

Bootstrap Aggregating Approach to Short-term Load Forecasting using Meteorological Parameters for Demand Side Management in the North-Eastern Region of India

Dipu Sarkar · Taliakum Ao ·
Sravan Kumar Gunturi

Received: date / Accepted: date

Abstract Electricity is an essential commodity that must be generated in response to demand. Hydroelectric power plants, fossil fuels, nuclear energy, and wind energy are just a few examples of energy sources that significantly impact production costs. Accurate load forecasting for a specific region would allow for more efficient management, planning, and scheduling of low-cost generation units and ensuring on-time energy delivery for full monetary benefit. Machine learning methods are becoming more effective on power grids as data availability increases. Ensemble learning models are hybrid algorithms that combine various machine learning methods and intelligently incorporate them into a single predictive model to reduce uncertainty and bias. In this study, several ensemble methods were implemented and tested for short-term electric load forecasting. The suggested method is trained using the influential meteorological variables obtained through correlation analysis and the past load. We used real-time load data from Nagaland's load dispatch centre in India and meteorological parameters of the Nagaland region for data analysis. The synthetic minority over-sampling technique for regression (SMOTE-R) is also employed to avoid data imbalance issues. The experimental results show that

Dipu Sarkar (Corresponding Author)
Department of Electrical and Electronics Engineering
National Institute of Technology Nagaland, Dimapur-797103, India
Phone: +91-9089643843
E-mail: dipusarkar5@rediffmail.com

Taliakum Ao
Department of Electrical and Electronics Engineering
National Institute of Technology Nagaland, Dimapur-797103, India

Sravan Kumar Gunturi
Department of Electronics and Instrumentation Engineering,
National Institute of Technology Nagaland, Dimapur-797103, India

the Bagging methods outperform other models with respect to mean squared error and mean absolute percentage error.

Keywords Demand side management · Ensemble Learning · Random Forest · Extra Trees · Short-term load forecasting · SMOTE.

1 Introduction

Load forecasting is a pivotal step for electric grids' efficient operation and management in the day-ahead electricity market[1]. Owing to the rapid growth of renewable energy and the implementation of electric vehicles and other emerging technologies, the electric power grid has undergone significant changes in the last few years, both on the supply and demand lines[2]. To sustain an equilibrium between supply and demand in this setting, electricity suppliers must take advantage of the next-generation power grid's opportunities. Furthermore, in the event of unusually high peak energy demand, electric utility companies must be prepared in advance with contingency plans. As a result, it is essential to forecast future load demand to reduce the cost of generating electricity[15]. Electricity load forecasting is the process of predicting future load based on various features, including weather conditions, time details such as the month and hour, economic conditions, energy tariffs, regional conditions are examples. Since accurate electricity load forecasting is critical in the power system, even minor improvements in load forecasting accuracy can result in substantial cost savings and environmental benefits[5]. Besides that, any forecast error results in a substantial cost increase. As a result, effective load forecasting is critical for power-grid construction, investment, and transactions in order to ensure reliable and cost-effective power system operation[20].

In recent times, researchers have suggested multiple models for forecasting electric load for different periods. Short-term load forecasting(SLF) models forecast electricity demand every half-hour or hour for the next few hours to two weeks. Medium-and long-term models are established when long-term load forecasting is needed[8]. Regression models are good at estimating stationary time series [19]. Since the time series of load demand is nonlinear and non-stationary, the Auto-Regressive Moving-Average method (ARMA)[30,6] and Support Vector Regression (SVR)[28,32] methods were proposed. A single model is incapable of accurately capturing the inherent characteristics of electricity load demand. Hence, [30] integrated the ARMA approach with the other machine learning methods to boost the efficiency of combined load forecasting. [21] developed a hybrid model that combines clustering with the integrated ARMA model.

The SLF problem has been also studied using a variety of sophisticated data-driven models. Fuzzy logic[24], artificial neural networks (ANNs)[3,7], Extreme Learning[17,22], and exponential smoothing[25] are some of the models that have been used. Numerous hybrid models have been developed by combining multiple models in order to increase forecasting accuracy. Yang et al. suggested an advanced empirical mode decomposition approach to solve the

net outcome and the envelope fitting limitations inherent in traditional empirical mode decomposition-based models[31]. Liang et al. proposed a framework that combines EMD, minimal consistency, maximal relevance, and a neural network with an optimisation method[18]. Talaat et al. suggested a novel approach of a multi-layer front-propagated neural net and the grasshopper optimization technique to achieve accurate load forecasting[26]. The majority of the methods reported previously have the following limitations:

- The mean absolute percentage error is high for conventional methods.
- Only Time-series load data is considered.
- Most only focus on boosting strategies, not bagging approaches like Random Forest and Extra Trees.
- The majority of methods, such as SVR and artificial neural networks, need additional training and testing time.
- Several techniques have failed to solve the issue of data imbalance caused by oversampling.

The current research makes four significant contributions: (1) For specific-region SLF implementation, the proposed framework makes use of real-time load and meteorological data. (2) The implementation of the synthetic minority over-sampling technique for Regression addresses data imbalance (SMOTE-R). (3) The most advanced load forecasting models are checked and compared. (4) The proposed method achieves the least Mean Absolute Percentage Error by using Bootstrap aggregating models.

The remainder of the article is organised as follows: Section 2 addresses the theoretical basis of ensemble learning models. Section 3 discusses the proposed methodology. Section 4 gives the conclusion and future directions.

2 Theoretical Background for Ensemble Learning

Ensemble machine learning is a series of models that train a diverse group of potential learners and incorporate their hypotheses to teach the target's particular viewpoint[29]. Ensemble machine learning regressors were evaluated in this analysis, including Ada-Boost, Cat-Boost, Extreme-Gradient-Boost, Light-Gradient-Boost, Random Forest, and Extra-Trees[23].

2.1 Boosting Techniques:

Boosting is an ensemble technique in which every successive model tries to eliminate the previous model subsets' errors, as shown in Figure 1. Boosting templates such as Ada-Boost, Cat-Boost, Extreme-Gradient-Boost, and Light-Gradient-Boost were introduced. Ada-Boost[34] is used in aggregation with other algorithms to reduce the prediction error. Training a sequence of weak regressor models from the observed data package and combining them into a robust regressor is one way to build the most potent forecast model. Extreme-Gradient-Boost[11] expands Classification And Regression Tree(CART)-compliant

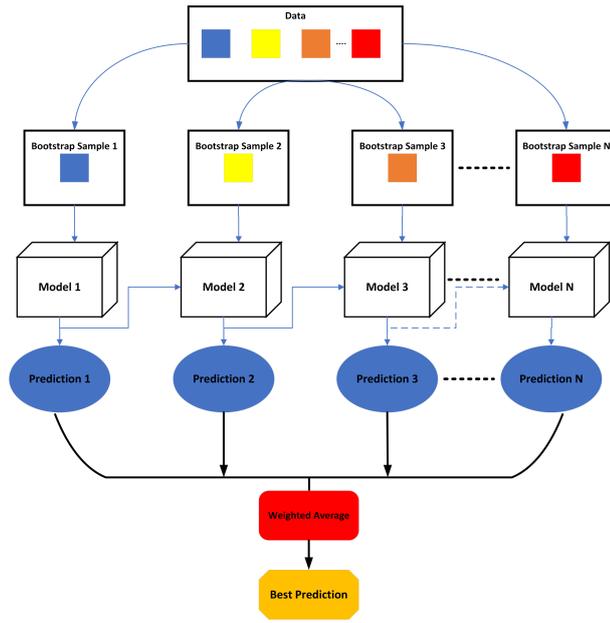


Fig. 1 Boosting model representation

classical boosting strategies. The fundamental role in the ensemble tree-boosting method predicts the group's new membership for each iteration. This is achieved in the most efficient way possible, ensuring that forecasts are made with soft regressors continuously constructed over the preceding regressor's fault. Following that, incorrectly classified models are assigned advanced weights, letting the regressor focus on subsequent recurrences' performance. The final categorization is critical because it requires the combined growth of all past decision trees. The evaluation of such regressors depends on the specific function, which defines the training failure and regularisation.

Cat-Boost[12, 14] is an application for gradient boosting that excels at handling categorical results. It makes no attempt to substitute discrete categorical attributes. It determines the average mark value by permuting a random dataset, such as adding the change's identical class value before the specified one. When confronted with definite category features, Cat-Boost avoids overfitting. Light-Gradient-Boost[16] is a decision tree-focused ensemble model. The advancing propagation process is used for this system. In each epoch, the residual is close-fitting with a negative slope for studying the decision tree. Light-gradient-Boost (introduced by Microsoft in 2016) is a significant Extreme-Gradient-Boost update that can fix issues by improving the sample size while the node is isolated. Light-Gradient-Boost employs a leaf-based method to increase forecast accuracy[13].

2.2 Bagging Techniques:

The objective of bagging is to achieve a standardised outcome by combining the outcomes of several decision trees. As seen in Figure 2, the Bootstrap Aggregating approach employs subsets to maintain a correct representation of the whole set. It is a technique of selecting several random samples from a training feature dataset. Random Forest[33] is a learning method that provides planning to construct multiple decision trees, resulting in a class that each tree averages or votes on. By adding a randomness layer to the bagging operation, Genuer et al. suggested a Random-Forest method[10]. Random-Forest is helpful for regression, grouping, and feature collection, among other things. Random-Forest has three distinct advantages. To begin, it reduces noise by selecting variables and data at random to create multiple regression trees. Second, Random-Forest can handle high-dimensional data and adapts to various datasets. It will work for both continuous and discrete data without the need for normalisation. Third, the parallel arithmetic's rapid learning speed reduces computational costs.

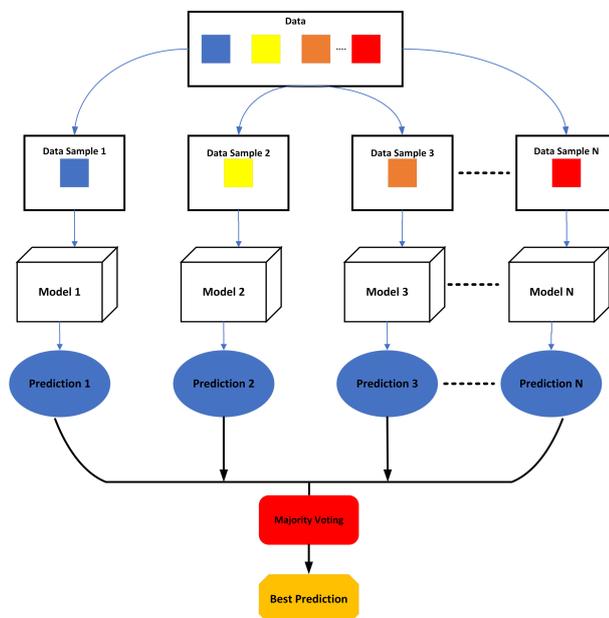


Fig. 2 Bagging model representation

Finally, an Extra-trees regressor uses a meta-learning algorithm to train several randomised decision trees. The averaging method is used to increase prediction accuracy and reduce over-fitting. Each decision tree is evaluated at each intersection by executing the following steps: 1) Select K attributes

randomly, 2) calculate the split for each feature, and 3) Select the feature that boosts the score using a normalised method of feature selection.

3 Proposed Method

This section addresses the proposed load forecasting model, which uses real-time load data and meteorological parameters as training data. The proposed scheme made use of data analysis collected from the Nagaland Load Dispatch Centre. The proposed system's block diagram is depicted in Figure 3. As shown

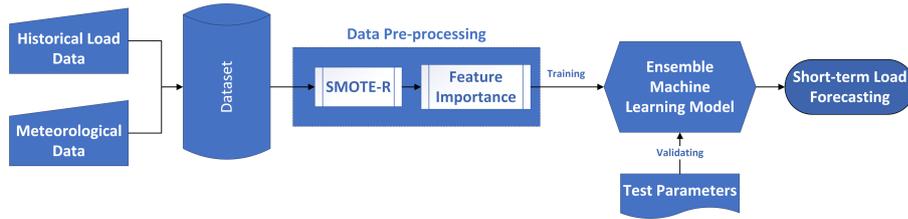


Fig. 3 Block diagram of proposed method

in the block diagram, the proposed system can be explained in three steps: (1) Feature Engineering. (2) Features extraction. (3) Use Ensemble Learning for Load forecasting.

3.1 Feature Engineering

Feature engineering is a method of converting operational variables into functions that best reflect the primary problem in analytical models and improve system performance for data that has not been used before. We used SMOTE-R to solve the data imbalance problem, which is discussed in the next section.

3.1.1 SMOTE for Regression (SMOTE-R)

In classification problems, SMOTE is a tool for sampling imbalanced class distributions. This approach has the advantage of combining normal class sub-sampling with minority class over-sampling. Bowyer et al. demonstrate the benefits of this method over other alternate sampling models on several real-world difficulties using several classification systems[4]. SMOTE-R, a version of SMOTE for regression tasks that correctly estimate rare extreme values, is proposed by Luis Torgo et al., who make a significant contribution[27]. The original SMOTE algorithm constructs fake cases with a rare target value using an over-sampling technique. To build these artificial cases, Fernandez et al. implemented an interpolation strategy[9]. The procedure involves selecting one of the k close neighbours for each event from a list of observations of uncommon

values. These two results result in a new scenario that combines attribute values with the two original examples' values. Both cases and synthetic cases belong to the same class as SMOTE is used to address classification issues using an unique class of concern in terms of the target variable. To build a general structure for predicting rare maximum values in a continuously varying target, the authors used a SMOTE-R technique combined with any other regression algorithm.

3.2 Important Feature Extraction

The feature pre-processing technique removes unimportant features from the CSV data before feeding it to the ensemble model. Feature importance metrics can aid in data interpretation, but they are often used to rank and pick features most relevant to a predictive model. The relative scores will indicate which features are most important to the target and, conversely, which features are least relevant. Using significance scores, feature importance can be used to construct a statistical model. Features with low scores will be removed, while features with high scores will be retained. The feature selection can help to simplify the problem being modelled, speed up the modelling process (deleting features is referred to as dimensionality reduction), and, in some cases, increase the model's efficiency.

To effectively deal with the system's unforeseen operating circumstances, the qualified model must be updated in actual implementations. As a result, a model upgrade mechanism is introduced to boost the generalization ability of the stability assessment model and achieve seamless online evaluation. Variable reduction is a technique for reducing the number of input variables. The more input features there are, the more complex the predictive modelling task becomes. Though dimensional reduction techniques are widely used for data visualisation, they may also be used for classifiers or regression models with higher-dimensional data.

3.3 Ensemble Load forecasting

As discussed in section 2, we use different ensemble learning models to forecast the load and we found that the Random Forest model is giving the best precision over the remaining ensemble models. The proposed ensemble load forecasting is illustrated using flow chart as seen in Figure 4.

4 Experimental Results

4.1 The Dataset and Feature Extraction

Ensemble SLF model is validated using real-time data from the Nagaland Load Dispatch Center in India's North-Eastern area. The dataset includes meteo-

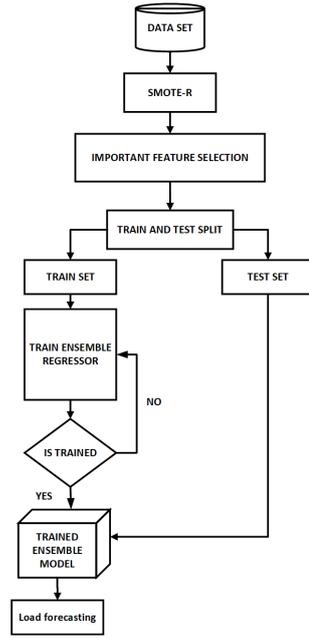


Fig. 4 Flow diagram of Proposed Ensemble SLF

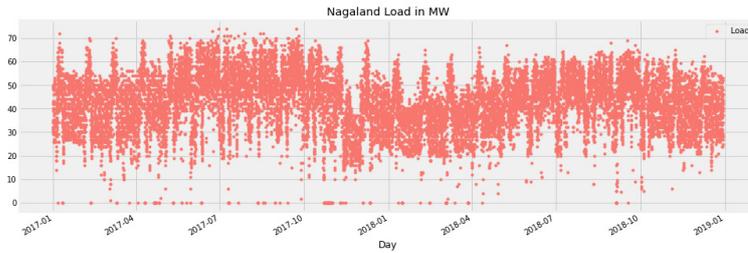
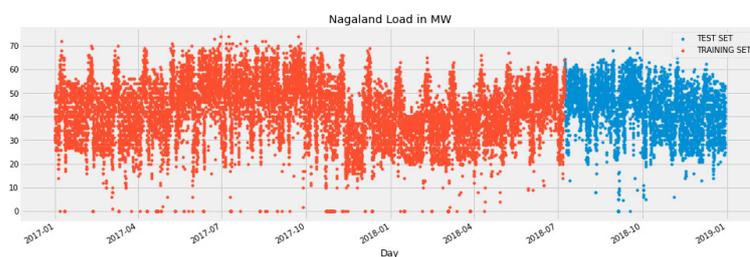


Fig. 5 Visualisation of Nagaland load consumption

rological parameters, categorical time data, and load demand for a defined sampling time. Table 1 provides additional information about the meteorological parameters used in the dataset. The proposed load forecasting model was trained using data from Nagaland's electric load demand between January 1st and December 31st, 2018. Hourly load data and 16 meteorological parameters were obtained over two years. Thus a total of 17520 samples (730 days*24 hours=17,520 samples) are being used as a database, as seen in Figure 5. As a result, the size of the data is 17520 x 17. This research used feature selection to increase the model's accuracy and effectiveness. Without feature selection, researchers would be forced to conduct additional unnecessary data training with and without specific features in order to assess whether or not those attributes have an impact on the prediction's accuracy. The current work divides the total data set into training and testing data sets with 80% and 20% split,

Table 1 Meteorological Parameter

Meteorological Parameter	Units
Temperature	Degree Centigrade
Wind Direction	Degrees
Humidity	Grams per cubic metre
Precipitation	Millimetres
Wind speed	Kilo meter per Hour
Pressure	Inches
Visibility	Kilo meter
Cloud Cover	Okta
Heat Index	Degree Centigrade
Dew point	Degree Centigrade
UV index	Index number
Wind Chill	Degree Centigrade
Feels Like	Degree Centigrade
Wind Gust	Kilo meter per Hour

**Fig. 6** Splitting load data in to Training and Testing sets

respectively as seen in Figure 6. Figure 7 illustrates the histogram plot of the dataset used in training. We selected features using the feature selection approach based on correlation coefficient and eliminated redundant features by evaluating the feature correlation matrix. The correlation matrix is an important data analysis metric used to analyse data to understand the relationship between different variables and make sound decisions. Each row and column in this matrix represents a variable with a value representing the coefficient of correlation between the variables defined by that row and column.

The Pearson correlation coefficient is the most basic method for determining the association between features and the target variable. The results' interval is $[-1, 1]$, where -1 represents a complete weak association (one variable rises while the other decreases), $+1$ represents a complete strong association, and 0 represents no linear association. Figure 8 shows the correlation map for the proposed system's training data. Each cell in the grid represents the correlation coefficient between two variables. It is a square matrix with each row containing a vector and all columns corresponding to the rows. The total number of columns is the same as the total number of rows. A high positive value (near 1.0) implies an excellent positive correlation, which means that when one variable's value rises, so does the other's value. The relationship

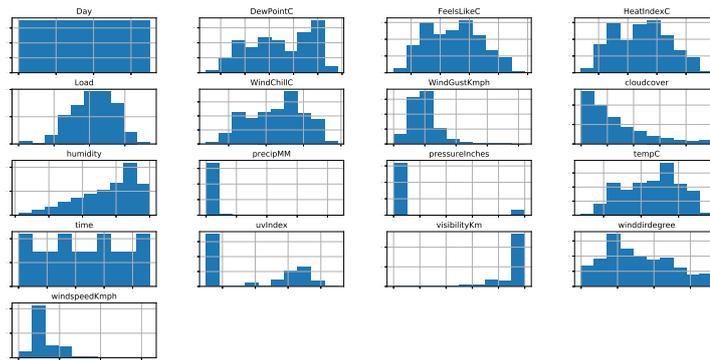


Fig. 7 Histogram

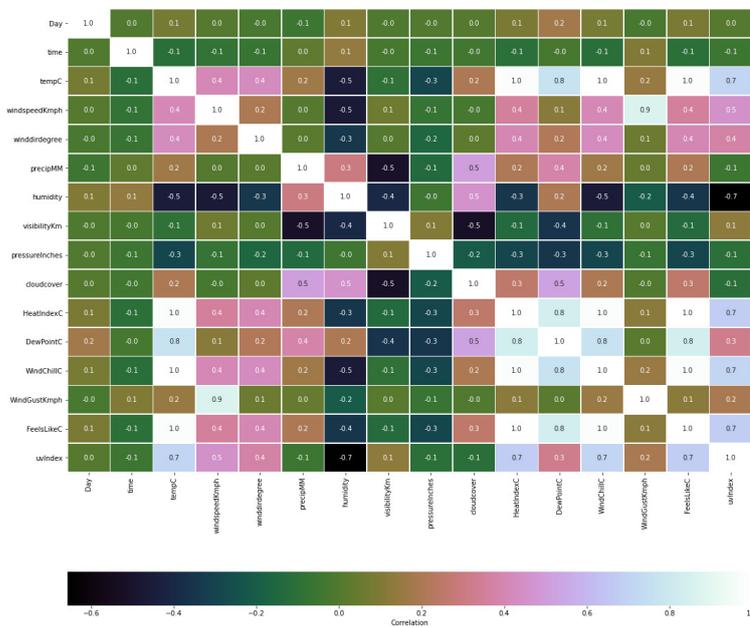


Fig. 8 Correlation plot

between a and b is the same as the relationship between b, and a. Diagonal cells reflect each variable’s correlation with itself. Hence, all diagonal elements are equal to one.

Along with correlation matrix, we have also used another method for selecting important features i.e., Recursive Feature Elimination (RFE). The RFE looks for a subset of the test-set attributes that begin with all of them and then successively eliminates them until the correct list remains in place. This

4.2 Evaluation of Ensemble Learning models

In the present study, we have used six ensemble regressor models including Ada-Boost, Extreme-Gradient-Boost, Cat-Boost, Light-Gradient-Boost, Random Forest, and Extra-trees Regressor models. Long Short-Term Memory(LSTM) and ARMA networks were also implemented and compared to demonstrate the proposed system's feasibility. First, we used cross-validation to determine which ensemble learning model will be the best fit for the proposed SLF. Figure 11 shows the comparison of different machine learning models cross validation performance with respect to Negative mean square error. The mean absolute error (MAE), mean absolute percentage error (MAPE), and mean square error (MSE) metrics have been used to assess the performance of ensemble learning algorithms. If \hat{y}_i indicates the expected value of i^{th} sample and y_i indicates the corresponding actual value. Then the MSE over $n_{samples}$ may be given as follows:

$$\mathbf{MSE} = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} (y_i - \hat{y}_i)^2. \quad (1)$$

The MSE calculates accuracy by comparing the amount of squared errors between forecasting and real values. The MSE value will evaluate the error variation; the smaller the MSE value, the more stable the prediction model will be. The MAE is a measure of the average amount of combined absolute error between predictions and actual values, which can help avoid error mutual cancellation. As a consequence, MAE will correctly represent the absolute prediction error.

$$\mathbf{MAE} = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} |(y_i - \hat{y}_i)|. \quad (2)$$

The relative difference between exact errors and real values is denoted by MAPE. The MAPE is a common metric for evaluating regression tasks.

$$\mathbf{MAPE} = \frac{100\%}{n_{samples}} \sum_{i=0}^{n_{samples}-1} \frac{|(y_i - \hat{y}_i)|}{|y_i|}. \quad (3)$$

The table 2 summarises machine learning models' evaluation using the MSE, MAE, and MAPE metrics without feature importance. As seen in the table, Random forest outperformed all other models, including LSTM, ARMA, and machine learning models. All state-of-the-art models were re-evaluated merely based on the important attributes obtained in the previous section 4.1. The Random forest algorithm, as seen in Table 3, has the lowest MSE(25.463), MAE(3.462), and MAPE(8.482) values of the other algorithms.

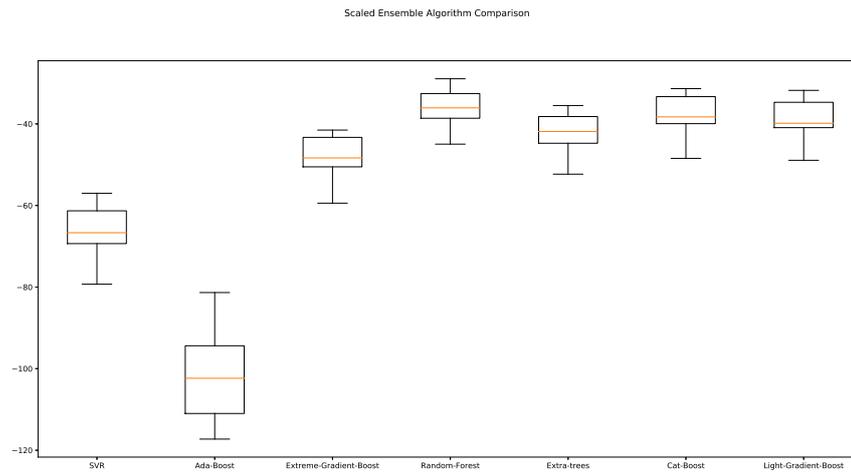


Fig. 11 Comparison of Machine learning models

Table 2 Evaluation of Machine Learning Models on Test set without Feature Importance

Model	MSE	MAE	MAPE
SVR	109.441	8.145	19.412
ARMA	42.316	4.194	10.584
LSTM	90.120	6.239	14.412
Ada-Boost	113.619	8.619	24.663
Extreme-Gradient-Boost	49.737	4.704	11.560
Light-Gradient-Boost	39.282	4.0767	10.335
Cat-Boost	37.535	4.016	10.031
Random Forest	33.895	3.758	9.675
Extra-Trees	39.543	4.077	10.408

Table 3 Evaluation of Machine Learning Models on Test set with Feature Importance

Model	MSE	MAE	MAPE
SVR	91.941	7.678	18.211
ARMA	30.359	3.818	9.415
LSTM	84.978	7.114	16.254
AdaBoost	65.124	6.415	16.670
Extreme-Gradient-Boost	34.007	4.167	10.104
Light-Gradient-Boost	26.009	3.649	8.612
Cat-Boost	28.968	3.649	8.612
Random Forest	25.463	3.462	8.482
Extra-Trees	28.461	3.722	9.057

5 Conclusion

In this analysis, an ensemble regressor was used to demonstrate a process of short term load forecasting in North-Eastern part of India. The suggested

scheme is dependent on meteorological parameters and historical patterns of load use. At the pre-processing data stage, SMOTE-R analysis was used to solve the data imbalance issue. Numerous ensemble algorithms based on boosting and bagging techniques were investigated, including Ada-Boost, Extreme-Gradient-Boost, Cat-Boost, Light-Gradient-Boost, Random Forest, and Extra-trees. The proposed design was tested and checked using historical load data collected from Nagaland's Load Dispatch Unit in real-time. For the Random Forest and Extra-tree Regressors, bagging methods yielded the lowest MAE values of 3.75 and 4.07, respectively. Boosting methods generated higher MAE and MAPE values than bagging methods and were thus unsuitable for real-time scenarios. Therefore, real-time use of bagging techniques such as Random forest and Extra-trees is facilitated. Additionally, the mean absolute error and mean absolute percentage error study revealed that bagging methods outperform the other methods. The authors plan to continue their work by developing a generative adversarial network-based energy forecasting scheme based on domain information.

Acknowledgements Not applicable

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding Statement

The authors received no financial support for the research, authorship, and publication of this article.

Authors Contribution

All the authors participated and contributed equally in the analysis and interpretation of the results and data, drafting the article or revising it critically and preparing the final version.

Availability of data and material

The data sets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Code availability

The custom code used in the current study is available from the corresponding author on reasonable request.

Ethics approval

Not applicable

Consent to participate

Not applicable

Consent to publish

Not applicable

References

1. Amato, U., Antoniadis, A., De Feis, I., Goude, Y., Lagache, A.: Forecasting high resolution electricity demand data with additive models including smooth and jagged components. *International Journal of Forecasting* **37**(1), 171–185 (2021)
2. Andriopoulos, N., Magklaras, A., Birbas, A., Papalexopoulos, A., Valouxis, C., Daskalaki, S., Birbas, M., Housos, E., Papaioannou, G.P.: Short term electric load forecasting based on data transformation and statistical machine learning. *Applied Sciences* **11**(1), 158 (2021)
3. Cecati, C., Kolbusz, J., Różycki, P., Siano, P., Wilamowski, B.M.: A novel rbf training algorithm for short-term electric load forecasting and comparative studies. *IEEE Transactions on Industrial Electronics* **62**(10), 6519–6529 (2015)
4. Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research* **16**, 321–357 (2002)
5. Chen, K., Chen, K., Wang, Q., He, Z., Hu, J., He, J.: Short-term load forecasting with deep residual networks. *IEEE Transactions on Smart Grid* **10**(4), 3943–3952 (2018)
6. Din, G.M.U., Mauthe, A.U., Marnierides, A.K.: Appliance-level short-term load forecasting using deep neural networks. In: 2018 International Conference on Computing, Networking and Communications (ICNC), pp. 53–57. IEEE (2018)
7. Dong, Y., Dong, Z., Zhao, T., Li, Z., Ding, Z.: Short term load forecasting with markovian switching distributed deep belief networks. *International Journal of Electrical Power & Energy Systems* **130**, 106942 (2021)
8. Fan, S., Hyndman, R.J.: Short-term load forecasting based on a semi-parametric additive model. *IEEE Transactions on Power Systems* **27**(1), 134–141 (2011)
9. Fernández, A., García, S., Galar, M., Prati, R.C., Krawczyk, B., Herrera, F.: *Learning from imbalanced data sets*. Springer (2018)
10. Genuer, R., Poggi, J.M., Tuleau-Malot, C.: Variable selection using random forests. *Pattern recognition letters* **31**(14), 2225–2236 (2010)
11. Georganos, S., Grippa, T., Vanhuysse, S., Lennert, M., Shimoni, M., Wolff, E.: Very high resolution object-based land use–land cover urban classification using extreme gradient boosting. *IEEE geoscience and remote sensing letters* **15**(4), 607–611 (2018)

12. Jhaveri, S., Khedkar, I., Kantharia, Y., Jaswal, S.: Success prediction using random forest, catboost, xgboost and adaboost for kickstarter campaigns. In: 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), pp. 1170–1173. IEEE (2019)
13. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.Y.: Lightgbm: A highly efficient gradient boosting decision tree. In: Advances in neural information processing systems, pp. 3146–3154 (2017)
14. Kumar, S., Singh, K., Kumar, S., Kaiwartya, O., Cao, Y., Zhou, H.: Delimitated anti jammer scheme for internet of vehicle: Machine learning based security approach. IEEE Access **7**, 113311–113323 (2019)
15. Li, F., Jin, G.: Research on power energy load forecasting method based on knn. International Journal of Ambient Energy pp. 1–6 (2019)
16. Li, F., Zhang, L., Chen, B., Gao, D., Cheng, Y., Zhang, X., Yang, Y., Gao, K., Huang, Z., Peng, J.: A light gradient boosting machine for remaining useful life estimation of aircraft engines. In: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), pp. 3562–3567. IEEE (2018)
17. Li, S., Goel, L., Wang, P.: An ensemble approach for short-term load forecasting by extreme learning machine. Applied Energy **170**, 22–29 (2016)
18. Liang, Y., Niu, D., Hong, W.C.: Short term load forecasting based on feature extraction and improved general regression neural network model. Energy **166**, 653–663 (2019)
19. Miraftebadeh, S.M., Foiadelli, F., Longo, M., Pasetti, M.: 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. IEEE (2019)
20. Moradzadeh, A., Moayyed, H., Zakeri, S., Mohammadi-Ivatloo, B., Aguiar, A.P.: Deep learning-assisted short-term load forecasting for sustainable management of energy in microgrid. *Inventions* **6**(1), 15 (2021)
21. Nepal, B., Yamaha, M., Yokoe, A., Yamaji, T.: Electricity load forecasting using clustering and arima model for energy management in buildings. *Japan Architectural Review* **3**(1), 62–76 (2020)
22. Park, R.J., Song, K.B., Kwon, B.S.: Short-term load forecasting algorithm using a similar day selection method based on reinforcement learning. *Energies* **13**(10), 2640 (2020)
23. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011)
24. Sadaei, H.J., Enayatifar, R., Abdullah, A.H., Gani, A.: Short-term load forecasting using a hybrid model with a refined exponentially weighted fuzzy time series and an improved harmony search. *International Journal of Electrical Power & Energy Systems* **62**, 118–129 (2014)
25. Sudheer, G., Suseelatha, A.: Short term load forecasting using wavelet transform combined with holt-winters and weighted nearest neighbor models. *International Journal of Electrical Power & Energy Systems* **64**, 340–346 (2015)
26. Talaat, M., Farahat, M., Mansour, N., Hatata, A.: Load forecasting based on grasshopper optimization and a multilayer feed-forward neural network using regressive approach. *Energy* **196**, 117087 (2020)
27. Torgo, L., Ribeiro, R.P., Pfahringer, B., Branco, P.: Smote for regression. In: Portuguese conference on artificial intelligence, pp. 378–389. Springer (2013)
28. Velasco, L.C.P., Polestico, D.L.L., Abella, D.M.M., Alegata, G.T., Luna, G.C.: Day-ahead load forecasting using support vector regression machines. *Int. J. Adv. Comput. Sci. Appl.(IJACSA)* **9**(3), 22–27 (2018)
29. Wan, S., Yang, H.: Comparison among methods of ensemble learning. In: 2013 International Symposium on Biometrics and Security Technologies, pp. 286–290. IEEE (2013)
30. Wei, L., Zhen-gang, Z.: Based on time sequence of arima model in the application of short-term electricity load forecasting. In: 2009 International Conference on Research Challenges in Computer Science, pp. 11–14. IEEE (2009)

31. Yang, X., Cheng, G., Liu, H.: Improved empirical mode decomposition algorithm of processing complex signal for iot application. *International Journal of Distributed Sensor Networks* **11**(10), 862807 (2015)
32. Yang, Y., Che, J., Deng, C., Li, L.: Sequential grid approach based support vector regression for short-term electric load forecasting. *Applied Energy* **238**, 1010–1021 (2019)
33. Zhang, D., Qian, L., Mao, B., Huang, C., Huang, B., Si, Y.: A data-driven design for fault detection of wind turbines using random forests and xgboost. *IEEE Access* **6**, 21020–21031 (2018)
34. Zhao, H., Yu, H., Li, D., Mao, T., Zhu, H.: Vehicle accident risk prediction based on adaboost-so in vanets. *IEEE Access* **7**, 14549–14557 (2019)