Received: 28 May 2021 | Revised: 19 August 2021 | Accepted: 10 September 2021

DOI: 10.1111/exsy.12832

ORIGINAL ARTICLE

Expert Systems WILEY

Machine learning-based model for prediction of power consumption in smart grid- smart way towards smart city

Shamik Tiwari[1] | Anurag Jain[1] | Nada Mohamed Osman Sid Ahmed[2] | Charu[3] | Lulwah M. Alkwai[2] | Alaa Kamal Yousif Dafhalla[2] | Sawsan Ali Saad Hamad[2]

[1]School of Computer Science, University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India

[2]College of Computer Science and Engineering, University of Ha'il, Ha'il, Kingdom of Saudi Arabia

[3]Department of CSE & IT, Jaypee Institute of Information Technology, Noida, India

Correspondence

Anurag Jain, School of Computer Science, University of Petroleum and Energy Studies, Bidholi, Dehradun-248007, Uttarakhand, India.
Email: anurag.jain@ddn.upes.ac.in

Funding information

Scientific Research Deanship at the University of Ha'il, Grant/Award Number: RG-20093

Abstract

A smart city is an idea that is realized by the computing of a large amount of data collected through sensors, cameras, and other electronic methods to provide services, manage resources and solve daily life problems. The transformation of the conventional grid to a smart grid is one step in the direction towards smart city realization. An electric grid is composed of control stations, generation centres, transformers, communication lines, and distributors, which helps in transferring power from the power station to domestic and commercial consumers. Present electric grids are not smart enough that they can estimate the varying power requirement of the consumer. Also, these conventional grids are not enough robust and scalable. This has become the motivation for shifting from a conventional grid to a smart grid. The smart grid is a kind of power grid, which is robust and adapts itself to the varying needs of the consumer and self-healing in nature. In this way, the transformation from a conventional grid to a smart grid will help the government to make a smart city. The emergence of machine learning has helped in the prediction of the stability of the grid under the dynamically changing requirement of the consumer. Also, the usage of a variety of sensors will help in the collection of real-time consumption data. Through machine learning algorithms, we can gain an insight view of the collected data. This has helped the smart grid to convert into a robust smart grid, as this will help in avoiding the situation of failure. In this work, the authors have applied logistic regression, decision tree, support vector machine, linear discriminant analysis, quadratic discriminant analysis, naïve Bayes, random forest, and k-nearest neighbour algorithms to predict the stability of the grid. The authors have used the smart grid stability dataset freely available on Kaggle to train and test the models. It has been found that a model designed using the support vector machine algorithm has given the most accurate result.

KEYWORDS

classifier algorithm, machine learning, prediction model, smart city, smart grid

1 | INTRODUCTION

The use of IT, AI, and IOT to collect, assess, and merge crucial details from basic systems in existing cities to raise facilities and make them more robust is termed a smart city. A smart city is an idea that is realized by the computing of a large amount of data collected through sensors, cameras, and other electronic methods to provide services, manage resources and solve daily life problems (Je & Huh, 2021). Through smart cities, needs like city services, public safety, daily livelihood, or commercial and industrial activities can be fulfilled in a timely, efficient, and intelligent manner. Power or energy is the backbone of all the systems (Wang et al., 2020). The smart grid is a kind of power grid, which is robust and adapts itself to the varying needs of the consumer and self-healing in nature. This makes it one of the strong pillars for the realization of the smart city concept (Ahmad et al., 2020).

These days, conventional electricity grids are transforming into smart grids. Conventional grids have unidirectional communication and power is lost in the generation, transmission, and distribution. Smart grids are introduced to resolve concerns of traditional grids (Fang et al., 2011). In smart grids, communication is bi-directional without restrictions. Smart grid systems are designed with bi-directional communication tools, control systems, and information systems. These advanced tools comprise cutting-edge phasor networks. These phasor networks contain phasor measurement units (PMUs), Phasor Data Concentrators (PDC), and Supervisory Control and Data Acquisition (SCADA) systems. These advanced sensors permit operators to measure grid stability. Smart grids also have smart digital meters that provide customers improved information and spontaneous feedbacks, sense the faults, automated feeder switches for re-routing of power in case of grid failure, and batteries with surplus energy to meet consumer demand in the future. This exciting transformation of the electric grid produces both opportunities and challenges to advance the capabilities of the present electricity distribution system. Forecasting of electricity plays an important role in the smart grid to minimize operational costs and effective management. Load and price prediction provide future trends (Arif et al., 2020). The latest developments in information technologies and the accessibility of bulky and varied data in power grids are opening the ways for the deployment of intelligent algorithms. Compared to conventional computational methods, machine learning methods have gained a benefit from their fundamental generalization competency. One of the key challenges of using smart grids is the capacity to manage electric and communication networks, where instantaneous information such as energy, demand, power quality, cost, and so forth, is related to the nodes of the system. Therefore, the attention lies in designing intelligent systems using machine learning algorithms that can make decisions, even in uncertain conditions (Miraftabzadeh et al., 2019). The grid stability is determined by the balance between power generation and demand. The flow of power has become bidirectional with the increase in decentralized production of power. These modifications need a smart control centre for upholding the balance between production and energy demand. Machine learning methods are followed to categorize the stability circumstances based on the response of the heterogeneous users to examine the stronger association among the parameters of the grid stability and input space (Moldovan & Salomie, 2019).

The prediction of smart grid stability is an inspiring research area since this information can be very valuable for the recognition of the components that causes volatilities in the smart grid and subsequently it is very valuable to decide configurations in which the grid is stable even if some members show irregularities (Azad et al., 2019). Collection and analysis of a large volume of data is the major task in the implementation of a smart city. Smart grids can use machine learning to make critical decisions about load at peak hours, change in consumer demand, and the reliability of the grid at peak load. This has helped in the prediction of the stability of the grid under the dynamically changing requirement of the consumer. Through machine learning algorithms, we can gain an insight view of the collected data. This has helped the smart grid to convert into a robust smart grid, as this will help in avoiding the situation of failure (Hossain et al., 2019).

In this manuscript, the authors have used different machine learning (logistic regression, decision tree, support vector machine, linear discriminant analysis, quadratic discriminant analysis, naïve Bayes, random forest, and k-nearest neighbour) algorithms with smart grid stability dataset to predict the stability of the grid. Support Vector Machine algorithm-based model has given the most accurate result. The rest of the article is structured into four sections. Section 2 provides the literature review in the field of machine learning applications for the smart grid. Section 3 and Section 4 present the methodology, result, and their analysis respectively. Finally, the conclusion is discussed in Section 5.

2 | LITERATURE REVIEW

Several studies have examined the application of the appropriate machine learning algorithm to handle multiple challenges in the area of power grid operations, controls, and supervision. To implement the smart city concept, the current utilization of smart grids is offering fresh research directions for the real-time utilization of intelligent systems in smart grids, especially in grid stability prediction. The contributions of other researchers in the same domain are discussed below.

Samanta and Chanda have investigated the influence of smart grid stability on synchronous generators. Authors have also studied instabilities through voltage, rotor angle, and frequency steadiness. A new voltage stability index is also proposed to decide the hazardous lines of a smart grid (Samanta & Chanda, 2017). In a work by Panda and Das (2019), a feature ranking and regression study are applied to analyse the association of the parameters of the decentralized control system such as reaction time and grid stability condition. The experiment is performed based on the simulated data, which includes different features of heterogeneous consumers. A study on the impact of local wind farm power variation and the upgrading approaches on grid voltage stability is studied by Qinyu et al., (2018) using a simulated model to assess the upgrading impacts of grid voltage stability with diverse reactive power compensation systems. A software termed a Grid Stability Awareness System to observe and examine the real-time smart grid stability is proposed in Ma et al., (2017). This tool has five modules namely an oscillation monitoring module, an event detection module, a transient instability monitoring module, a voltage stability monitoring module, and an angle difference monitoring module. These modules can address the voltage stability, small-signal stability, and transient stability in power grids. One of the major challenges for the machine learning algorithms-based grid stability prediction is to promise the learning speed and prompt response for the optimal factors of the applied algorithm when handling the bulky and high dimensional data. A grid stability evaluation technique using the Bayesian optimized Light GBM optimization method is suggested to handle this issue in Wang et al., (2019). The Multidirectional Long Short-Term Memory (MLSTM) method is projected by Alazab et al. (2020) to forecast the stability of the smart grid networks. The results of MLSTM are compared with traditional machine learning models like Long Short-Term Memory, Gated Recurrent Units, and Recurrent Neural Networks. The comparative investigation demonstrates the superiority of the projected model on other models. Recently, a fuzzy inference-based prediction system described by an inherently enhanced interpretability-accuracy trade-off for apparent and perfect forecast the decentralized smart grid control stability is proposed (Gorzałczany et al., 2020). Moreover, the authors also suggested the hierarchy of input attributes from the grid stability perspective.

Quitzow and Rohde (2021) have discussed the case study of Berlin for realizing a smart city through the smart grid. The authors have discussed their energy-related policies for implementing a smart city from the smart grid. Further, they have discussed dominant imaginaries for the smart city through the smart grid. The economic imperative, environmental solution, and experimental challenges are the important dominant imaginaries that show smart grid technology. Tanwar et al. (2018) have discussed the role of IoT and smart grid in the implementation of the smart city. The authors have highlighted the significance of energy in all sectors. They have also discussed the different issues in the development of the smart city. Heterogeneity, unplanned development of the city, adaptability by people are some major issues in the development of the smart city. Koutitas (2018) have mentioned the smart grid as the backbone and fundamental element of the smart city. The author has mentioned the smart grid as integration of conventional power grid with information and communication technology. Further, the smart grid as an anchor of the smart city is also explained in detail. Bonetto and Rossi (2017) have highlighted the dependency of a smart city on the efficiency and reliability of the smart grid. Cooperation of power electronics, information and communication technology, and economics are necessary for the implementation of the smart grid. Challenges in the implementation of the smart grid for a smart city are discussed in detail.

Considering the nature of the smart grid stability problem and its contribution to the realization of smart city two key objectives are proposed in this work and described in Section 3:

- Pursue improvements in smart grid stability predictions with prominent machine learning algorithms.
- Assess the performance of these machine learning models.

3 | METHODOLOGY

3.1 | Dataset details

The authors have used the smart grid stability dataset (Arzamasov et al., 2018) created by Vadim Arzamasov. Vadim has donated this dataset to the University of California (UCI) Machine Learning Repository. It is freely available for research work purposes on the site of Kaggle. The dataset contains 12 features and 60,000 rows of data. Based on those 12 features, the stability value for the grid is calculated and the grid is classified into stable and unstable. Description of features are as follows:

- The first four features, tau1, tau2, tau3, and tau4 indicate the reaction time of each participant in the network. This value lies in the range of 0.5 to 10. Feature tau1 is corresponding to the supplier node while tau2, tau3, tau4 are corresponding to the consumer node.
- The next four features p1, p2, p3, and p4 are indicating the power produced and consumed by each participant in the network. A positive value is indicating power has been produced where a negative value is indicating that power has been consumed by the participant. Feature p1 is representing the supplier node while p2, p3, p4 are representing the consumers. The real value in the range -2.0 to -0.5 has been used for p2, p3, and p4 while p1 will be the negative sum of p2, p3, and p4.
- The next four features g1, g2, g3, and g4 are indicating the coefficient of price elasticity for each participant in the network. This value lies in the range of 0.05 to 1.00. Feature g1 is corresponding to the supplier node while g2, g3, g4 are corresponding to the consumer node. Here g stands for gamma.
- The first 12 features are predictive while the last two features are dependent. The next two features are stab and stabf respectively. The value of stab is calculated and if it comes out to be positive then stabf will have unstable labels whereas on having a negative value of stab, the label of stabf will be stable.

3.2 | Dataset pre-processing

The following steps have been executed under the data preprocessing task:

- The content of the dataset was generated from simulation work, so there is no missing value in this dataset. So no need to apply any missing value generation technique.

- All features of the dataset are used in the calculation of stability value, so no feature selection algorithm is used.
- All features have numerical values so feature coding is done.
- Data has no outliers as shown in box plots in Figures 1–3 respectively. These boxplots also reveal information about a statistical data set's shape, variability, and median.

3.3 | Modelling

The authors have divided the available dataset into two parts having 70% and 30% data respectively. 70% data is used for learning the model while the remaining 30% data is utilized for testing the designed model.

For the training of models, authors have used eight different classifier algorithms namely linear discriminant analysis, quadratic discriminant analysis, logistic regression, decision tree, support vector machine, naïve Bayes, random forest, and k-nearest neighbour.

3.4 | Classifier details

- Logistic regression: The method of modelling the probability of a discrete result given an input variable is known as logistic regression. A binary result is the most prevalent outcome model in logistic regression. For multiple discrete outcomes, multimodal logistic regression is used to model the situation. Logistic regression is a useful analysis tool for categorization problems. This classification algorithm is used to find the probability of a particular feature based on another feature (Lamba et al., 2021). Based on the probability dependent feature is assigned a value of 0 or 1. It uses the Sigmoid function to restrict the output in the range of [0,1]. The hypothesis of logistic regression is given in Equation 1:

$$h_\theta(x) = \frac{1}{1+e^{\left(-\theta^T x\right)}} \tag{1}$$

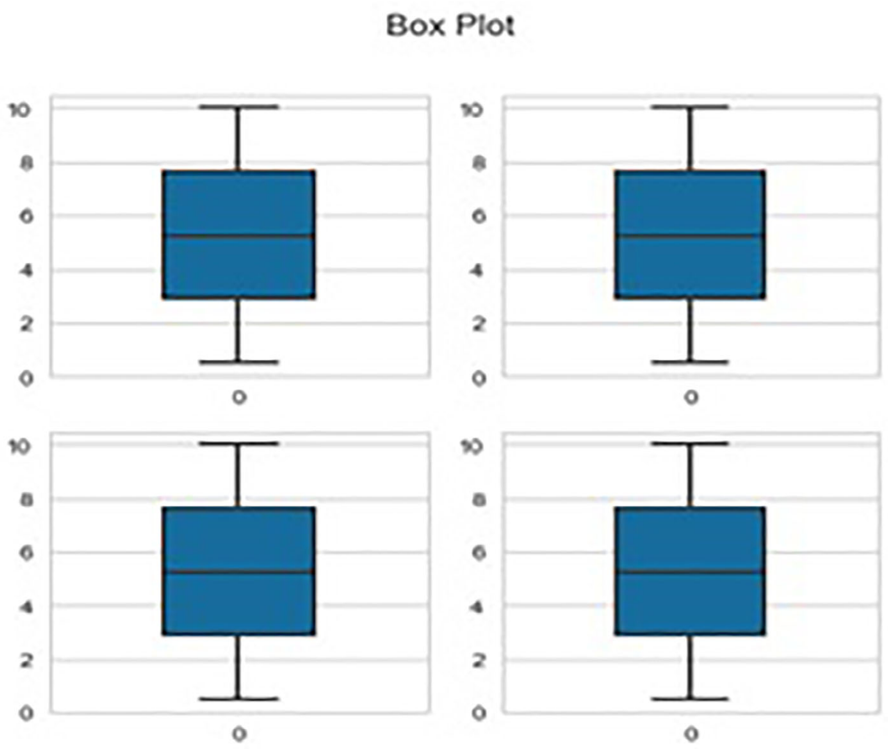

FIGURE 1 Box plots show no outlier for tau1, tau2, tau3, and tau4 features respectively

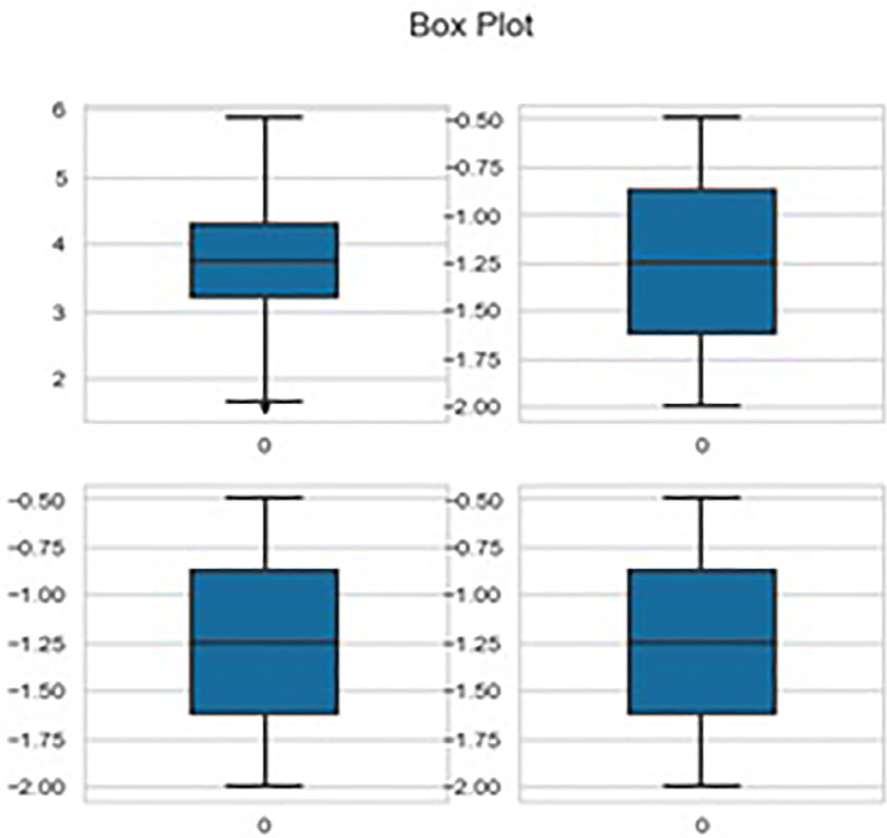

FIGURE 2 Box plots show no outlier for p1, p2, p3, and p4 features respectively

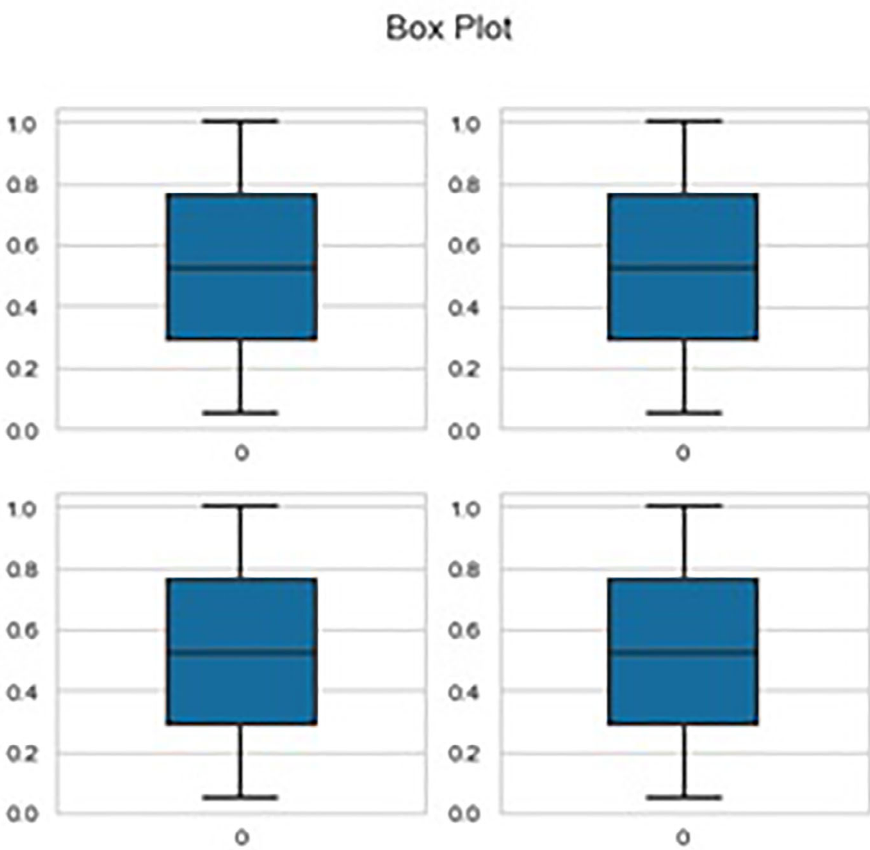

FIGURE 3 Box plots show no outlier for g1, g2, g3, and g4 features respectively

where $\theta^T = [\theta_0\, \theta_1\, \theta_2\, \theta_3 \ldots \theta_j]$ and feature vector $x = [x_0\, x_1\, x_2\, x_3 \ldots x_j]$

- Decision tree: Decision tree is a supervised learning-based tree-structured classifier approach, which consists of decision node (specify a test or choice of some feature, with one branch for each outcome) and leaf nodes (specify the predicted value or classification of the problem) (Priyanka, 2020). In a decision tree classifier, constructing an optimal decision tree is a critical task. A given collection of attributes can be used to create a variety of decision trees. While some of the trees are more accurate than others, due to the exponential size of the search field, identifying the best tree is computationally impossible. However, numerous efficient algorithms have been developed to quickly construct a reasonably accurate, albeit imperfect, decision tree. A greedy algorithm-based approach is used for the creation of a decision tree by taking a group of locally optimal decisions about attribute selection for partitioning data. Take, for instance, Classification and Regression Trees (CART) algorithm.
- Random forest: Random forest is a supervised learning-based meta-heuristic classifier, which consists of multiple decision trees that are independent but they operate in an ensemble way (Sharma et al., 2020). Each decision tree gives its class as output and class, which is in majority among the output of all decision trees, will become the class of random forest. When to stop and which attribute to split are some of the factors, which can affect the performance of decision tree classification.
- Linear discriminant analysis is a dimensionality reduction method that is generally used for supervised classification applications (Devi & Radhika, 2018). It is used for modelling differences in classes that is, splitting two or more classes. It is used to project the features in higher dimension space into a reduced dimension space.
- Quadratic discriminant analysis: It is a variation of linear discriminant analysis that permits for the non-linear partition of data (Ghojogh & Crowley, 2019). It uses a quadratic decision surface to isolate features of two or more classes of the classification problem.
- Naïve Bayes: It is a form of the probabilistic classifier, which applies the Bayes theorem. It is generally used in various kinds of classification problems like disease diagnosis, spam, document filtering, and so forth (Rani et al., 2021). Independence among the participating features involved in the process is the basic assumption for this algorithm. Its calculations are based on the formula of the Bayes rule given in Equation (2).

$$P\left(\frac{y}{X}\right) = \frac{P(X/y)P(y)}{P(X)} \tag{2}$$

Here y is the class variable and X is the feature vector $= [x_1\, x_2\, x_3\, x_4 \ldots x_n]$

- Support vector machine: Classification and regression both are used in this, but most are used for classification (Zidi et al., 2017). SVM works by finding an optimal line that performs the partitioning of datasets in classes. Margin or gap is the distance between data points of the class. Hyperplane should be chosen, and such that margin should be max. Data points of both classes, which are closed to the hyperplane, are known as vectors. It is a classifier that is the furthest away from the training data. The perpendicular distance between the hyperplane and the training observations is calculated. The shortest such distance is known as margin, and it is measured between the hyperplane and the observation. As a result, the maximal margin hyperplane is the hyperplane with the highest margin, that is, the hyperplane with the greatest distance between it and the training observations. We can classify testing data using that hyperplane.
- K-nearest neighbour: K-nearest neighbour classifier is instance-based and classifies similar instances. In KNN each point is in multidimensional space, and nearest neighbours are organized; the value of ‘K' may vary for every neighbour (Prasath et al., 2017). The efficiency of the KNN algorithm varies with the number of neighbours to be used. By default, it takes only one neighbour. It is simple to understand and gives amazing results. It is calculations are based on the Euclidian distance formula given in Equation (3).

$$d(p,q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + (q_3 - p_3)^2 \ldots (q_n - p_n)^2} \tag{3}$$

4 | RESULTS AND ANALYSIS

Testing of proposed models is carried out on a machine having window 10 platforms, with the hardware configuration of 8 GB RAM, i3 @GHz processor with the environment developed using Python 3.7.

Accuracy, sensitivity, precision, F-score, and AUC (Tiwari, 2020a, 2020b) are the most important parameters computed to examine the performance of any classifier model. The area under the curve (AUC) summarizes the Receiver Operating Characteristic (ROC) curve that calculates the classifier's ability to distinguish between classes. The AUC shows how well the model differentiates among negative and positive classes. A higher value of AUC is desirable. Accuracy represents the correctness of the classifier in predicting the correct class. These metrics are defined as:

$$\text{Precision} = \text{Tp}/(\text{Tp}+\text{Fp})$$

$$\text{Sensitivity} = \text{Tp}/(\text{Tp}+\text{Fn})$$

$$\text{F}-\text{Score} = (2*\text{Precision}*\text{Sensitivity})/(\text{Precision}+\text{Sensitivity})$$

$$\text{Accuracy} = (\text{Tp}+\text{Tn})/(\text{Tp}+\text{Tn}+\text{Fn}+\text{Fp})$$

where Tp, Tn, *Fp*, and *Fn* are the truly projected positive, truly negative cases, false-positive cases, and false-negative cases respectively. Results in terms of these metrics are presented in Table 1.

Figure 4 illustrates the comparison of different classifiers on the scale of accuracy. The support vector machine classifier has given the best accuracy.

TABLE 1 Results for each machine learning model in terms of performance metrics

Model/metrics	Accuracy	Precision	Recall	F-score	AUC (ROC)
Logistic regression	0.81	0.8	0.78	0.81	0.89
Decision tree	0.9	0.89	0.89	0.9	0.89
Support vector machine	0.98	0.97	0.97	0.97	0.99
Linear discriminant analysis	0.81	0.8	0.78	0.81	0.89
Quadratic discriminant analysis	0.86	0.85	0.85	0.86	0.94
Random forest	0.95	0.95	0.94	0.95	0.99
K-nearest neighbours	0.89	0.89	0.87	0.89	0.95
Naïve Bayes	0.83	0.83	0.79	0.82	0.91

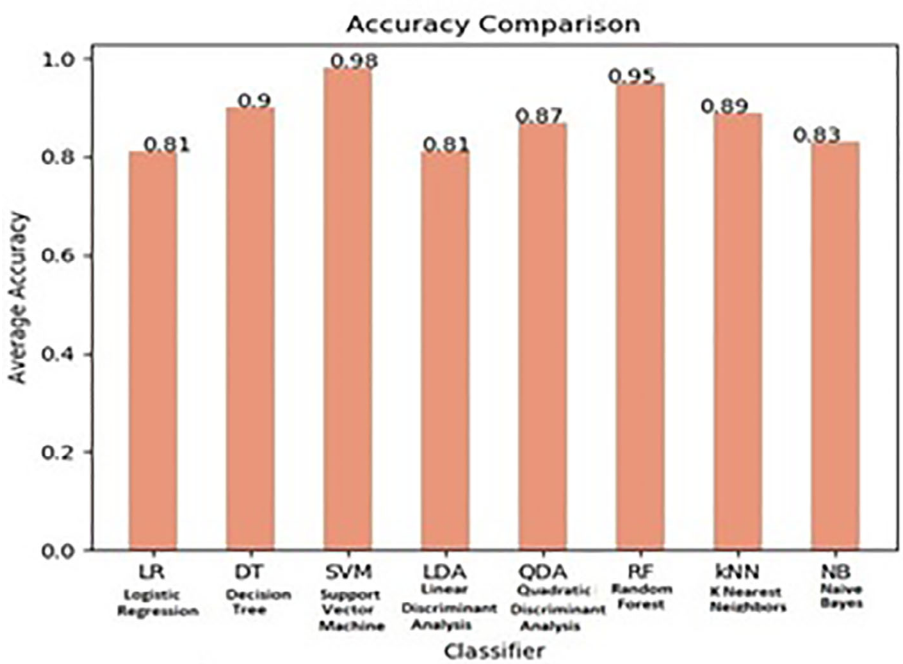

FIGURE 4 Classifier comparison on the scale of accuracy

Precision is a measure of consistency. It represents the ability of the classifier to return only relevant cases. Figure 5 displays the comparison of different classifiers on the scale of precision. The support vector machine classifier has given the best accuracy.

Sensitivity or recall represents the ability of the classifier to recognize all relevant cases correctly. Figure 6 shows the comparison of different classifiers on the scale of sensitivity. The support vector machine classifier has given the best accuracy.

There is always a trade-off between precision and recall. Which parameter between precision and recall should be maximized, depends upon the problem. However, there is another metric through which we can consider both precision and recall at the same time, it is called F-score. Now

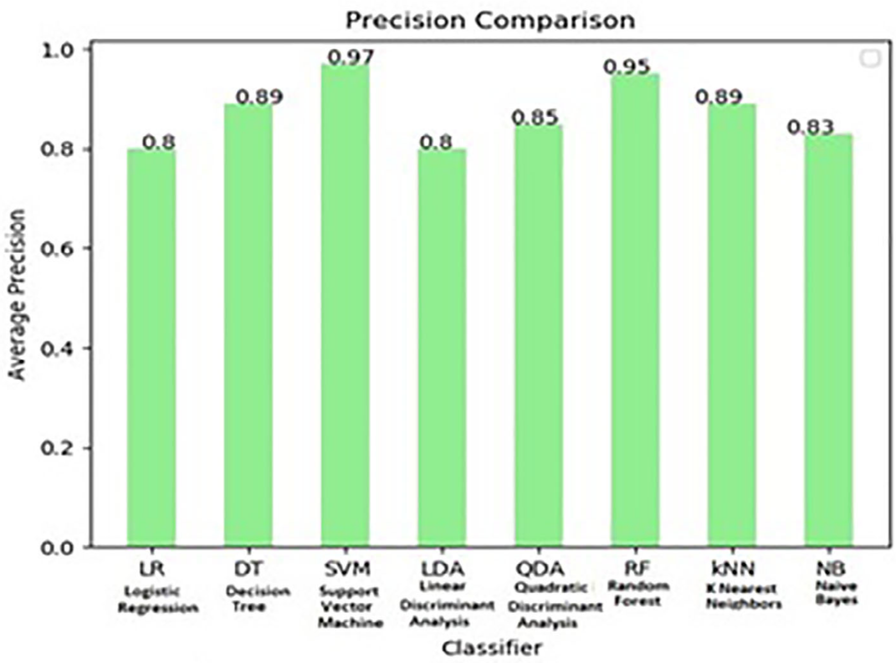

FIGURE 5 Classifier comparison on the scale of precision

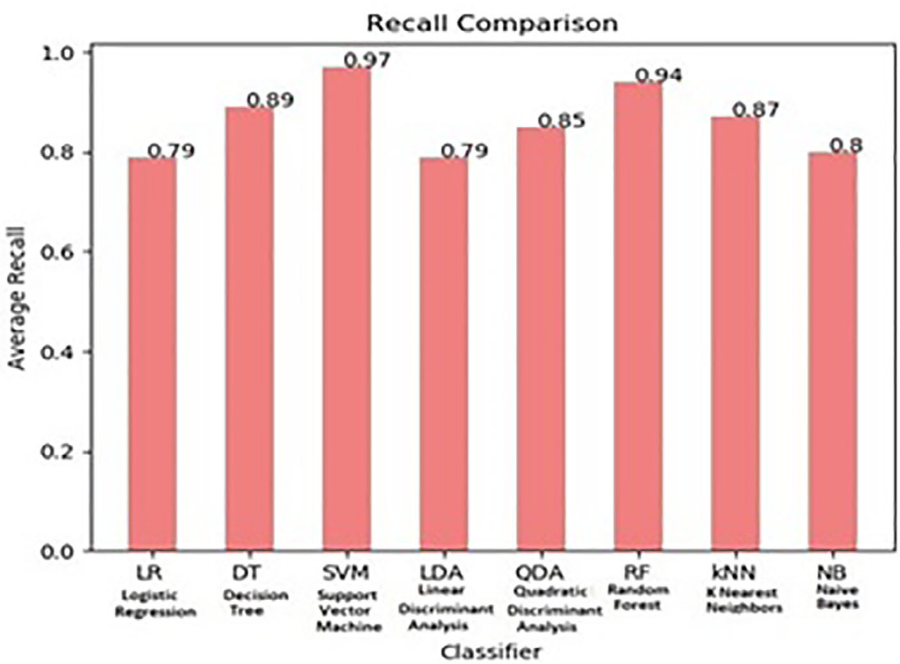

FIGURE 6 Classifier comparison on the scale of recall

instead of making a balance between two metrics, the aim is to maximize a single parameter F-score. F-score is a harmonic mean of precision and sensitivity (recall). Figure 7 displays the comparison of different classifiers on the scale of the F-score. The support vector machine classifier has given the best accuracy.

To further strengthen and validate the results, the ROC Curve is designed and the AUC is calculated for all eight models. The high area under the ROC Curve (AUC) score shown in Figure 8 further strengthens the result shown in previous figures.

From the results shown in Figures 4–8, it can be concluded that the model designed using a support vector machine classifier is the most accurate in the prediction of grid stability.

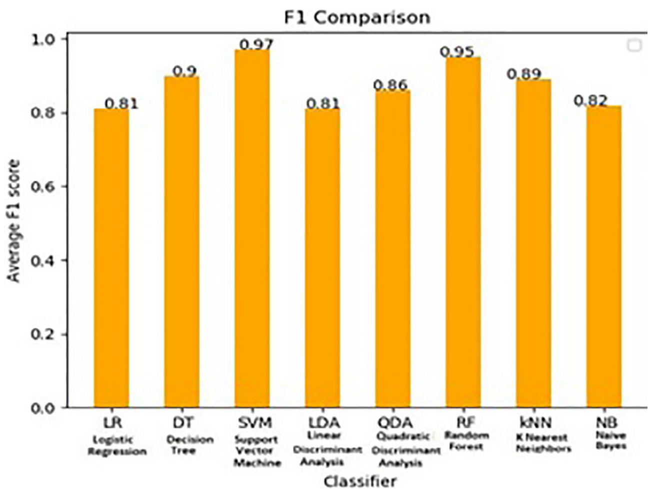

FIGURE 7 Classifier comparison on the scale of F1 score

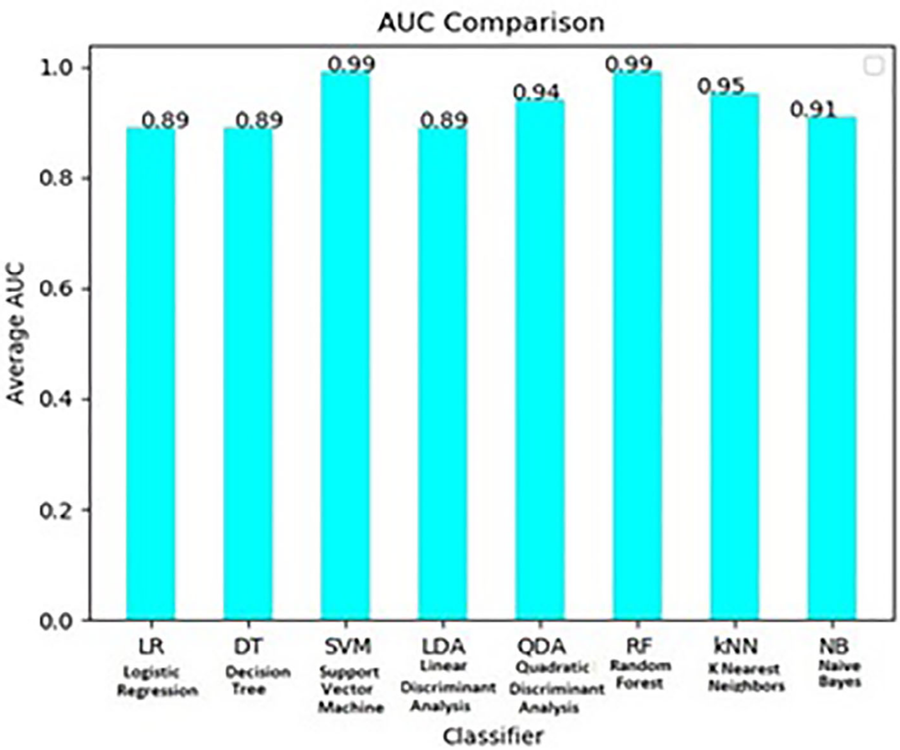

FIGURE 8 Classifier comparison on the scale of the AUC

5 | CONCLUSION

The power requirement of the consumer is always dynamic. The stability of the grid depends on how efficiently the system is handling the fluctuating power requirement of the customer. With the emergence of machine learning, we found a technique that can be used for the prediction of power consumption and stability of the grid. The authors have used the smart grid stability dataset available on the site of Kaggle for the training and testing of different machine learning models. From the simulation result, it has been found that a model designed using a support vector machine (SVM) classifier has shown the highest accuracy on the scale of precision, recall, F-score, and accuracy. It has been further validated by the result of the AUC score. Results can be verified through the code available at https://github.com/shamiktiwari/smartgrid.

In the future, the model can be improved by considering the surrounding and geographical conditions. Also by using deep learning-based classifiers, the accuracy level can be improved.

ACKNOWLEDGEMENT

This project has been funded by Scientific Research Deanship at the University of Ha'il – Saudi Arabia through project number RG-20093.

CONFLICT OF INTEREST

There is no conflict of interest regarding the publication of this article.

DATA AVAILABILITY STATEMENT

Smart grid stability dataset (Arzamasov et al., 2018) that supports the finding of this study is openly available at Kaggle. It was created by Vadim Arzamasov. Vadim has donated this dataset to the University of California (UCI) Machine Learning Repository. It is freely available for research work purposes on https://www.kaggle.com/pcbreviglieri/smart-grid-stability.

ORCID

Shamik Tiwari https://orcid.org/0000-0002-5987-7101

Anurag Jain https://orcid.org/0000-0001-5155-022X

REFERENCES

Ahmad, T., Zhang, H., & Yan, B. (2020). A review on renewable energy and electricity requirement forecasting models for smart grid and buildings. *Sustainable Cities and Society*, *55*, 102052.

Alazab, M., Khan, S., Krishnan, S. S. R., Pham, Q. V., Reddy, M. P. K., & Gadekallu, T. R. (2020). A multidirectional LSTM model for predicting the stability of a smart grid. *IEEE Access*, *8*, 85454–85463.

Arif, A., Javaid, N., Anwar, M., Naeem, A., Gul, H., & Fareed, S. (2020). Electricity load and price forecasting using machine learning algorithms in smart grid: A survey. In L. Barolli, F. Amato, F. Moscato, T. Enokido, & M. Takizawa (Eds.), *Workshops of the International Conference on Advanced Information Networking and Applications* (pp. 471–483). Springer.

Arzamasov, V., Böhm, K. & Jochem, P. Towards concise models of grid stability in 2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)

Azad, S., Sabrina, F., & Wasimi, S. (2019). Transformation of smart grid using machine learning. In *2019 29th Australasian Universities Power Engineering Conference (AUPEC)* (pp. 1–6).

Bonetto, R., & Rossi, M. (2017). Smart grid for the smart city. In V. Angelakis, E. Tragos, H. Pöhls, A. Kapovits, & A. Bassi (Eds.), *Designing, Developing, and Facilitating Smart Cities* (pp. 241–263). Springer.

Devi, S. S., & Radhika, Y. (2018). A survey on machine learning and statistical techniques in bankruptcy prediction. *International Journal of Machine Learning and Computing*, *8*(2), 133–139.

Fang, X., Misra, S., Xue, G., & Yang, D. (2011). Smart grid—The new and improved power grid: A survey. *IEEE Communications Surveys & Tutorials*, *14*(4), 944–980.

Ghojogh, B. and Crowley, M. (2019). Linear and quadratic discriminant analysis: Tutorial. arXiv preprint arXiv:1906.02590.

Gorzałczany, M. B., Piekoszewski, J., & Rudziński, F. (2020). A modern data-mining approach based on genetically optimized fuzzy systems for interpretable and accurate smart-grid stability prediction. *Energies*, *13*(10), 25–59.

Hossain, E., Khan, I., Un-Noor, F., Sikander, S. S., & Sunny, S. H. (2019). Application of big data and machine learning in smart grid, and associated security concerns: A review. *IEEE Access*, *7*, 13960–13988.

Je, S. M., & Huh, J. H. (2021). Estimation of future power consumption level in smart grid: Application of fuzzy logic and genetic algorithm on big data platform. *International Journal of Communication Systems*, *34*(2), e4056.

Koutitas, G. (2018). The smart grid: Anchor of the smart city. In S. McClellan, J. Jimenez, & G. Koutitas (Eds.), *Smart Cities* (pp. 53–74). Springer.

Lamba, R., Gulati, T., Al-Dhlan, K. A., & Jain, A. (2021). A systematic approach to diagnose Parkinson's disease through kinematic features extracted from handwritten drawings. *Journal of Reliable Intelligent Environments*, *7*, 253–262. https://doi.org/10.1007/s40860-021-00130-9

Ma, J., Feuerborn, S., Black, C., & Venkatasubramanian, V. M. (2017). A comprehensive software suite for power grid stability monitoring based on synchrophasor measurements. In *2017 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)* (pp. 1–5).

Miraftabzadeh, S. M., Foiadelli, F., Longo, M., & Pasetti, M. (2019). A survey of machine learning applications for power system analytics. In *2019 IEEE International Conference on Environment and Electrical Engineering and 2019 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)* (pp. 1–5).

Moldovan, D., & Salomie, I. (2019). Detection of sources of instability in smart grids using machine learning techniques. In *15th International Conference on Intelligent Computer Communication and Processing (ICCP)* (pp. 175–182).

Panda, D. K., & Das, S. (2019). Regression analysis of grid stability under decentralized control. In *2019 International Conference on Engineering, Science, and Industrial Applications (ICESI)* (pp. 1–6).

Prasath, V.B., Alfeilat, H.A.A., Hassanat, A., Lasassmeh, O., Tarawneh, A.S., Alhasanat, M.B. and Salman, H.S.E. (2017). Distance and similarity measures effect on the performance of K-nearest neighbor classifier—A review. arXiv preprint arXiv:1708.04321

Priyanka, D. K. (2020). Decision tree classifier: A detailed survey. *International Journal of Information and Decision Sciences*, *12*(3), 246–269.

Qinyu, B., Lin, Y., Jiayi, M., Tianqi, L., Xiaotian, Z., & Youyin, W. (2018). Analysis of influence with connected wind farm power changing and improvement strategies on grid voltage stability. In *2018 China International Conference on Electricity Distribution (CICED)* (pp. 2029–2033).

Quitzow, L., & Rohde, F. (2021). Imagining the smart city through smart grids? Urban energy futures between technological experimentation and the imagined low-carbon city. *Urban Studies*. https://doi.org/10.1177/00420980211005946

Rani, P., Kumar, R., Jain, A., & Chawla, S. K. (2021). A hybrid approach for feature selection based on genetic algorithm and recursive feature elimination. *International Journal of Information System Modeling and Design*, *12*(2), 17–38.

Samanta, S. K., & Chanda, C. K. (2017). Investigate the impact of smart grid stability analysis on synchronous generator. In *2017 IEEE Calcutta Conference (CALCON)* (pp. 241–247).

Sharma, D., Kumar, R., & Jain, A. (2020). A systematic review of risk factors and risk assessment models for breast cancer. In N. Marriwala, C. C. Tripathi, D. Kumar, & S. Jain (Eds.), *Mobile Radio Communications and 5G Networks, Lecture Notes in Networks and Systems* (Vol. 140, pp. 509–519). Springer.

Tanwar, S., Tyagi, S., & Kumar, S. (2018). The role of internet of things and smart grid for the development of a smart city. In Y. C. Hu, S. Tiwari, K. Mishra, & M. Trivedi (Eds.), *Intelligent communication and computational technologies* (pp. 23–33). Springer.

Tiwari, S. (2020a). A blur classification approach using deep convolution neural network. *International Journal of Information System Modeling and Design*, *11*(1), 93–111.

Tiwari, S. (2020b). A comparative study of deep learning models with handcraft features and non-handcraft features for automatic plant species identification. *International Journal of Agricultural and Environmental Information Systems*, *11*(2), 44–57.

Wang, R., Liu, Y., Ye, X., Tang, Q., Gou, J., Huang, M., & Wen, Y. (2019). Power system transient stability assessment based on bayesian optimized LightGBM. In *2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2)* (pp. 263–268).

Wang, Z., Ogbodo, M., Huang, H., Qiu, C., Hisada, M., & Abdallah, A. B. (2020). Aebis: Ai-enabled blockchain-based electric vehicle integration system for power management in smart grid platform. *IEEE Access*, *8*, 226409–226421.

Zidi, S., Moulahi, T., & Alaya, B. (2017). Fault detection in wireless sensor networks through SVM classifier. *IEEE Sensors Journal*, *18*(1), 340–347.

AUTHOR BIOGRAPHIES

Dr. Shamik Tiwari, currently working as Sr. Associate Professor in the School of Computer Science, University of Petroleum and Energy Studies, Dehradun. He has rich experience of around eighteen years as an academician. His interest areas are digital image processing, computer vision, biometrics, machine learning especially deep learning, and health informatics. He has written many national and international publications, including books in these fields.

Dr. Anurag Jain, currently working as Associate Professor in the School of Computer Science, University of Petroleum and Energy Studies, Dehradun. He is in the field of academia for the last eighteen years. He has published around 44 research papers in renowned journals and conferences. Fourteen students have successfully defended their M. Tech. thesis under his guidance. His research area is scheduling and load balancing in cloud computing, healthcare, machine learning, and data science. Currently, he is guiding 5 Ph.D. students in their research work.

Prof. Nada Mohamed Osman Sid Ahmed is serving as an Assistance professor in the College of Computer Science & Engineering, University of Hai'l. Prior to this, she has served as an Academic Affairs Secretary at the Engineering and Medical Sciences Academy, Sudan. Her fields of Interest are Nanotechnology, Electronics, Renewable Energy, and Telecommunication.

Dr. Charu Gandhi an Associate Professor in the Department of Computer Science and Engineering, Jaypee Institute of Information Technology, Noida, Uttar Pradesh India (Deemed to be University). She holds a doctorate degree in Computer Science from Kurukshetra University, Kurukshetra, India and Master of Technology from Banasthali University, Rajasthan, India She has an academic and research experience of 15 years in the areas of Computer Networks, Mobile Computing, Information Security and Internet of Things. She is a member of CSI, SM-IEEE, EAI, ACM and ISTE.

Dr. Lulwah is currently working as an Assistant Professor in university of Hail, Saudi Arabia. She played various roles of the department chair of Computer Engineer, Coordinator of Research and various committees at college and university levels. She obtained her Masters and PhD in Computer Science from Old Dominion University, USA. She graduated Bachelor of Science in Computer Science from King Saud University, Saudi Arabia. She was a member of the Web Science and Digital Libraries (WS-DL) research group at ODU. She has good knowledge and skills in the field of Data Science, Digital Preservation, Web Science, Digital libraries, recommendation system, Internet Archive, Natural language processing, and several computer programming languages includes Python, Java, Java Script, VB, VB Script, C#, C++, C, PHP, Oracle, SQL. She is publishing and or presenting research papers in various national and international journals, conferences, workshops and seminars. She is attending many training courses.

Dr. Alaa Kamal Yousif Dafhalla earned her B.Sc. degree from the University of Gezira, Sudan in computer engineering faculty of engineering and technology, and pursued her M.Sc study in computer engineering and network at the University of Gezira in 2007. She finished her Ph. D at University Malaysia Perlis on July 2018 in Computer Engineering "Bio-Inspired chameleon technique in MAC protocol for energy-efficient wireless networks". Her research interests include Wireless Network protocols development, Modeling and simulation, Software Engineering, Network Programming, and Optimization methodologies with robust engineering, programming with Open sources that involves Network modeling and simulation.

Dr. Sawsan Ali Saad Hamad, received B.Sc. degree in Electrical Engineering from the University of Khartoum, Sudan, in 2003. Received M.Sc. and PhD in Electrical, Electronic and Systems Engineering from The National University of Malaysia, Malaysia, in 2008 and 2016, respectively. She worked as an Assistant Professor in the College of Communication Engineering at Future University, Sudan. She is currently the Head of the Department of Computer Engineering in the female branch, University of Ha'il, Saudi Arabia. Her current research interests include Wireless Communications, Mobile networks, Small Cells, Internet of Things.

How to cite this article: Tiwari, S., Jain, A., Ahmed, N. M. O. S., Charu, Alkwai, L. M., Dafhalla, A. K. Y., & Hamad, S. A. S. (2021). Machine learning-based model for prediction of power consumption in smart grid- smart way towards smart city. *Expert Systems*, e12832. https://doi.org/10.1111/exsy.12832