Automatika

Journal for Control, Measurement, Electronics, Computing and Communications

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/taut20

Assessment and classification of grid stability with cost-sensitive stacked ensemble classifier

Karthikeyan Ramasamy, Arivoli Sundaramurthy & Durgadevi Velusamy

To cite this article: Karthikeyan Ramasamy, Arivoli Sundaramurthy & Durgadevi Velusamy (2023) Assessment and classification of grid stability with cost-sensitive stacked ensemble classifier, Automatika, 64:4, 783-797, DOI: 10.1080/00051144.2023.2218164

To link to this article: https://doi.org/10.1080/00051144.2023.2218164

Published online: 06 Jun 2023.

AUTOMATIKA
2023, VOL. 64, NO. 4, 783–797
https://doi.org/10.1080/00051144.2023.2218164

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Assessment and classification of grid stability with cost-sensitive stacked ensemble classifier

Karthikeyan Ramasamy[a], Arivoli Sundaramurthy [a] and Durgadevi Velusamy[b]

[a]Department of Electrical and Electronics Engineering, M. Kumarasamy College of Engineering, Thalavapalayam, India; [b]Department of Information Technology, Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam, India

ABSTRACT
Smart Grid is an intelligent power grid with a bidirectional flow of electricity and information, that applies intelligent techniques to operate the grid autonomously near the stability limit. An intelligent technique is developed to identify and predict the abnormalities due to changes in customer behaviour and the unexpected disruption in the grid. A cost-sensitive stacked ensemble classifier (CS-SEC) is proposed for predicting the operations in smart grid that combines four cost-sensitive base classifiers, namely Extreme gradient boosting, Naive Bayes, Nu-support vector machine and Random forest at level-1 and the support vector machine as the meta classifier in level-2. The meta classifier uses the probability of prediction of the first-level classifiers with stratified 5-fold cross-validation to predict the decentralized smart grid stability. The proposed stacked ensemble classifier achieved an accuracy of 98.6% with specificity, recall and precision of 98.34%, 99.0% and 99.06%, respectively. Extensive experimental evaluation and results show that the proposed CS-SEC provides an accurate prediction of grid stability compared with other state-of-the-art models. The results reveal the robustness and competency of the proposed CS-SECs with optimized parameters.

ARTICLE HISTORY
Received 13 August 2022
Accepted 22 May 2023

KEYWORDS
Classification algorithms; machine learning; Smart Grids; stability analysis; support vector machines

1. Introduction

Smart cities are formed by the interconnection of smart technological devices and intelligent equipment integrated using information and communication technologies (ICT). Intelligent devices are a network of related equipment that transmit data using wireless communication technologies like wireless sensor networks (WSNs), wireless mesh networks (WMNs), wireless fidelity (WiFi) technologies and so on. These devices utilize artificial intelligence for automatic decision making to improve the quality of life. However, smart cities [1] require integrating socio-economical and environmental development for sustainable growth and development that includes advanced technologies leveraging ICT to create efficient infrastructures like smart parking, smart grid, traffic management, healthcare and so on. Smart Grid technologies play a vital role in the modernization of the traditional power grid with intelligent techniques that perform data analytics to improve the reliability and availability of electric power for the customers to harness the sustainability and development of society.

Smart grid (SG) is a cyber-physical system that integrates a bidirectional flow of electricity and information and communication technology (ICT) [2]. The smart grid establishes the connection between the producers and consumers with advanced metering infrastructure, wireless networks and high-speed internet technologies for autonomous grid operation. SG also helps to integrate renewable, and non-renewable energy resources [3], electric vehicles with plug-in technologies [4]. The recent development in smart grid infrastructure [5] with digitized metering and advanced communication technologies help the consumer access energy consumption, pricing data and responses accordingly during peak hours of energy demand to reduce or shift their electricity utilization. With effective monitoring and control, the SG can identify and locate faults [6] for the earlier restoration of the grid and to resume its normal operation after the power outages.

The integration [7] of several technologies incurs the technical problems and challenges associated with stability that need to be addressed by SG. The issues like data storage and management [8], cyber-attacks and vulnerabilities [9], power demand and the most prominent challenge of SG is to maintain the stability [10] for effective and reliable operation of the grid. The stable operation of the grid is maintained when the amount of electricity supplied is equal to the amount consumed in electrical grids. It is essential to meet the energy requirement, so the power grid is integrated with renewable and non-renewable resources to

CONTACT Arivoli Sundaramurthy sarivoliapeee@gmail.com Department of Electrical and Electronics Engineering, M. Kumarasamy College of Engineering, Thalavapalayam, Karur-639113, TN, India

balance the demand with supply. Even though the integration of renewable energy [11] plays a critical role in balancing demand with supply in the grid, it affects the grid stability. It requires maintaining the system to operate in equilibrium. The other factors affecting the grid stability are power oscillation [12], power thefts [13] and so on. Hence, it is essential to maintain the grid's stability by analysing the factors affecting the stable operation of the grid. This motivated us to develop an intelligent machine learning model for predicting the grid stability for reliable power transmission, distribution and maintaining the smart grid in a stable state.

Smart grid applies artificial intelligent techniques [14] to optimize energy utilization with real-time monitoring of energy consumption, enhancing reliability and security for resilience operation of the grid. Artificial intelligence, machine learning and deep learning algorithms have gained huge popularity in handling high-dimensional data in recent years. The performance of any machine learning algorithm [15] can be enhanced by selecting important and relevant features that help to discriminate the data into the appropriate category of operation. Feature selection or dimensionality reduction plays a vital role in improving the machine learning algorithm's performance while handling high dimensional data by selecting significant, relevant and consistent features for model evaluation, improving the performance and reducing the computational cost of the model. The following feature selection and dimensionality techniques are applied to choose the features with greater significance for predicting the stability of the grid, namely principal component analysis [16], random forest variable importance measure [17], posterior feature removal [18], forward feature construction [19] and correlation method [20]. The proposed study utilizes the Pearson correlation coefficient measure for selecting a significant and relevant feature from the dataset.

There is no previous study that compared the performance of stacked ensemble learning technique using feature selection for classification of stability in smart grid networks to the best of our knowledge. This study fills the gap by evaluating the performance of stacked ensemble-learning technique. The results of this technique are compared with those of four traditional machine-learning methods, namely the Extreme gradient boosting machine (GBM), Naive bayes (NB), Nu-support vector machine (Nu-SVM) and random forest.

In this study, we have developed four machine learning algorithms, namely extreme gradient boosting algorithm (XGBoost), Nu-SVM, random forest algorithm (RF) and NB for stability analysis. A cost-sensitive stacked ensemble classifier (CS-SEC) is proposed by combining the four base classifiers at the first level and support vector machine as the meta learner at the second level used for predicting the stability of the decentralized smart grid. The dataset is taken from the University of California Irvine (UCI) machine learning repository [21]. The dataset consists of three factors, namely, grid participant's reaction times under varying grid conditions (tau), each participant's electricity generation/consumption volumes (P), and the cost-sensitivity (g) that are used to predict whether the grid is stable or not. The experimental tests for evaluating the performance of the CS-SEC and the other algorithms developed. The results indicate that the proposed CS-SEC model has a high performance with 98.57% accuracy and fewer false predictions. The proposed model achieves a sensitivity of 99.0% and specificity of 98.34% proving with the Area under the curve (AUC) of 0.99, which proves the effectiveness and robustness in better classifying the state of the smart grid operation.

Most studies applied different machine learning methods for classification of system (stable or unstable). However, in some cases, individual machine learning techniques classified the system but with poor performance. One solution to address this issue is to use ensemble-learning techniques. These techniques integrate the decisions of multiple machine learning models to improve the overall classification performance. Using a stacked ensemble learning technique, classification of the system status is found. Their results show that this technique performs much better than any of the individual machine learning techniques, namely Extreme GBM, NB, Nu-SVM and random forest.

The primary significance of this paper can be summarized as follows.

(i) We have proposed a CS-SEC for predicting the stability of the smart grid.
(ii) The significant features are selected using the Pearson correlation co-efficient.
(iii) The CS-SEC combines four heterogeneous classifiers with structural diversity and learning capability, namely XGBoost, Nu-SVM, RF and NB in the first level and SVM in the second level.
(iv) A weight-based cost-sensitive computation is used in the classifiers to reduce the misclassification cost of prediction.

The contributions of this paper can be summarized as follows:

(i) Review of stacked ensemble learning technique
(ii) Use of Pearson correlation feature selection method to identify the most important features
(iii) A comparative analysis of traditional machine learning and ensemble learning methods using the performance metrics such as TPR, FPR, FNR and accuracy.

The flow of the paper is presented as follows: Section 2 describes the previous literature works related to stability issues in smart grids. Section 3 describes the dataset used, the feature selection algorithm and the details about the implementation of the proposed CS-SEC used to predict the stability of the smart grid. Section 4 describes the different evaluation metrics used for measuring the performance of the proposed CS-SEC and the other developed machine learning algorithms. The discussion about the significance of the proposed algorithm and the developed classifiers in the classification of stability of the smart grid is presented in Section 5. In section 6, we summarize the findings with directions for future research.

2. Related work

In recent years, artificial intelligence has a profound application in different domains like health care analytics [22,23], cyber security [24], smart grid [25], social network analysis [26] and so on. Machine learning algorithms are used in power system for a variety of applications like time series classification using convolution neural network [27,28,29], multivariate convolutional neural networks [30] and envelop time series model [31]. The artificial intelligence and machine learning algorithms are applied in different SG applications to predict energy consumption [32], load forecasting [33], stability analysis [34], demand-response prediction [35], fault diagnosis [36], power flow optimization [37] and so on. This section presents a discussion of the state-of-the-art approaches related to the grid stability analysis in smart grid control using machine learning algorithms.

The grid stability analysis is an important measure used to enhance the responsiveness, efficiency and reliable operation of an SG. The parameters associated with the stable operation of the grid are used to develop a data-driven predictive model for a Decentralized Smart Grid Control (DSGC). Researchers have widely used machine learning and deep learning algorithms to evaluate the stability of a grid. The most widely used artificial neural network (ANN), extreme learning machine (ELM) and Random forest assess the grid's stability. The authors in [38] has developed a machine learning model using Artificial neural networks with the Gram-Schmidt orthogonalization technique for significant feature selection to assess the voltage stability of the grid. The authors have used n features to attain a better performance with a minimum error of 0.2391 RMSE in predicting the voltage stability.

In [39], the performance of an artificial intelligence approach based on kernel extreme learning machine is analysed for long-term voltage stability assessment through the measure of MSE and RMSE with the highest estimation error of 0.01%.

A machine learning model is developed in [40] using a decision tree with different sampling techniques like Random Over-Sampling Examples, Synthetic minority oversampling technique, Borderline-Synthetic minority oversampling technique and Adaptive Synthetic Sampling to handle the imbalanced dataset. The decision tree-based machine learning algorithms are applied for assessing the performance of the power system to forecast the short-term voltage stability with non-linear SMOTE to handle imbalanced data. The Decision tree algorithm with the non-linear SMOTE technique achieved an accuracy of 97.78% which is superior to the performance achieved by other sampling methods.

An active machine learning algorithm for monitoring the stability to predict the states of the power system is developed in [41] using three algorithms, namely Random forest, Artificial neural networks and Support Vector Machines. The performance evaluation shows that the RF has relatively better predictive accuracy with 90.01% than SVM and ANN. A multi-directional long short-term memory(LSTM) is applied in [42] for grid stability analysis and prediction. The developed multidimensional LSTM achieved an accuracy of 99%, which is comparatively better than other algorithms like Recurrent neural network, Gated recurrent unit and LSTM.

In [43], the authors used eight base classifiers and gradient boosting decision tree as the meta-classifier to achieve the real-time prediction of rock mass class based on the stacking ensemble classifier. The comparison between the stacking ensemble classifier and other individual classifiers is examined. It demonstrates that for small and unbalanced data, the ensemble learning model outperforms individual classifiers in terms of learning and generalization.

In [44], Algorithms for ensemble learning were employed to examine the stability of 444 slope instances. For the assessment of slope stability, various ensemble learning techniques (Ada Boost, GBM, bagging, extra trees (ET), random forest (RF), hist gradient boosting, voting and stacking) are investigated and contrasted. The tenfold CV approach is used to increase the classification models' capacity for generalization. The stacking model has the highest accuracy (84.74%) and best performance.

The authors in [45] integrate the most recent advances in artificial intelligence theory and suggest a brand-new PV power forecast model based on the stacking method. The analysis of the calculation example demonstrates that the proposed stacking combines the benefits of the single-model prediction algorithm and looks at the data space and structure from various perspectives to allow different algorithms to work in conjunction with one another and produce the best prediction outcomes. In terms of predicting PV power, stacking provides high application value and excellent accuracy.

In [46], through feature selection methods, the stacking based ensemble learning technique was identified to produce improved performance over the other ensemble and machine learning techniques for anomaly intrusion detection in smart grid.

In [47], to identify and categorize GPS spoofing attacks on unmanned aerial vehicles, supervised machine learning techniques like Gaussian Naive Bayes (GNB), Classification and Regression Decision Tree, Logistic Regression, Random Forest, Linear-Support Vector Machine, Artificial Neural Network and Unsupervised machine learning techniques namely Principal Component Analysis, K-means clustering and Auto encoder have been used. In terms of accuracy, probability of detection, chance of misdetection, probability of false alarm, processing time, training time, prediction time and memory size, the performance of various supervised models was compared with that of unsupervised models.

The findings demonstrate that in terms of identifying and categorizing GPS spoofing attempts, the Classification and Regression Decision Tree model performs better than other supervised and unsupervised models.

The application of various ML approaches to improve the reliability, stability, efficiency, security and responsiveness of the smart grid is reviewed in [48]. The authors also discussed the challenges in implementing ML-based solutions in smart grids.

A review of ML application is done in [49] to study the issues related to smart grid safety, reliability, forecasting and energy management.

The authors in [50] developed machine learning algorithms, namely linear regression, random forest, GBM and multi-layer perceptron, to predict the stability of the power system. The author selected significant features using Binary particle swarm optimization. The BPSO-MLP algorithm achieved a classification accuracy of 93.8% in predicting stability and is better when compared to other algorithms.

The authors in [51] developed six machine learning algorithms like linear identification classification, GNB classification, K-Nearest Neighbour (kNN) classification, CART decision tree (DT) classification and AB (Ada-boost classification), RBF SVM (Kernel support vector machine) to predict the stability of the network. The test performance shows that the detection results of NB and support vector machine achieved 97.1% accuracy in predicting grid stability.

The smart grid stability analysis was done in [52] using machine learning algorithms namely SVM (support vector machine), logistic regression (LR), KNN method, NB, Decision Tree (DT), Random Forest (RF), Stochastic Gradient Descent (SGD), Extreme gradient boosting (XGB) classifiers. The experimental test result shows that the XGB classifier outperformed all other classifier models with 97.5% accuracy in finding the stable operation of the grid.

Smart grid plays a role to minimize the power loss during transmission, to predict dynamic needs of customer, utilize the renewable power during peak demand and maintain stability in grid operation. Machine learning helps to predict the stability during dynamically varying power demand and thereby avoids breakdown situation.

In this study, we have proposed a cost-sensitive stacked ensemble model with four base classifier models combining XGB, NB, Nu-SVC and RF to predict and classify the stability of the smart grid control. The comparison results show that the performance of the stacked ensemble is superior to the performance of a single model.

3. Materials and methodologies used

In this section, we briefly discuss the dataset and methods used for the classification of system stability of a smart grid. The following subsection describes the dataset and its characteristics, followed by the data preprocessing technique. Next, the discussion of data partitioning and the proposed ensemble stacking algorithm with cost-sensitive learning is proposed to improve classification accuracy.

3.1. Dataset used

The dataset is taken from the University of California Irvine machine learning repository, prepared by Vadim Arzamasov to illustrate the smart distribution grid stability control. The dataset considered here has input factors related to total energy balance (assumed energy generated or used in each grid area), response time for participants to adjust consumption and/or production in response to price changes, called response time energy price increase and the percentage of price fluctuations, all other factors are equal. The database consists of 10,000 records with 12 feature attributes where the predictable features are response time of power producer, consumer −1, consumer −2, consumer −3, the energy balance of power producer, consumer −1, consumer −2, consumer −3 along with the price efficiency and flexibility of power producer, consumer −1, consumer −2 and consumer −3. The response or target variable is categorical, either stable or unstable situation of the grid. The output variable is approximately labelled as 0 for stable and 1 for unstable of the grid. The in-depth analysis shows that the dataset is highly imbalanced, with 3620 records belonging to stable. The remaining 6380 sample represents the instability nature of the grid, corresponding to a 36.2:63.8 ratio of unbalanced data. The description of the entire dataset is illustrated in Table 1.

Three key factors of the model are:

(i) Pn: Power balance, illustrating the power produced when $n = 1$ or power consumed when $n = 2$, 3 and 4

Table 1. Attribute details of electrical grid stability simulated data set.

S.No	Name of attribute	Nature of attribute	Description of attribute	Type	Min	Max
1	tau1	Input	Reaction time of electricity producer in sec	Numerical	0.5	10
2	tau2	Input	Reaction time of electricity consumer 1 in sec	Numerical	0.5	10
3	tau3	Input	Reaction time of electricity consumer 2 in sec	Numerical	0.5	10
4	tau4	Input	Reaction time of electricity consumer 3 in sec	Numerical	0.5	10
5	p1	Input	Nominal power produced	Numerical	1.5	5.9
6	p2	Input	Nominal power consumed by consumer 1	Numerical	−2	−0.5
7	p3	Input	Nominal power consumed by consumer 2	Numerical	−2	−0.5
8	p4	Input	Nominal power consumed by consumer 3	Numerical	−2	−0.5
9	g1	Input	Gamma coefficient proportional to price elasticity of producer	Numerical	0.05	1
10	g2	Input	Gamma coefficient proportional to price elasticity of consumer 1	Numerical	0.05	1
11	g3	Input	Gamma coefficient proportional to price elasticity of consumer 2	Numerical	0.05	1
12	g4	Input	Gamma coefficient proportional to price elasticity of consumer 3	Numerical	0.05	1
13	stabf	Output	Either unstable class or stable class	Categorical (Discrete)	0	1

(ii) taun : Individual participants reaction time during change in an electricity price

(iii) gn: Price elasticity co-efficient.

3.2. Preprocessing

Data preprocessing is a technique used to enhance the algorithm's performance by transforming the unprocessed or raw data into an appropriate form for efficient handling of data by the machine learning algorithm. Then, the feature selection technique removes highly correlated, insignificant and inconsistent data. Two tasks were performed in data preprocessing, namely normalization and feature selection.

3.2.1. Data normalization

Data Normalization is a technique used to transform or standardize the data to have a similar distribution. The most widely used method for data normalization is rescaling or min–max normalization and z-score normalization. In this study, we have applied z-score normalization, a standardization technique with a mean of 0 and having a standard deviation of 1. This scaling technique transforms the values centred around the average value with unit standard deviation. The z-score normalization is defined as given in Equation (1).

$$E' = \frac{E - \overline{M}}{\sigma_M} \tag{1}$$

where

E' and E are new and old for each data entry, $\overline{M}$ is the mean, and σ_M is the standard deviation.

3.2.2. Feature selection and data splitting

Feature selection plays a crucial role in machine learning applications that helps to improve the performance of the algorithms by eliminating irrelevant, inconsistent, redundant features from the training process. The attribute subset selection enhances the performance and reduces the computational time in the training and testing process. This study has applied the Pearson correlation measure that measures the statistical relationship between the linearly correlated feature attributes. The correlation coefficient between attributes is estimated using the Equation (2).

$$r = \frac{\sum_{i=1}^{n} (a_i - \overline{a})(b_i - \overline{b})}{\sqrt{\sum_{i=1}^{n} (a_i - \overline{a})^2 \sum_{i=1}^{n} (b_i - \overline{b})^2}} \tag{2}$$

where a and b are variables.$\overline{a}$ and $\overline{b}$ are mean of that variables, respectively.

The correlation of attributes before and after using feature selection techniques are depicted in Figure 1(a,b), respectively. After using the feature selection technique, the attributes p1, p2, p3 and p4 are dropped out as it they are least significant.

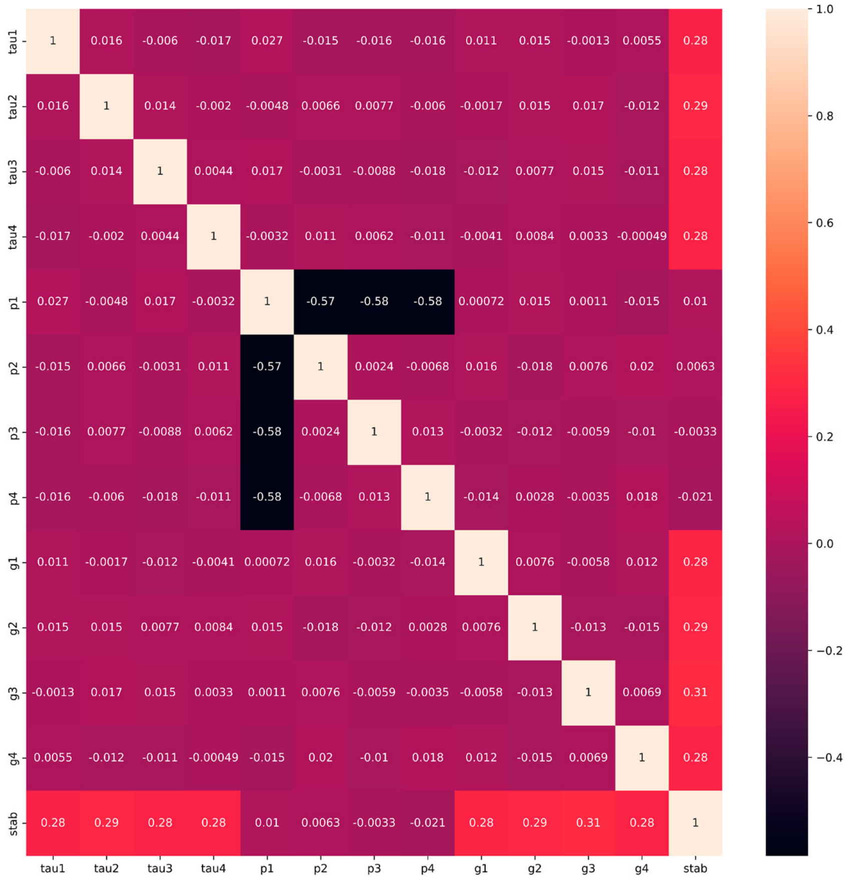
The correlation of attributes before using feature selection techniques is shown in Figure 1(a).

Data Splitting: The dataset consists of 10,000 records split into training and testing data in the ratio of 90% training and remaining for testing the machine learning model. The model is trained using the training dataset, and the performance of the trained model is evaluated using the unseen test dataset. We have applied a stratified 5-fold cross-validation technique to balance the classes in training and data validation. In 5-fold cross-validation, the training dataset is divided into 5 different folds, and we train the model with K-1 fold (i.e. 4 fold), and the Kth fold is used for validation of the classifier. Finally, the unseen test dataset is used to evaluate the classifier model's ability to predict the stability of the smart grid.

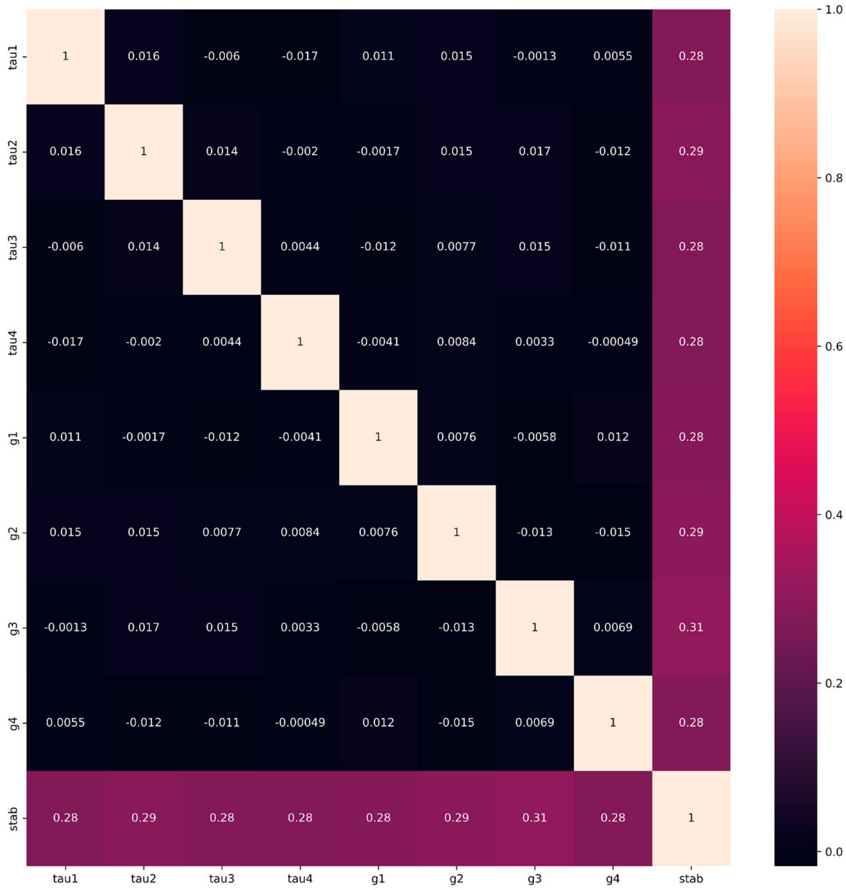
3.3. Proposed stacked ensemble classifier

Let us consider a dataset $D = (x_N, y_N)$ with N observations and m classes having n-feature attributes such that the i^{th} sample $X_i = (x_{i1}, x_{i2}, \ldots, x_{in})$ and $y_i \in \{0, 1\}$. Say $y_i = 0$ implies the stable class and $y_i = 1$ implies the grid is in unstable state. The dataset D is splitted into training T_r and testing T_e dataset.

The proposed stacked heterogeneous ensemble learning model combines four base classifier models, and the predictive results of the base classifiers are given as input to the meta classifier for training the model.



(a)



(b)

Figure 1. (a) Correlation of attributes before using feature selection techniques. The correlation of attributes after using feature selection techniques is shown in Figure 1(b). (b) Correlation of attributes after using feature selection techniques.

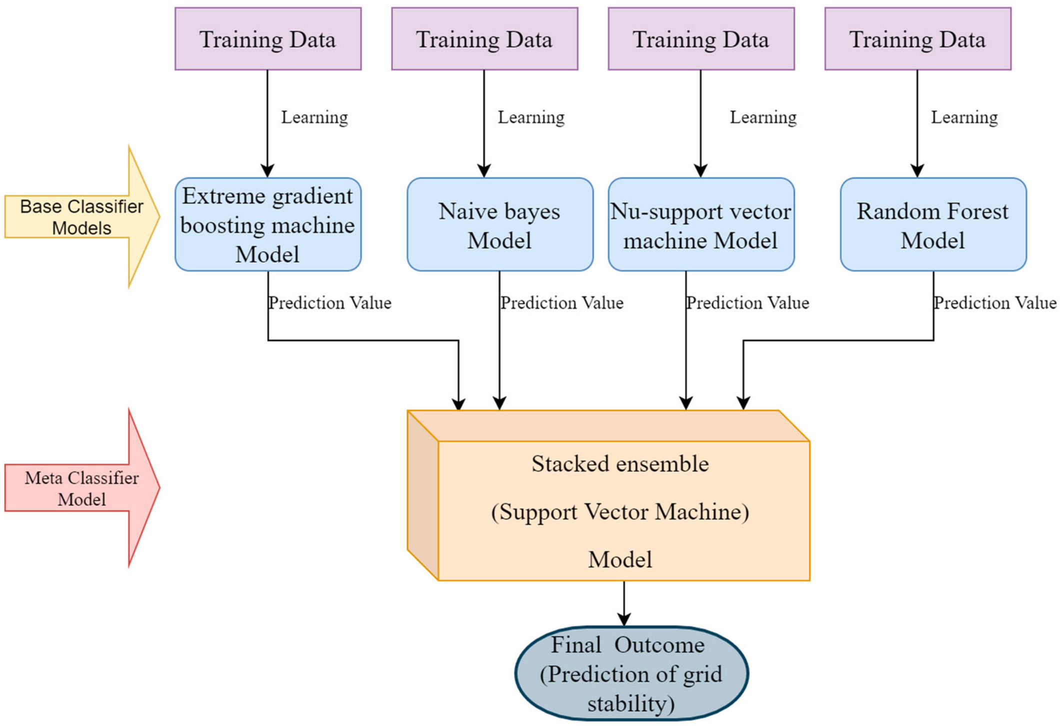


Figure 2. Implementation of meta classifier from base classifier.

The base classifiers of the stacked ensemble model, such as XGB, NB, NuSVC and RF with SVC as meta-classifier, combined to improve the predictions of grid stability when compared with a single model.

Meta-classifier combines multiple classification models by an ensemble learning technique (stacking). The base classifiers (individual classification models) are trained based on a complete training set, then the meta-classifier is trained on the outputs of the base classifiers as features. The advantage of meta-classifier or ensemble models is that suppose XGB gives an accuracy of 95.6%, NB gives an accuracy of 84.7%, NuSVC gives an accuracy of 97.3%, RF gives an accuracy of 90.16%. So the end model will give an accuracy that will be greater than 97.3% which is trained using meta classifier. As the base classifiers often consist of different learning algorithms, the stacking ensembles are often heterogeneous. The algorithm below summarizes stacking.

Algorithm Stacking

1. Input Training Data $TD = \{xi, yi\}_{i=1}^{m}$
2. Output Ensemble Classifier E
3. Step 1: Learn Base-Classifiers
4. for t = 1 to T
5. learn e_t based on TD
6. end for
7. Step 2: Construct New Data Set of Predictions
8. for i = 1 to m
9. $TDe = \{x_i', y_i\}; x_i' = \{e_1(x_i), \ldots\ldots, e_T(x_i)\}$
10. end for
11. Step 3: Learn a meta-Classifier
12. Learn E based on $_{TDe}$
13. Return E

The process of implementation of meta classifier from base classifier is shown in Figure 2. SVM is a supervised machine learning algorithm which uses kernel trick technique to transform the data and then based on these transformations it finds an optimal boundary between the possible outputs. It does some extremely complex data transformations, then figures out how to seperate the data based on the labels or outputs defined. For more complex data points, the boundary that the algorithm calculates doesn't have to be a straight line.

SVM is quite complicated when it comes to what model parameters (Maximal margin, kernels and cost) to pick, if it is used as base classifiers. So SVM can't be a base classifier.

The architecture of the proposed stacked ensemble model is shown in Figure 3. The stacked ensemble model utilized stratified K-fold cross-validation with the probability of prediction results for training the model. We have used a 5-fold cross-validation technique to train and validate the model.

Later the unseen test set is used to evaluate the performance of the model with evaluation metrics like accuracy, sensitivity, specificity, precision, MCCR and AUC measures. In general, stacking of dissimilar classifiers improves the performance of classification. The stacked ensemble model incorporates the underlying structure dissimilarity to enhance the predictive performance of the classifier by exploiting the strength of these classifiers by combining different prediction results.

4. Performance evaluation measure

In this section, we discuss the different predictive performances of the evaluation metrics used to compute

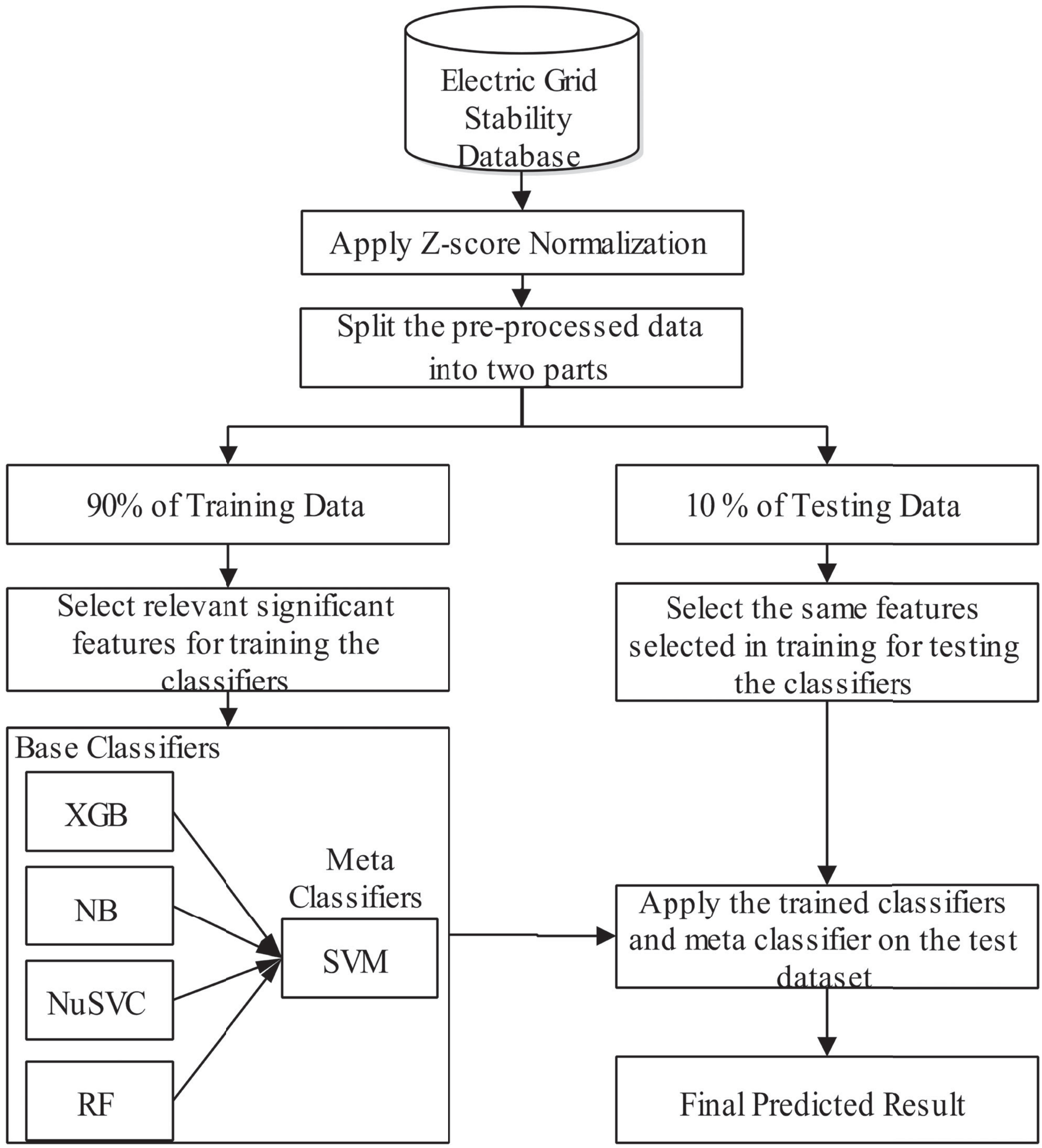


Figure 3. Proposed stacked heterogeneous ensemble learning workflow architecture.

the classifiers. The result of classification is represented in the confusion matrix table with four results: True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). A situation where positive feedback is expected to be a positive category. False negatives are conditions where the actual observation is positive, and the actual outcome is negative results. False positives are conditions where the real observations belong to the negative category but are likely to be predicted as positive. True negativity predicted as truly is the condition in which observations from the negative category are expected to be negative. Performance evaluation in classification is justified by an accuracy classifier. However, the accuracy metric alone is not considered when the data set is imbalanced. Hence, other metrics like precision, recall and specificity are considered to assess the classifier's performance in predicting the grid stability of a smart grid.

Accuracy is the percentage of correct overall predictions across all observations in the dataset. The true recall rate is the accuracy of the projections in the positive category and the percentage of correct predictions for positive observations. In addition to accuracy, Performance evaluation is done with metrics like accuracy, sensitivity, specificity, precision, f1 score, kappa, the AUC and Matthew's correlation coefficient.

Accuracy is the ratio of correctly classified class to that of total class in the database and is mathematically determined using Equation (3).

$$\text{Accuracy(Acc)} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

Sensitivity/True Positive rate and Specificity/Negative positive rate is mathematically estimated using Equations (4) and (5), respectively.

$$Sensitivity(Sens) = \frac{TP}{TP + FN} \tag{4}$$

$$\text{Specificity(Spec)} = \frac{TN}{TN + FP} \tag{5}$$

Precision is the ratio of correctly predicted positive observations to the sum of the predicted positive observations and is mathematically represented as in Equation (6).

$$\text{Precision(Prec)} = \frac{TP}{TP + FP} \tag{6}$$

F1 score is estimated mathematically through the harmonic mean of the precision and recall as in Equation (7).

$$\text{F1Score(F1)} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

Kappa is measured using random classifier output and the accuracy of the algorithm as in Equation (8).

$$\text{Kappa(K)} = \frac{P_0 - P_e}{1 - P_e} \tag{8}$$

where

K is Kappa coefficient value
P_o is total ratio of the main diagonal of the observed frequency
P_e is total of the peripheral ratio of the observed frequency.

Balanced accuracy is defined as the average of recall obtained on each class and is represented in Equation (9).

$$\text{Balanced Accuracy(Bal Acc)} = \frac{\text{Sensitivity} + \text{Specificity}}{2} \tag{9}$$

The AUC is a measure of the classifier's ability to distinguish classes.

(i) If AUC = 1, then the classifier correctly predicts all the Positive and the Negative class
(ii) If AUC = 0, then the classifier predicts all the Positive class as Negative and the Negative class as
(iii) Positive.
(iv) If $0.5 < \text{AUC} < 1$, then true positive and true negative classes are predicted more correctly

Matthews Correlation Coefficient evaluates the quality of the classifier and is defined as the correlation between predicted classes and ground truth. It can be calculated based on values from the confusion matrix as in Equation (10).

$$r = \frac{(TP * TN) - (FP * FN)}{\sqrt{(\mathrm{TP} + FP)(\mathrm{TP}+\mathrm{FN})(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}} \tag{10}$$

4.1. Cost matrix for binary classification

The minor class is designated as a positive class in a binary class task, whereas the major class is designated as a negative class. The confusion matrix is a representation tool that lists both the numbers that were correctly categorized in the data rows as well as the ones that were misclassified. The amount of data rows that have a positive class and are categorized as such is known as the true positive (TP); if they are categorized as a negative class, they are known as the false negative (FN). The amount of data rows with a negative class and that are categorized as such is known as the true negative TN; if they are categorized as a positive class, they are known as the false negative FP. TP, FN, TN and TN each represent a different metric in a cost matrix as shown below.

		1 Actual Positive	0 Actual Negative
1	Predict Positive	True Positive C(1,1)	False Positive C(1,0)
0	Predict Negative	False Negative C(0,1)	True Negative C(0,0)

In datasets, the stable class to unstable class ratio is frequently out of proportion, particularly in the electrical smart grid dataset. The classification model's imbalance class issue caused it to concentrate on the major class (i.e. unstable), which has less significance for decision-making. Modifying the model-building process and the metrics used to gauge the effectiveness of the categorization is the answer to this issue.

4.2. Cost-sensitive metrics

To evaluate the efficiency of the binary classification models, four metrics are used. These metrics are the probability of detection (TPR), probability of misdetection (FNR), probability of false alarm (FPR) and accuracy.

$$TPR = \frac{TP}{(TP + FN)} * 100$$
$$FPR = \frac{FP}{(TN + FP)} * 100$$
$$FNR = \frac{FN}{(TP + FN)} * 100$$
$$Accuracy = \frac{(\mathrm{TP} + TN)}{(TP + TN + FP + FN)} * 100$$

where TP is the number of correct predicted unstable condition; TN is the number of predicted stable condition; FP is the number of incorrect predicted unstable condition; FN is the number of incorrect predicted stable condition

5. Experimental result and discussion

In this section, we discuss the experimental settings and the performance of the classifier models in the prediction and classification of the stable operation of a smart grid. The machine learning algorithms are developed using Google colaboratory in an Intel Core i5-11 35G7 windows 10 operating system operating at 2.40 GHz, having 16 GB main memory and 512 SSD.

The algorithm's performance is measured with accuracy, sensitivity, specificity, precision, f1 score, kappa, balanced accuracy, AUC and Matthews correlation coefficient. The dataset is split into 90 % training, and the remaining is used for testing the classifier. The dataset consists of 10,000 samples with 12 attributes. We have utilized 9000 samples for training consisting of 3620 records in class 0 and 6380 records belonging to class 1. The test data has 1000 records, with 362 data in stable and 638 records in an unstable state.

This study has developed a CS-SEM with four base classifiers, namely XGB, NB, NuSVC and RF. The performance of the proposed CS-SEM with other classifier models is evaluated using the 90% training data set and validated on the unseen test data set with the train-test split approach. The proposed CS-SEM is trained with stratified 5-fold cross-validation. The probability of prediction of the base classifiers on the training dataset and class labels are used by the SVM meta-classifier to identify the stability status of the SDGC as in a stable or unstable state. Then, the unseen test data are used for validating the model in predicting the states of the grid stability. Initially, we have trained the classifier models with all the 12 feature attributes, and the performance of the classifier with different seeds is executed 10 times. The average of the performance testing results is reported in Table 2.

From Table 2, we observe that the performance of the developed ML algorithms in terms of accuracy, sensitivity, specificity, precision, kappa, f1 score, AUC and

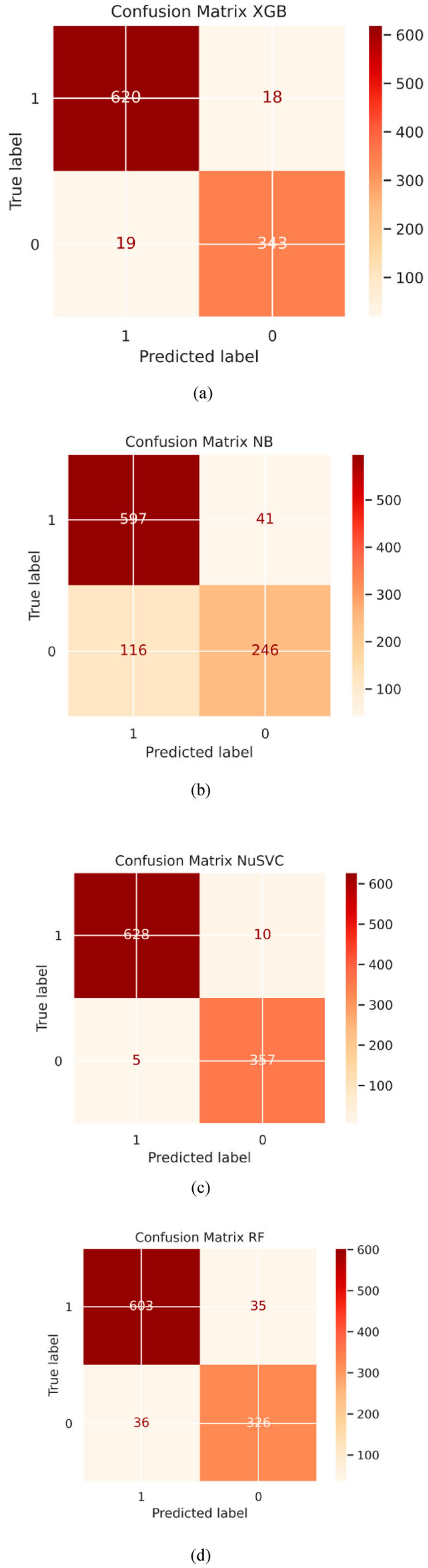


Figure 4. Confusion Matrix for XGB, NB, NuSVC and RF, Confusion Matrix for XGB, (b) Confusion Matrix for Naive Bayes classifier, (c) Confusion Matrix for NuSVC, (d) Confusion matrix for Random forest.

Table 2. Performance of algorithms when all the features are considered.

Classifier	Sens	Spec	Acc	Prec	F1	K	Bal	AUC	MCCR
XGB	0.97	0.93	95.61	0.96	0.97	0.90	0.95	0.95	0.90
NB	0.94	0.69	84.7	0.84	0.89	0.65	0.81	0.81	0.66
NuSVC	0.97	0.97	97.3	0.98	0.98	0.94	0.97	0.97	0.94
RF	0.96	0.79	90.16	0.89	0.93	0.78	0.88	0.88	0.79
SE-SVC	0.98	0.98	97.9	0.99	0.98	0.95	0.98	0.98	0.95

Table 3. Performance of algorithms when significant features are selected.

Classifier	Sens	Spec	Acc	Prec	F1	K	Bal	AUC	MCCR
XGB	0.97	0.94	96.20	0.97	0.97	0.92	0.96	0.96	0.92
NB	0.94	0.68	84.30	0.84	0.88	0.64	0.81	0.81	0.65
NuSVC	0.98	0.99	98.50	0.99	0.99	0.97	0.99	0.99	0.97
RF	0.94	0.89	92.22	0.94	0.94	0.83	0.92	0.92	0.83
SE-SVC	0.99	0.98	98.57	0.99	0.99	0.97	0.99	0.99	0.97

Table 4. Confusion matrix.

		Actual value	
		1	0
Predicted value	1	True Positive	False Positive
	0	False Negative	True Negative

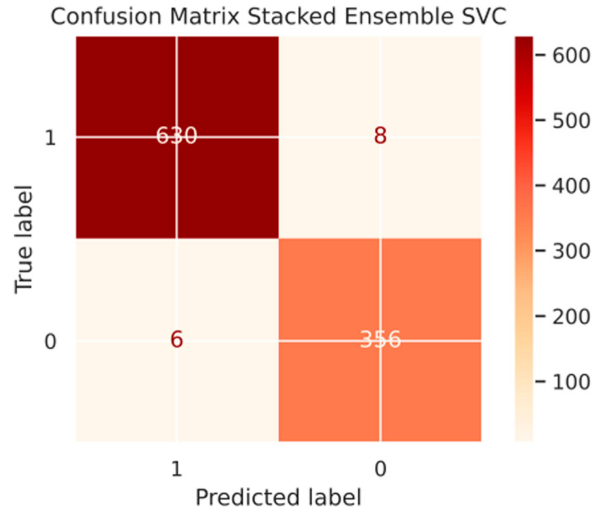


Figure 5. Confusion matrix for stacked ensemble SVC.

MCCR is evaluated and tested. The testing results of the predictive model show an insight that the CS-SEM can perform comparatively better compared to other classifier models.

To analyse and evaluate the classifier's performance and its efficiency in computation is experimented with eight significant features using the statistical Pearson Correlation measure. The experimental evaluation is done with eight highly significant and relevant features executed with different seeds 10 times. The results of the execution are reported in Table 3.

From Table 3, it can be observed that the performance of the developed ML algorithms is evaluated and tested through performance metrics. The testing results of the predictive model show an insight that the CS-SEM is able to perform comparatively better in comparison with other classifier models.

Complete understanding about the performance of algorithm is known from the confusion matrix. For a binary classification problem, we have a 2×2 matrix as shown in Table 4 with four values namely True Positive, False Positive, True Negative and False Negative.

The Confusion matrix obtained for the base classifiers for a single run of the four base classifiers are shown in Figure 4.

The confusion matrix of the proposed stacked ensemble classifier with low values of false negatives and false positive is shown in Figure 5.

It is clear from Table 5 that performance measurements are used to assess and test the effectiveness of the built ML algorithms. The prediction model's testing results provide insight into how much better the CS-SEM performs when compared to other classifier models.

From the table, it is clear that SE-SVC outperformed the other base classifier with 99.01% True Positive Rate, 2.20% False positive rate, 0.94% False Negative rate and 98.6% accuracy.

The receiver operating characteristic curve for the developed classifier models on the given data set is shown in Figure 6.

From the ROC curve, it can be visualized that the proposed CS-SEM performed better with and without feature selection. The AUC is 0.99, which illustrates the effectiveness and robustness of the proposed CS-SEM classifier model.

Table 6 shows the comparison of proposed work with the previous work. The proposed method produced the better result than the previous other methods.

In this subsection, the proposed CS-SEM is compared with other classifier models developed in previous studies. The comparison is made in terms of a number of features selected for model evaluation, such as accuracy, sensitivity, specificity, precision and AUC of the classifiers.

As sensitivity increases, specificity tends to decrease and vice versa. Highly sensitive tests will lead to positive

Table 5. Performance of base classifier and meta classifier based on cost-sensitive metrics.

Classifier	TP	TN	FP	FN	TPR (%)	FPR (%)	FNR (%)	Accuracy
XGB	620	343	18	19	97.03	4.99	2.97	96.3
NB	597	246	41	116	83.73	14.29	16.27	84.3
NuSVC	628	357	10	5	99.21	2.72	0.79	98.5
RF	603	326	35	36	94.37	9.70	5.63	92.9
SE-SVC	630	356	8	6	99.06	2.20	0.94	98.6

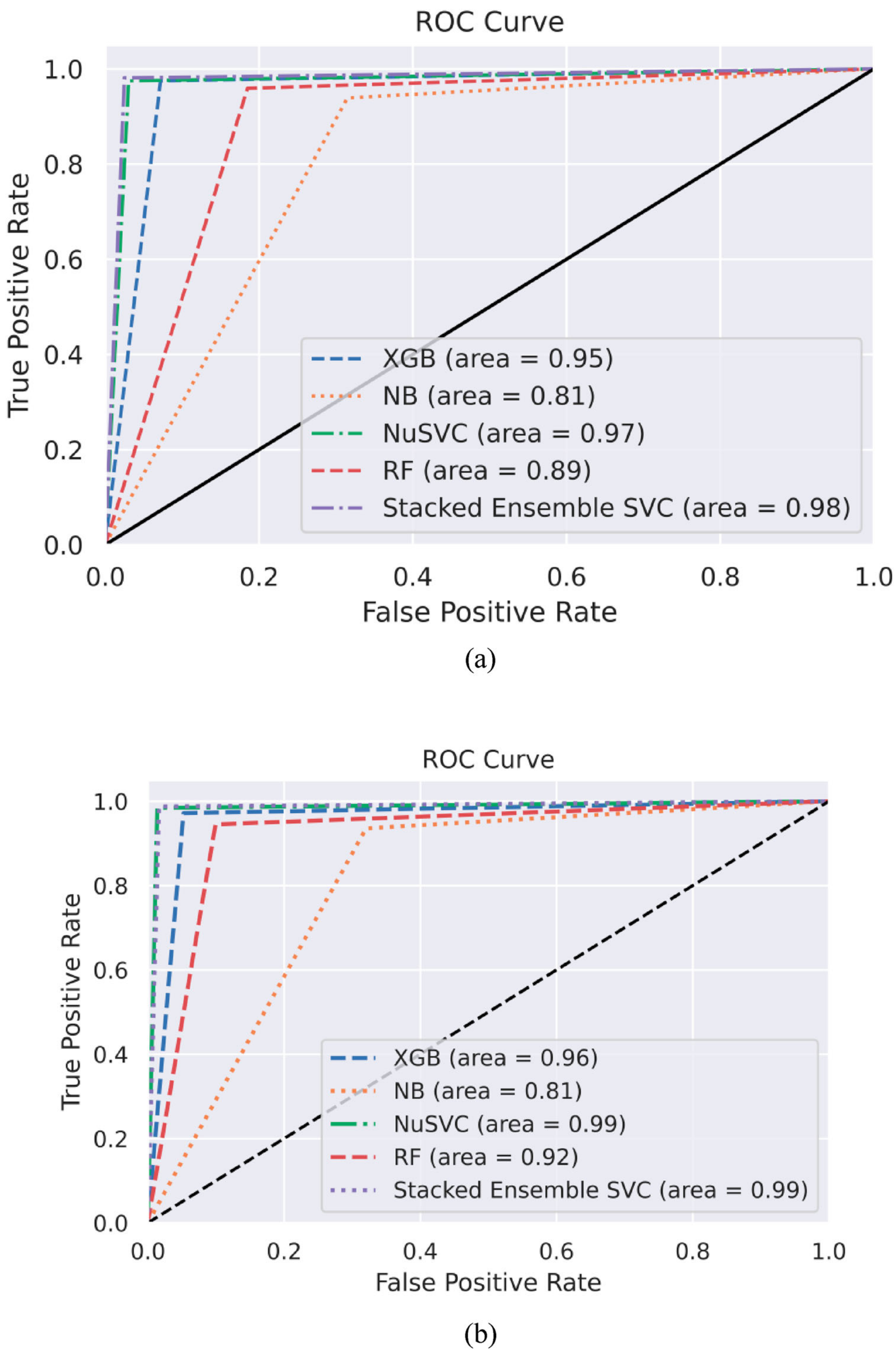


Figure 6. ROC with all features and Feature selection, ROC with all features, ROC with feature selection.

findings, whereas highly specific tests will lead to negative findings.

The proposed study and the previous literature works used the same dataset publicly available in the UCI machine learning repository.

6. Conclusion

The smart grid is the recent advancement in the traditional power grid with a bidirectional flow of electricity and information communication technology.

Table 6. Comparison of previous work with proposed work.

Method	No. of features	Acc.	Sens.	Spec.	AUC	Prec.
Multi-layer perceptron classifier [50]	12	93.8	93.8	93.8	93.8	93.8
RBF SVM [51]	12	97.1	97.4	93.7	99.6	98.1
XGBoost classifier [52]	12	97.5	98.4	95.8	99.9	97.6
Optimizable SVM [53]	13	99.9	99.9	99.9	100	99.8
Optimized DL model [54]	12	97.5	–	–	–	98.6
Optimized artificial neural network [55]	12	97.27	–	–	–	96.79
Ensemblebagging algorithm [56]	12	90	–	–	–	90
Proposed cost-sensitive stacked ensemble classifier	12	98.6	99	98	99	99

In this study, the stability of the smart grid is analysed and assessed to minimize the instability problem in the decentralized smart grid to maintain a stable operation of the grid. The stability analysis provides uninterrupted power to the customers, reduces power losses and meets demand-supply needs. Here a CS-SEM model is proposed to predict the stability of the smart grid. The proposed model combines four ML classifiers: Extreme GBM, NB, Nu-SVM and Random Forest at level-1 with SVM as meta classifiers at level-2 in the stacking model. The performance of the developed model is evaluated using eight significant and discriminant features selected using the Pearson correlation measure. The experimental results show that the proposed CS-SEM has the ability to predict the stability of decentralized smart grid effectively with 98.57% accuracy, 98.6% sensitivity, 98.3% specificity, 99% precision with 98.5% AUC. The proposed model has obtained the best performance result in predicting the stability of the smart grid than other prediction models.

As the extension of future work, the stability analysis of the smart grid can be done with deep learning algorithms to handle large dataset and improve the prediction of grid stability of smart grid.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Arivoli Sundaramurthy http://orcid.org/0000-0002-2927-3956

References

[1] Kaluarachchi Y. Implementing data-driven smart city applications for future cities. Smart Cities. 2022;5(2): 455–474. DOI:10.3390/smartcities5020025.

[2] Kulkarni V, Sahoo SK, Thanikanti SB, et al. Power systems automation, communication, and information technologies for smart grid: a technical aspects review. TELKOMNIKA Telecomm Comp Electron Cont. 2021;19(3):1017–1029. DOI:10.12928/telkomnika.v19i3.16428.

[3] Islam MM, Ali MI, Ceh B, et al. Renewable and non-renewable energy consumption driven sustainable development in ASEAN countries: do financial development and institutional quality matter? Environ Sci Poll Res. 2022;29(23):34231–34247. DOI:10.1007/s11356-021-18488-x.

[4] Darab C. Electric vehicles in smart grids. In: Information Resources Management Association, editor. Research anthology on smart grid and microgrid development, ch054. Hershey (PA): IGI Global; 2022. p. 1192–1211.

[5] Sai Pandraju TK, Samal SRS, Saravanakumar R., et al. Advanced metering infrastructure for low voltage distribution system in smart grid based monitoring applications. Sustain Comp: Inform Sys. 2022;35:100691. DOI:10.1016/j.suscom.2022.100691.

[6] Yugeswar Reddy O, Chatterjee S, Chakraborty AK, et al. Rapid fault identification in standalone DC microgrids. Simul Model Pract Theory. 2022;116:102489. DOI:10.1016/j.simpat.2022.102489.

[7] Erdiwansyah M, Nasaruddin HH, Zaki M, et al. A critical review of the integration of renewable energy sources with various technologies. Protec Cont Mod Pow Syst. 2021;6(1). DOI:10.1186/s41601-021-00181-3.

[8] Zainab A, Ghrayeb A, Syed D, et al. Big Data management in smart grids: technologies and challenges. IEEE Access; 9:73046–73059. DOI:10.1109/ACCESS.2021.3080433.

[9] Olowu TO, Dharmasena S, Hernandez A, et al. Energy, environment, and sustainability. New Res Direct Solar Energ Techn. 2021: 31–51. DOI:10.1007/978-981-16-0594-9_3.

[10] Fan D, Ren Y, Feng Q, et al. Restoration of smart grids: current status, challenges, and opportunities. Renew Sustain Energ Rev. 2021;143:110909. DOI:10.1016/j.rser.2021.110909.

[11] Shahid. Smart grid integration of renewable energy systems. 2018 7th International Conference on Renewable Energy Research and Applications (ICRERA); Paris, France, 2018. p. 944–948. IEEE. DOI:10.1109/ICRERA.2018.8566827.

[12] da Cunha GL, Fernandes RAS, Fernandes TC. Small-signal stability analysis in smart grids: an approach based on distributed decision trees. Elect Pow Syst Res. 2022;203:107651. DOI:10.1016/j.epsr.2021.107651.

[13] Omidvar Tehrani S, Shahrestani A, Yaghmaee MH. Online electricity theft detection framework for large-scale smart grid data. Electr Power Syst Res. 2022;208: 107895. DOI:10.1016/j.epsr.2022.107895.

[14] Omitaomu, and Niu. Artificial intelligence techniques in smart grid: a survey. Smart Cities. 2021;4(2):548–568. DOI:10.3390/smartcities4020029.

[15] Singh R, Singh A, Bhattacharya. A machine learning approach for anomaly detection to secure smart grid systems. In: Research anthology on smart grid and microgrid development, ch.40. IGI Global. p. 911–923.

[16] Omuya EO, Okeyo GO, Kimwele MW. Feature selection for classification using principal component analysis

and information gain. Exp Syst Appl. 2021;174:114765. DOI:10.1016/j.eswa.2021.114765.

[17] Otchere DA, Ganat TOA, Ojero JO, et al. Application of gradient boosting regression model for the evaluation of feature selection techniques in improving reservoir characterisation predictions. J Petrol Sci Eng. 2022;208:109244. DOI:10.1016/j.petrol.2021.109244.

[18] Zhu L, Zhang W, Tu G. Generalization enhancement of artificial neural network for turbulence closure by feature selection. Adv Aerodyn. 2022;4(1):1–24. DOI:10.1186/s42774-021-00088-5.

[19] Xu J, Yuan M, Ma Y. Feature selection using self-information and entropy-based uncertainty measure for fuzzy neighborhood rough set. Comp Intellig Sys. 2022;8(1):287–305. DOI:10.1007/s40747-021-00356-3.

[20] Zhou Y, Cheng G, Jiang S, et al. Building an efficient intrusion detection system based on feature selection and ensemble classifier. Comput Netw. 2020;174:107247. DOI:10.1016/j.comnet.2020.107247.

[21] Vadim Arzamasov. Electrical Grid Stability Simulated Data Set. 2021 [cited 2021 April]. Available from: https://archive.ics.uci.edu/ ml/datasets/ Electrical Grid Stability Simulated Data.

[22] Ramasamy K, Balakrishnan K, Velusamy D. Detection of cardiac arrhythmias from ECG signals using FBSE and JAYA optimized ensemble random subspace K-nearest neighbor algorithm. Biomed Sign Proc Cont. 2022;76:103654. DOI:10.1016/j.bspc.2022.103654.

[23] Velusamy D, Ramasamy K. Ensemble of heterogeneous classifiers for diagnosis and prediction of coronary artery disease with reduced feature subset. Comp Meth Prog Biomed. 2021;198:105770. DOI:10.1016/j.cmpb.2020.105770.

[24] Kotenko I, Izrailov K, Buinevich M. Static analysis of information systems for IoT cyber security: a survey of machine learning approaches. Sensors (Basel). 2022;22(4):1335. DOI:10.3390/s22041335.

[25] Acarali D, Rajesh Rao K, Rajarajan M, et al. Modelling smart grid IT-OT dependencies for DDoS impact propagation. Comput Secur. 2022;112(102528):102528. DOI:10.1016/j.cose.2021.102528.

[26] Yuan Y, Cheng D, Zhou Z, et al. Minimum conflict consensus with budget constraint based on social network analysis. Comput Ind Eng. 2022;168(108098):108098. DOI:10.1016/j.cie.2022.108098.

[27] Dempster A, Petitjean F, Webb GI. ROCKET: exceptionally fast and accurate time series classification using random convolutional kernels. Data Min Knowl Discov. 2020;34(5):1454–1495. DOI:10.1007/s10618-020-00701-z.

[28] Kirichenko L, Radivilova T, Bulakh V. Classification of fractal time series using recurrence plots. Sci Technol. 2018: 719–724. DOI:10.1109/INFOCOMMST.2018.8632010.

[29] Han L, Yu C, Xiao K, et al. A new method of mixed gas identification based on a convolutional neural network for time series classification. Sensors. 2019;19(9):1960. DOI:10.3390/s19091960.

[30] Liu CL, Hsaio WH, Tu YC. Time series classification with multivariate convolutional neural network. IEEE Trans Ind Electron. 2019;66(6):4788–4797. DOI:10.1109/TIE.2018.2864702.

[31] Lee YJ, Pao HK, Shih SH, et al. Compressed learning for time series classification. 2016 IEEE International Conference on Big Data (Big Data); 2016. DOI:10.1109/BigData.2016.7840688.

[32] Putra MAP, Kim DS, Lee JM. Composite multi-directional LSTM for accurate prediction of energy consumption. 2022 International Conference on Information Networking (ICOIN); 2022. DOI:10.1109/ICOIN53446.2022.9687290

[33] Mishra S, Ganthia BP, Sridharan A, et al. Optimization of load forecasting in smartgrid using artificial neural network based NFTOOL and NNTOOL. J Phys Conf Ser. 2022;2161(1):0012068. DOI:10.1088/1742-6596/2161/1/012068.

[34] Samanta SK, Chanda CK. Analysis of voltage stability in smart grid system due to demand load variation. In Lecture notes in electrical engineering, Singapore: Springer Singapore; 2022. p. 123–136.

[35] Shreenidhi HS, Ramaiah NS. A two-stage deep convolutional model for demand response energy management system in IoT-enabled smart grid. Sust Energ, Grid Network. 2022;30(100630):100630. DOI:10.1016/j.segan.2022.100630.

[36] Yi Y, Chen Z, Wang L. Intelligent aging diagnosis of conductor in smart grid using label-distribution deep convolutional neural networks. IEEE Trans Instrum Meas. 2022;71:1–8. DOI:10.1109/TIM.2022.3141160.

[37] Nagarajan C, Tharani B, Saravanan S, et al. Performance estimation and control analysis of AC-DC/DC-DC hybrid multi-port intelligent controllers based power flow optimizing using STEM strategy and RPFC technique. Int J Robot Cont Syst. 2021;2(1):124–139. DOI:10.31763/ijrcs.v2i1.562.

[38] Ghaghishpour A, Koochaki A. An intelligent method for online voltage stability margin assessment using optimized ANFIS and associated rules technique. ISA Trans. 2020;102:91–104. DOI:10.1016/j.isatra.2020.02.028.

[39] Villa-Acevedo WM, López-Lezama JM, Colomé DG, et al. Long-term voltage stability monitoring of power system areas using a kernel extreme learning machine approach. Alexandria Eng J. 2022;61(2):1353–1367. DOI:10.1016/j.aej.2021.06.013.

[40] Zhu L, Lu C, Dong ZY, et al. Imbalance learning machine-based power system short-term voltage stability assessment. IEEE Trans Ind Inf. 2017;13(5):2533–2543. DOI:10.1109/TII.2017.2696534.

[41] Malbasa V, Zheng C, Chen PC, et al. Voltage stability prediction using active machine learning. IEEE Trans Smart Grid. 2017;8(6):3117–3124.

[42] Alazab M, Khan S, Krishnan SSR, et al. A multi-directional LSTM model for predicting the stability of a smart grid. IEEE Access. 2020; 8:85454–85463. DOI:10.1109/ACCESS.2020.2991067.

[43] Hou S, Liu Y, Yang Q. Real-time prediction of rock mass classification based on TBM operation big data and stacking technique of ensemble learning. J Rock Mech Geotech Eng. 2022;14(1):123–143. DOI:10.1016/j.jrmge.2021.05.004.

[44] Lin S, Zheng H, Han B, et al. Comparative performance of eight ensemble learning approaches for the development of models of slope stability prediction. Acta Geotech. 2022;17(4):1477–1502. DOI:10.1007/s11440-021-01440-1.

[45] Guo X, Gao Y, Zheng D, et al. Study on short-term photovoltaic power prediction model based on the stacking ensemble learning. Energy Reports. 2020;6:1424–1431.

[46] Khoei TT, Aissou G, Hu WC, et al. Ensemble learning methods for anomaly intrusion detection system in smart grid. 2021 IEEE International Conference

on Electro Information Technology (EIT); 2021: p. 129–135. DOI:10.1109/EIT51626.2021.9491891.

[47] Khoei TT, Gasimova A, Ahajjam MA, et al. A comparative analysis of supervised and unsupervised models for detecting gps spoofing attack on UAVs. 2022 IEEE International Conference on Electro Information Technology (eIT); 2022: p. 279–284, DOI:10.1109/eIT53891.2022.9813826.

[48] Azad S, Sabrina F, Wasimi S. Transformation of smart grid using machine learning. 2019 29th Australasian Universities Power Engineering Conference (AUPEC); 2019. DOI:10.1109/AUPEC48547.2019.211809.

[49] Bomfim TS. Evolution of machine learning in smart grids. 2020 IEEE 8th International Conference on Smart Energy Grid Engineering (SEGE); 2020. doi:10.1109/SEGE49949.2020.9182023.

[50] Moldovan D, Salomie I. Detection of sources of instability in smart grids using machine learning techniques. 2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP); 2019. DOI:10.1109/ICCP48234.2019.8959649.

[51] Li C. RETRACTED: stability analysis of distributed smart grid based on machine learning. IOP Conf Ser: Earth and Env Sci. 2021;692(2):0022125. DOI:10.1088/1755-1315/692/2/022125.

[52] Ghosh A, Kole A. A comparative analysis of enhanced machine learning algorithms for smart grid stability prediction. TechRxiv. 2021. DOI:10.36227/techrxiv.16863145.v1

[53] Abu Al-Haija Q, Smadi AA, Allehyani MF. Meticulously intelligent identification system for smart grid network stability to optimize risk management. Energies. 2021;14(21):6935. DOI:10.3390/en14216935.

[54] Breviglieri P, Erdem T, Eken S. Predicting smart grid stability with optimized deep models. SN Comp Sci. 2021;2(2). DOI:10.1007/s42979-021-00463-5

[55] Chahal A, Gulia P, Gill NS, et al. Performance analysis of an optimized ANN model to predict the stability of smart grid. Complexity. 2022;2022:1–13. DOI:10.1155/2022/7319010.

[56] Aziz F, Lawi A. Increasing electrical grid stability classification performance using ensemble bagging of C4.5 and classification and regression trees. Int J Elect Comp Eng (IJECE). 2022;12(3):2955. DOI:10.11591/ijece.v12i3.pp2955-2962.