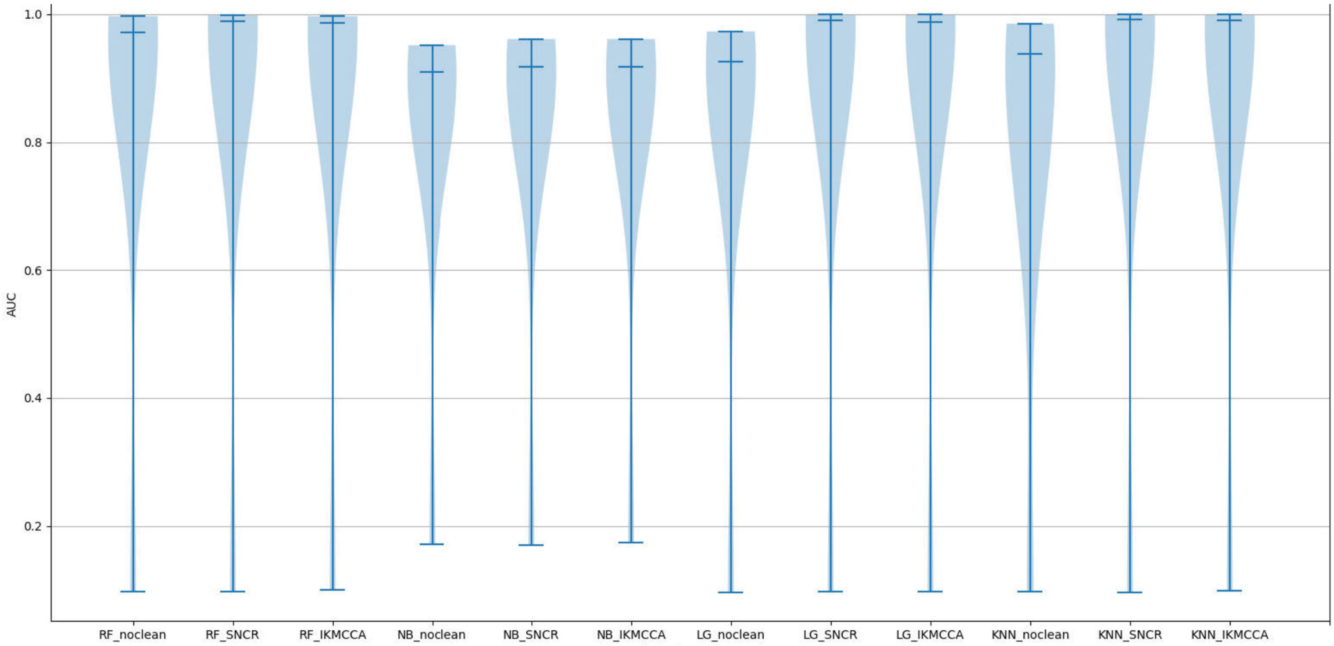


FIGURE 5. The comparison results via violin plot for the two datasets in terms of AUC.

RF. Random forest (RF) is a common classification model that has been widely used in many machine learning fields. This is an integrated learning classification model, which creates a series of decision trees by randomly dividing data and uses the voting results of the decision tree for classification. This model has a strong processing capacity for imbalanced datasets.

NB. The naive Bayesian model (NB) is a simple but robust classification model. This model is based on the Bayesian principle and has been proven to have better classification results than complex models in multiple application areas, such as support vector machine models.

LR. Logistic regression (LR) is a variant of the regression method that is essentially a linear classifier. This model has strong interpretability, and the fitted parameters can represent the impact of each feature on the result.

KNN. The k-nearest neighbor (KNN) algorithm is relatively mature in theory. It considers the k nearest samples to a certain example, and the voting result of most examples determines the example category.

F. EXPERIMENTAL DESIGN

To evaluate the model performance of different query strategies, we first divide the preprocessed dataset into m groups and then use random stratified sampling to generate the instances training set and test set.

The experiment is performed m times ($m = 10$). In our experiments, we use the implementation of these classifiers provided by scikit-learn to avoid internal threats to validity and use the default value for the classifier's hyperparameters. The pseudo-code for the experimental setup is shown in Algorithm 2.

Algorithm 2 Steps of the Experimental Setup

```
Input: dataset $T$ = {IBM TUSZ preprocessed dataset,
       epilepsy seizures dataset on UCI's machine
       learning repository } learning models = {RF,
       NB, LG, KNN} strategies = {No Clean, SNCR,
       IKMCCA }
Output: metrics = { AUC, recall, F1 }
for data in dataset do
    Randomly stratify the current dataset into $m$ folds;
    Define training set and test set  for  classifier in
    learning models do
        for  strategy in strategies do
            train the classifier on the training set;
            report performance  AUC, recall, F1 on test
            set;
        end
    end
end
```

V. EXPERIMENTAL RESULTS

In this section, we report experimental results for the comparison with and without removing the overlapping instances to answer RQ1 and RQ2.

A. RESULT ANALYSIS FOR RQ1

RQ1: How is the prediction performance affected by the class overlap problem of epilepsy seizures?

To answer this RQ, we conduct the experiments on the two EEG seizure datasets via RF, NB, LR, and KNN classifiers by using the SNCR, IKMCCA, and No Clean strategies. In the IKMCCA method, the ratio $p\%$ is set to the ratio of the

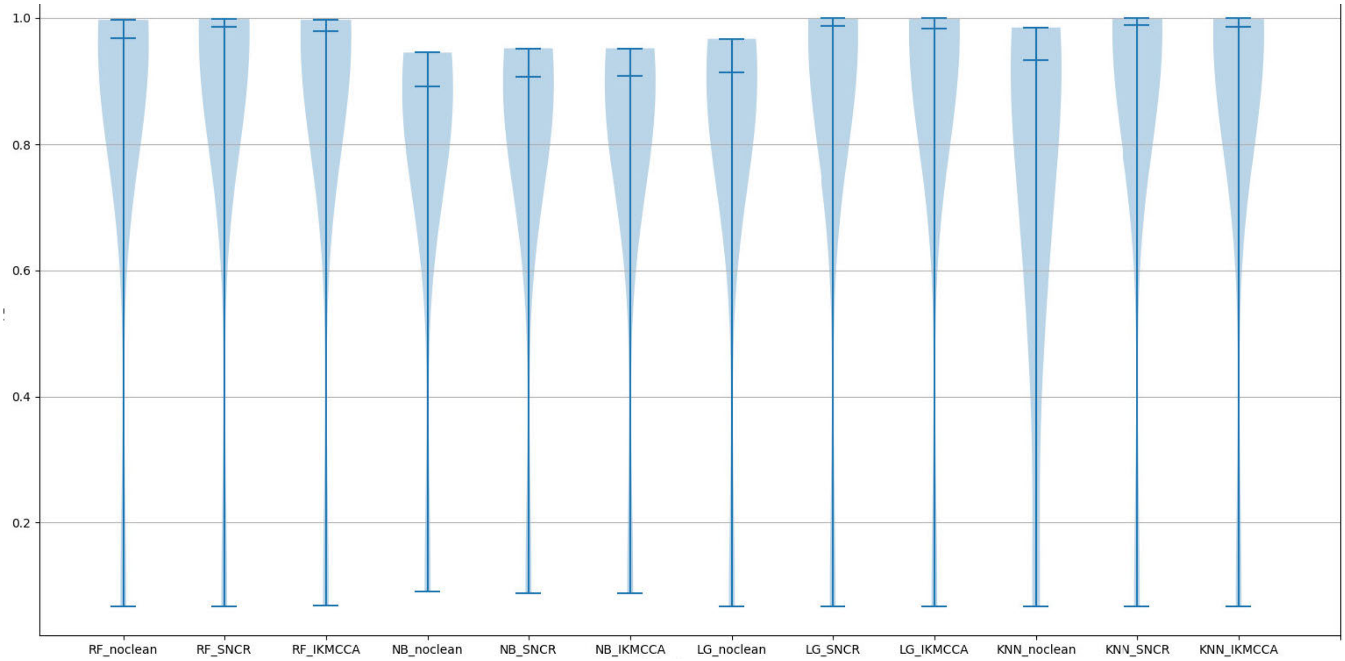


FIGURE 6. The comparison results via violin plot for the two datasets in terms of F1.

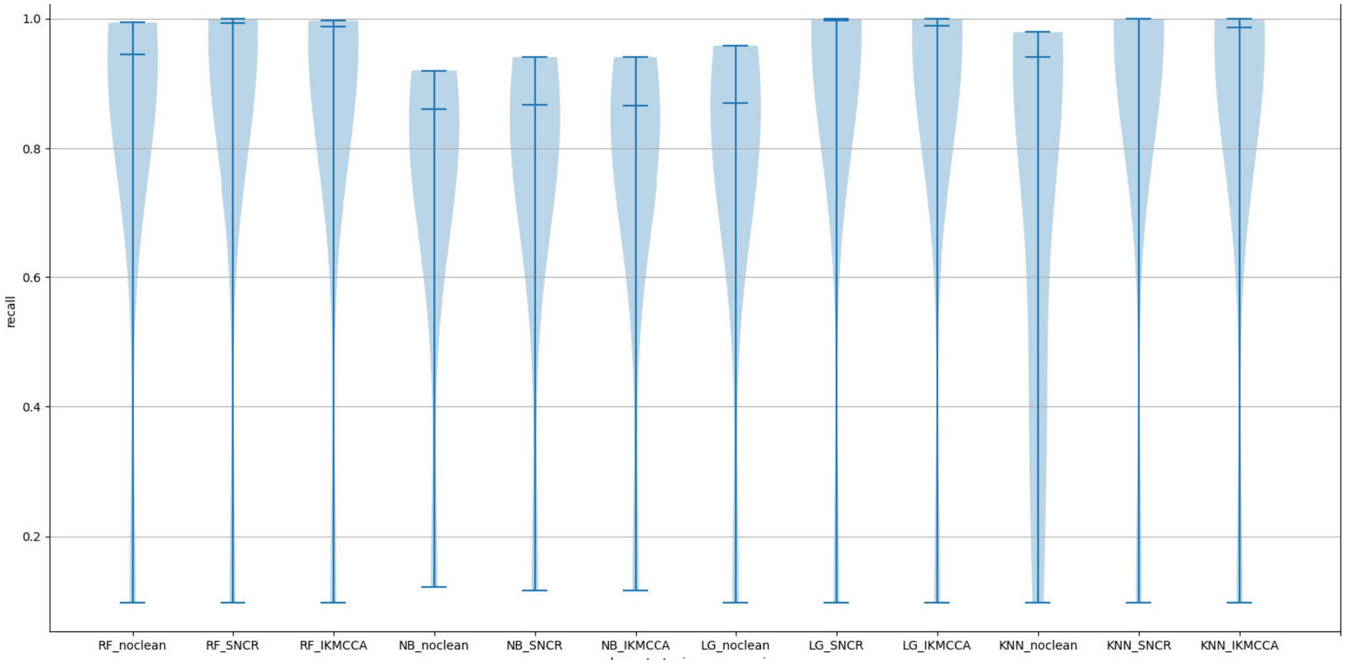


FIGURE 7. The comparison results via violin plot for the two datasets in terms of recall.

number of instances in the minority class to the number of instances in the majority class.

The comparison results via violin plot are shown for each learning model in Figure 5 to Figure 7. From these figures, we can observe that using the SNCR strategy can achieve the best performance (i.e., median value) in terms of AUC, recall, and F1-measure on the RF, NB, LR, and KNN classifiers. In particular, (1) compared with the No Clean strategy, it is better to solve the class overlap problem by using the cleaning strategies, and (2) compared with IKMCCA, the SNCR method performs better in the two epilepsy seizures datasets.

The graphic display of comparison results in terms of different performance measures does not clearly show the differences between different strategies. Then, to compare the performance of different strategies on the difference training datasets from a statistical point of view, the non-parametric Friedman test at a confidence level of 95% is used to conduct a statistical analysis of the results.

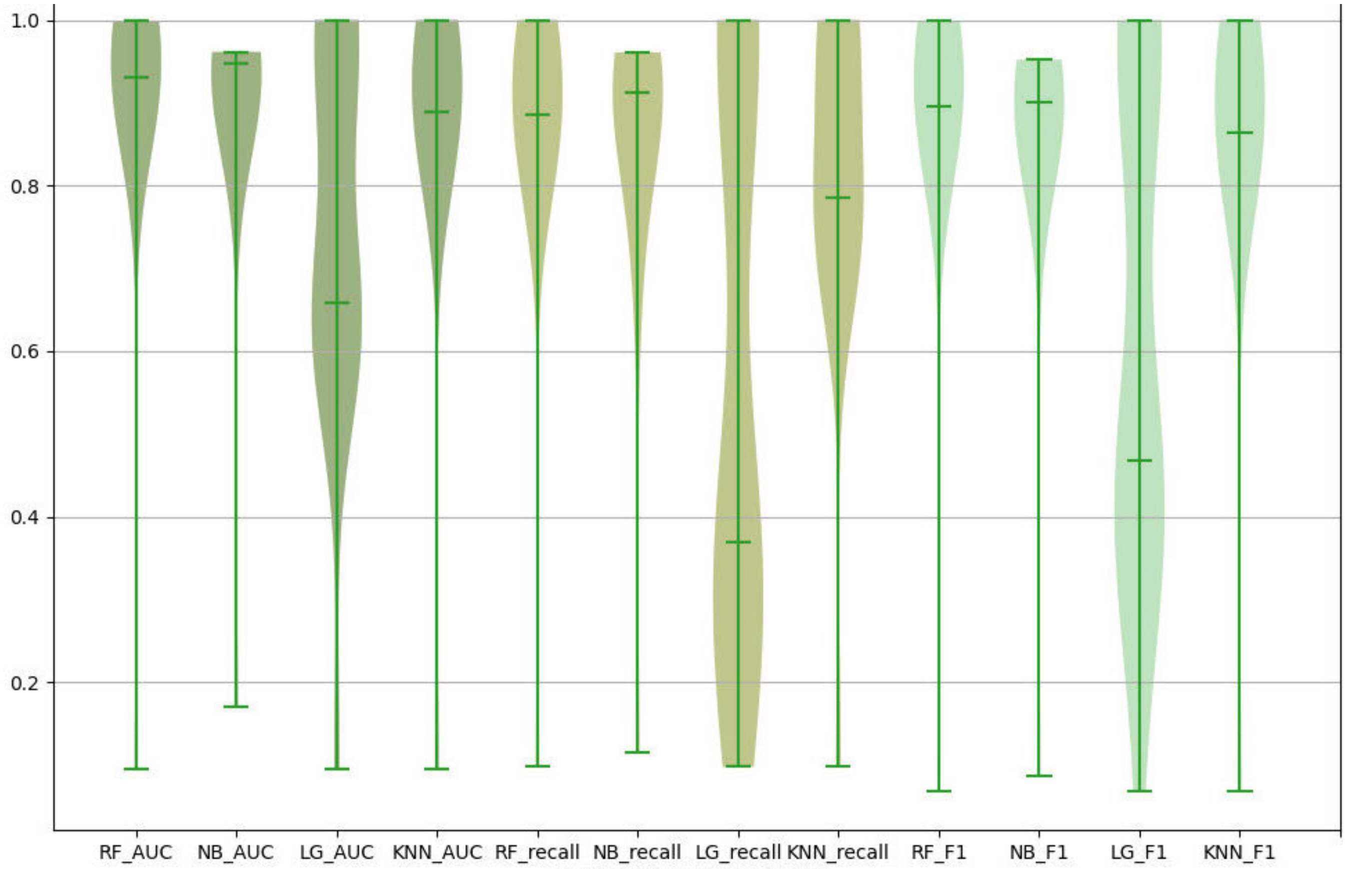


FIGURE 8. The violin plot on SNCR strategy in terms of AUC, recall, and F1.

Firstly, we define the null hypotheses (H0) and alternative hypotheses (H1) as follows:

H0: There is no difference between the strategies on the different datasets.

H1: There is a difference between the strategies on the different datasets.

Secondly, we set the significance level α to 0.05.

Then, we find that the calculated value is smaller than the critical value for a 0.05 significance level. Hence, the null hypothesis is rejected and we can conclude that there is a difference between these three strategies.

To reveal the differences between different strategies, we further adopt a post hoc statistical analysis method. The mean ranks results of 12 approaches in terms of AUC, recall, and F1 are shown in Table 3.

In the end, to further compare these strategies, we compared the effect size of the No Clean strategy with those of the other two strategies. Cohen's *d* effect size is used, and the final results are shown in Table 4.

Summary for RQ1: We can find a statistically significant performance improvement after removing the overlapping instances. Therefore, removing the overlapping instances before building the EEG-based early epileptic seizure detection prediction models is needed. Moreover, SNCR can achieve better results than IKMCCA for all the classifiers. Therefore, we recommend SNCR to consider the class overlap problem when dealing with EEG-based early epileptic seizures.

B. RESULT ANALYSIS FOR RQ2

RQ2: Which classification model performs best on epilepsy seizures in terms of different performance measures?

According to the violin plots of Figure 5 to Figure 7 and Cohn's *d* effect size in Table 4, the SNCR strategy can better deal with the class overlap issue. Then, to answer which classifier has the best classification performance when using the SNCR strategy, we select some of the previous experiments by only focusing on the SNCR strategy in terms of AUC, recall, and F1 for different classifiers (i.e., RF, NB, LG, and KNN). We want to further reveal which of the four classifiers has strong generalization ability and robustness when dealing with the class overlap issue, which is a valuable and meaningful exploration, which can guide other researchers to use a more robust classifier in future studies.

Firstly, we use the violin plots of different classifiers according to the experimental results in terms of different evaluation measures. The violin plots on the SNCR strategy in terms of AUC, recall, and F1 measures are shown in Figure 8.

TABLE 3. The mean rank results of 12 approaches in terms of AUC, RECALL, and F1.

Performance Measure	RF_ noclean	RF_ SNCR	RF_ IKMCCA	NB_ noclean	NB_ SNCR	NB_ IKMCCA	LG_ noclean	LG_ SNCR	LG_ IKMCCA	KNN_ noclean	KNN_ SNCR	KNN_ IKMCCA
AUC	7.375	8.575	6.925	6.3	7.925	7.525	2.175	6.025	5.525	4.85	8.025	6.55
Recall	7.125	8.9	6.75	6.35	7.925	7.325	2.225	6.525	5.4	4.85	8.6	6.025
F1	7.675	8.725	6.975	6.5	7.575	7.475	2.175	6.15	5.525	5	7.625	6.6

TABLE 4. The comparison results between the methods removing the class overlap instances and the methods without removing the class overlap instances in terms of AUC, Recall and F1.

Classifier	Removing VS Without	AUC	Recall	F1
RF	SNCR vs No Clean	0.121	0.295	0.079
	IKMCCA vs No Clean	0.044	0.075	0.067
NB	SNCR vs No Clean	0.073	0.156	0.042
	IKMCCA vs No Clean	0.041	0.074	0.034
LG	SNCR vs No Clean	0.306	0.456	0.453
	IKMCCA vs No Clean	0.126	0.141	0.109
KNN	SNCR vs No Clean	0.349	0.582	0.434
	IKMCCA vs No Clean	0.037	0.110	0.101

TABLE 5. The mean ranks of results of four classifiers in terms of AUC, Recall, and F1.

Measure	RF	NB	LG	KNN
AUC	2.75	3.65	2.125	2.475
Recall	2.675	2.65	1.8	2.675
F1	2.45	3.25	2.125	2.425

According to Figure 8, the NB classifier performs better than other classifiers based on the median value.

Secondly, the non-parametric Friedman test with post-hoc Nemenyi test at a confidence level of 95% is used to conduct a statistical analysis of the results. We can find that there is a difference between different classifiers.

In this scenario, the count of all the training datasets is 20, and the count of compared algorithms is 4. The q_a is queried, and $q_a = 2.569$. The CD is defined as 0.7416 according to Equation (5). The mean ranks results of the four classifiers in terms of AUC, recall, and F1 are shown in Table 5.

In addition, to perform a thorough comparison of the four algorithms, we compare the effect size on the 20 training sets, and the results are shown in Table 6.

TABLE 6. The comparison results among different classifiers in terms of AUC, Recall and F1.

Classifier	Compared classifier	AUC	Recall	F1
NB	RF	0.075	0.131	0.150
NB	LG	0.655	0.897	0.837
NB	KNN	0.059	0.154	0.037
RF	LG	0.692	0.980	0.930
RF	KNN	0.127	0.276	0.108
LG	KNN	0.576	0.770	0.844

Summary for RQ2: After investigating which classifier performs best on EEG-based early epileptic seizure detection datasets for SNCR, we can find a statistically significant improvement in favor of the NB classifier. Therefore, we recommend the NB classifier to build the EEG-based early epileptic seizure detection prediction models in the future.

VI. THREATS TO VALIDITY

In this section, we mainly discuss potential threats to the validity of our empirical study.

A. THREATS TO CONSTRUCT VALIDITY

Only two open datasets are evaluated in our empirical studies. However, the first dataset is the world's largest publicly-available dataset of epilepsy seizures, and the representativeness of our findings can be guaranteed. The second dataset is downloaded from UCI's machine learning repository. This dataset is preprocessed, and its instances have not been transformed by using FFT.

B. THREATS TO INTERNAL VALIDITY

We do not choose all the classification models, which have been considered in the previous studies for EEG-based early epileptic seizure detection. To alleviate this threat, we only choose some representative classification models in our empirical study.

C. THREATS TO EXTERNAL VALIDITY

The two datasets used in our study are free and open datasets. Other commercial and private datasets have not been considered because of intellectual property issues. This may threaten the generalization of our empirical studies.

VII. CONCLUSION AND FUTURE WORK

EEG-based early epileptic seizure detection prediction relies on a large amount of labeled data; classifiers are used to construct models to achieve early detection of an epileptic seizure. However, in actual work, due to the many interference factors encountered in the EEG data collection process, data label errors are inevitable. Some instances have the same measured value but have different seizure types. This kind of error is called the class overlap issue. From the perspective of resolving the class imbalance problem and the class overlap problem, we propose a novel SNCR strategy. Then, we designed experiments to investigate whether using this strategy to solve class overlap can improve the classifier's performance. We conduct the empirical studies to compare the performance using different models on two open datasets. The results show that the SNCR strategy can achieve significantly better performance in terms of AUC, recall, and F1. In other words, the class overlap issue has a performance impact on prediction; it also shows that when removing the class overlap instance, strategies such as oversampling should be considered to solve the class imbalance problem.

This strategy can be used in other application domains, such as software defect predictions, to solve the class imbalance problem [33]. Also, for cross-project software defect prediction, the class imbalance problem can be solved using this SNCR strategy [34], [35]. Finally, the features of EEG data in our study are constructed based on feature engineering in a manual way. In contrast, deep learning can automatically learn semantic features. Then analyzing the impact of the class overlap problem on the leaned semantic features is another interesting problem and can be investigated in our future work.

REFERENCES

[1] K. G. Hartmann, R. T. Schirrmeister, and T. Ball, "EEG-GAN: Generative adversarial networks for electroencephalograhic (EEG) brain signals," 2018, *arXiv:1806.01875*. [Online]. Available: http://arxiv.org/abs/1806.01875

[2] U. R. Acharya, S. Vinitha Sree, G. Swapna, R. J. Martis, and J. S. Suri, "Automated EEG analysis of epilepsy: A review," *Knowl.-Based Syst.*, vol. 45, pp. 147–165, Jun. 2013.

[3] H. U. Ryu, J. P. Hong, S.-H. Han, E. J. Choi, J. H. Song, S.-A. Lee, and J. K. Kang, "Seizure frequencies and number of anti-epileptic drugs as risk factors for sudden unexpected death in epilepsy," *J. Korean Med. Sci.*, vol. 30, no. 6, pp. 788–792, 2015.

[4] Z. Chen, G. Lu, Z. Xie, and W. Shang, "A unified framework and method for eeg-based early epileptic seizure detection and epilepsy diagnosis," *IEEE Access*, vol. 8, pp. 20080–20092, 2020.

[5] J. M. Antelis, L. Montesano, A. Ramos-Murguialday, N. Birbaumer, and J. Minguez, "On the usage of linear regression models to reconstruct limb kinematics from low frequency EEG signals," *PLoS ONE*, vol. 8, no. 4, Apr. 2013, Art. no. e61976.

[6] S.-H. Kim, C. Faloutsos, and H.-J. Yang, "Coercively adjusted auto regression model for forecasting in epilepsy EEG," *Comput. Math. methods Med.*, vol. 2013, Apr. 2013, Art. no. 545613.

[7] Y. Li, W. Cui, M. Luo, K. Li, and L. Wang, "Epileptic seizure detection based on time-frequency images of EEG signals using Gaussian mixture model and gray level co-occurrence matrix features," *Int. J. Neural Syst.*, vol. 28, no. 07, Sep. 2018, Art. no. 1850003.

[8] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: A systematic review," *J. Neural Eng.*, vol. 16, no. 5, Aug. 2019, Art. no. 051001.

[9] L. Gong, S. Jiang, R. Wang, and L. Jiang, "Empirical evaluation of the impact of class overlap on software defect prediction," in *Proc. 34th IEEE/ACM Int. Conf. Automated Softw. Eng. (ASE)*, Nov. 2019, pp. 698–709.

[10] M. Shepperd, D. Bowes, and T. Hall, "Researcher bias: The use of machine learning in software defect prediction," *IEEE Trans. Softw. Eng.*, vol. 40, no. 6, pp. 603–616, Jun. 2014.

[11] Z. Zheng, X. Wu, and R. Srihari, "Feature selection for text categorization on imbalanced data," *ACM SIGKDD Explor. Newslett.*, vol. 6, no. 1, pp. 80–89, Jun. 2004.

[12] L. Chen, B. Fang, Z. Shang, and Y. Tang, "Tackling class overlap and imbalance problems in software defect prediction," *Softw. Qual. J.*, vol. 26, no. 1, pp. 97–125, Mar. 2018.

[13] H. P. Martinez, Y. Bengio, and G. N. Yannakakis, "Learning deep physiological models of affect," *IEEE Comput. Intell. Mag.*, vol. 8, no. 2, pp. 20–33, May 2013.

[14] K. A. Robbins, J. Touryan, T. Mullen, C. Kothe, and N. Bigdely-Shamlo, "How sensitive are EEG results to preprocessing methods: A benchmarking study," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 5, pp. 1081–1090, May 2020.

[15] A. Shoka, M. Dessouky, A. El-Sherbeny, and A. El-Sayed, "Literature review on EEG preprocessing, feature extraction, and classifications techniques," *Menoufia J. Electron. Eng. Res.*, vol. 28, no. 1, pp. 292–299, Dec. 2019.

[16] V. Shah, E. von Weltin, S. Lopez, J. R. McHugh, L. Veloso, M. Golmohammadi, I. Obeid, and J. Picone, "The temple University hospital seizure detection corpus," *Frontiers Neuroinform.*, vol. 12, p. 83, Nov. 2018.

[17] S. Roy, U. Asif, J. Tang, and S. Harrer, "Seizure type classification using EEG signals and machine learning: Setting a benchmark," 2019, *arXiv:1902.01012*. [Online]. Available: http://arxiv.org/abs/1902.01012

[18] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for EEG decoding and visualization," 2017, *arXiv:1703.05051*. [Online]. Available: http://arxiv.org/abs/1703.05051

[19] K. Gadhoumi, J.-M. Lina, F. Mormann, and J. Gotman, "Seizure prediction for therapeutic devices: A review," *J. Neurosci. Methods*, vol. 260, pp. 270–282, Feb. 2016.

[20] K. Das, B. Giesbrecht, and M. P. Eckstein, "Predicting variations of perceptual performance across individuals from neural activity using pattern classifiers," *NeuroImage*, vol. 51, no. 4, pp. 1425–1437, Jul. 2010.

[21] Z. Kurth-Nelson, M. Economides, R. J. Dolan, and P. Dayan, "Fast sequences of non-spatial state representations in humans," *Neuron*, vol. 91, no. 1, pp. 194–204, Jul. 2016.

[22] U. Asif, S. Roy, J. Tang, and S. Harrer, "SeizureNet: Multi-spectral deep feature learning for seizure type classification," 2019, *arXiv:1903.03232*. [Online]. Available: http://arxiv.org/abs/1903.03232

[23] A. S. Zandi, M. Javidan, G. A. Dumont, and R. Tafreshi, "Automated real-time epileptic seizure detection in scalp EEG recordings using an algorithm based on wavelet packet transform," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 7, pp. 1639–1651, Jul. 2010.

[24] L. Vidyaratne, A. Glandon, M. Alam, and K. M. Iftekharuddin, "Deep recurrent neural network for seizure detection," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2016, pp. 1202–1207.

[25] Q. Lin, S.-Q. Ye, X.-M. Huang, S.-Y. Li, M.-Z. Zhang, Y. Xue, and W.-S. Chen, "Classification of epileptic EEG signals with stacked sparse autoencoder based on deep learning," in *Proc. Int. Conf. Intell. Comput.* Springer, 2016, pp. 802–810.

[26] W. Tang and T. M. Khoshgoftaar, "Noise identification with the k-means algorithm," in *Proc. 16th IEEE Int. Conf. Tools Artif. Intell.*, Nov. 2004, pp. 373–378.

[27] L. Liu, S. Li, and Y. Wang, "Eeg signal denoising and feature extraction based on wavelet packet transform," *Comput. Eng. Sci.*, vol. 37, no. 4, pp. 790–795, 2015.

[28] L. D. Mitchell, "Improved methods for the fast Fourier transform (FFT) calculation of the frequency response function," *J. Mech. Des.*, vol. 104, no. 2, pp. 277–279, Apr. 1982.

[29] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, Jun. 2002.

[30] D. Ahmedt-Aristizabal, T. Fernando, S. Denman, L. Petersson, M. J. Aburn, and C. Fookes, "Neural memory networks for seizure type classification," 2019, *arXiv:1912.04968*. [Online]. Available: http://arxiv.org/abs/1912.04968

[31] S. A. Hosseini, M. Akbarzadeh-T, and M. B. Naghibi-Sistani, "Qualitative and quantitative evaluation of eeg signals in epileptic seizure recognition," *Int. J. Intell. Syst. Appl.*, vol. 5, no. 6, p. 41, 2013.

[32] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 64, no. 6, Nov. 2001, Art. no. 061907.

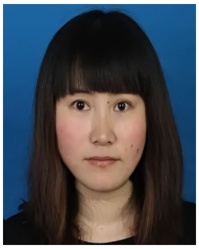
[33] L. Gong, S. Jiang, L. Bo, L. Jiang, and J. Qian, "A novel class-imbalance learning approach for both within-project and cross-project defect prediction," *IEEE Trans. Rel.*, vol. 69, no. 1, pp. 40–54, Mar. 2020.

[34] F. Li, Y. Qu, J. Ji, D. Zhang, and L. Li, "Active learning empirical research on cross-version software defect prediction datasets," *Int. J. Performability Eng.*, vol. 16, no. 4, p. 609, 2020.

[35] Q. Yubin, C. Xiang, C. Ruijie, J. Xiaolin, and G. Jiangfeng, "Active learning using uncertainty sampling and query-by-committee for software defect prediction," *Int. J. Performability Eng.*, vol. 15, no. 10, p. 2701, 2019.

YUBIN QU (Member, IEEE) was born in Nanyang, China, in 1981. He received the B.S. and M.S. degrees in computer science and technology from Henan Polytechnic University, China, in 2004 and 2008, respectively. Since 2009, he has been a Lecturer with the Information Engineering Institute, Jiangsu College of Engineering and Technology. He is the author of more than ten articles. His research interests include software maintenance, software testing, and machine learning.

XIANG CHEN (Member, IEEE) received the B.Sc. degree from the School of Management, Xi'an Jiaotong University, China, in 2002, and the M.Sc. and Ph.D. degrees in computer software and theory from Nanjing University, China, in 2008 and 2011, respectively. He is currently with the Department of Information Science and Technology, Nantong University, as an Associate Professor. In these areas, he has published over 60 papers in refereed journals or conferences, such as the IEEE Transactions on Software Engineering, IEEE Access, *Information and Software Technology*, *Journal of Systems and Software*, the IEEE Transactions on Reliability, *Journal of Software: Evolution and Process*, *Software Quality Journal*, *Journal of Computer Science and Technology*, ASE, ICSME, SANER, and COMPSAC. His main research interests include software engineering, particularly software maintenance, and software testing, such as software defect prediction, combinatorial testing, regression testing, and fault localization. He is a Senior Member of CCF, China, and a member of ACM. He serves as an Associate Editor for IEEE Access.



FANG LI was born in Baoji, China, in 1982. She received the M.S. degree in computer science and technology from Henan Polytechnic University, China, in 2008 and 2011, respectively. Since 2014, she has been a Lecturer with the Jiangsu College of Engineering and Technology. Her research interests include network ideological, political education, and computer application.

FAN YANG was born in Qiqihar, China, in 1980. He received the M.S. degree in computer science and technology from the Nanjing University of Aeronautics and Astronautics, China, in 2013. Since 2011, he has been a Lecturer with the Graphic Information Center, Jiangsu College of Engineering and Technology. His research interests include artificial intelligence and software engineering.

JUNXIA JI was born in Nantong, China, in 1981. She received the M.S. degree in rehabilitation medicine from Nanjing Medical University, China, in 2008. Her research interests include artificial intelligence and brain–computer interface.

LONG LI (Member, IEEE) received the Ph.D. degree from the Guilin University of Electronic Technology, Guilin, China, in 2018. He is currently a Lecturer with the School of Computer Science and Information Security, Guilin University of Electronic Technology. His research interests include cryptographic protocols, privacy-preserving technologies in big data, and the IoT.