Cureus Journal of Computer Science

Open Access Original Article

EStack Fusion Model For Energy Demand Forecasting

Jyothy Joseph [1] , Amal C R [1]

1. Department of Computer Science, Al Ameen College, Edathala, Aluva, IND

Corresponding authors: Jyothy Joseph, jyothyjoseph@gmail.com, Amal C R, amalcr105@gmail.com

Received 02/26/2025
Review began 03/03/2025
Review ended 03/22/2025
Published 03/26/2025

DOI: https://doi.org/10.7759/s44389-025-03476-w

Abstract

Precise forecasting of energy demand is important for good energy management and grid stability at different times. Energy demand is subject to the effects of both temperature and humidity along with other seasonal or periodic influences, making the forecasting task sufficiently complicated. This study presents the EStack Fusion Model: an ensemble method of XGBoost, Gradient Boosting, and Long Short-Term Memory (LSTM) networks in order to improve forecasting accuracy. For prediction, the model uses key features in terms of energy demand, temperature, humidity, holiday status, and month. Data preprocessing procedures, such as normalization and missing value treatment, have been implemented to ensure data quality. The dataset was split into training and testing sets, and various measures such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R^2) for evaluating the model's performance were used. EStack Fusion outperforms individual models, with a 5.7411 MSE, 1.3601 MAE, and 0.8391 R^2 showing a 1.45% improvement from the best standalone model. Proposed ensemble method captures non-linear dependencies and temporal variations. It is suitable for real-world problems such as smart grid management, demand response systems, and renewable energy forecasting. However, computational efficiency and scalability for real-time deployment remain crucial research issues. Future research will focus on optimal techniques, lightweight architectures, and adaptive learning models to further enhance prediction accuracy and efficiency.

Categories: IoT Applications, Data Science Methodologies, Machine Learning (ML)
Keywords: xgboost, gradient boosting, lstm, ensemble learning, estackfusion, energy demand forecasting, machine learning

Introduction

Forecasting energy demand plays an indispensable role in the efficient management of energy and in the stability of the grid. It enables energy suppliers as well as grid operators to predict changes, manage assets, and save operational costs for accurate forecasting of energy consumption. However, complexity in energy demand remains due to the effects of meteorological conditions such as temperature and humidity coupled with the variations of the seasons and holidays. These factors make energy demand forecasting using traditional methods often insufficient to capture dynamic and non-linear relationships [1-3]. Accurate forecasting becomes even more relevant when renewable energy sources are factored in, as these exhibit variations due to weather conditions. The demand patterns of electricity consumption also differ depending on the time of the year, day, and type of environment, making it imperative for forecasting models to consider such intricate links [4-6]. Kerala's energy consumption patterns are influenced by its tropical climate, varying seasonal patterns, and socio-economic factors, making it an ideal subject for studying energy demand. Traditional frameworks such as time series, autoregressive model application, and linear regression, which are widely used to forecast energy demand, are not able to cope with complex dependencies associated with the impact of fluctuations of renewable energy and external weather factors [7,8]. Machine learning (ML) presents promising advancements in this field. Algorithms such as XGBoost, Gradient Boosting, and Long Short-Term Memory (LSTM) networks have been successfully applied for energy demand forecasting [9,10]. Boosting models like XGBoost and Gradient Boosting are really great at finding some structure in relationships, but they quite misrepresent any temporal aspect and therefore are not helpful for sequential forecasting. In contrast to that, LSTM networks serve well for time-series predictions but are quite expensive in terms of computing resources and also overfit very quickly unless tuned properly. The performance of such models is further improved when used within ensemble learning frameworks that leverage the strengths of individual models to enhance prediction accuracy. This study proposes the EStack Fusion Model, a stacking-based ensemble ML model that integrates the complementary strengths of XGBoost, Gradient Boosting, and LSTM networks to predict future energy demand. The objective of this approach is to enhance prediction accuracy in energy demand forecasting, which can be important for energy management and policy formulation.

Literature review

Energy demand forecasting is important in ensuring proper energy management by balancing demand and supply in the least-cost manner. Different ML models have been attempted in energy forecasting. Piyal et

How to cite this article
Joseph J, C R A (March 26, 2025) EStack Fusion Model For Energy Demand Forecasting. Cureus J Comput Sci 2 : es44389-025-03476-w. DOI https://doi.org/10.7759/s44389-025-03476-w

al. [1] detailed their findings on energy demand forecasting in Bangladesh, in which they identified the potential of ML for demand forecasting in developing nations. Singh et al. [2] dealt with seasonal load forecasting and investigated meteorological variables as external factors that improve forecasting reliability. At the same time, a multi-model approach was adopted by Mubarak et al. [3], which entailed seasonal decomposition to improve short-term reactive energy forecasting. Other examples include Ali et al. [4], who applied ML techniques to heat demand forecasting for district heating networks, thus going beyond pure electrification in demand forecasting. While the opposite of this is to be found in AL-Musaylh et al. [5], who applied ground-based climatic and atmospheric data for forecasting of demand for electricity in Australia: climate information has been proven to increase accuracy in predicting cases. Chapagain et al. [6] provided proof of the need for implementing deep neural networks in short-term forecasting of electricity demand and additionally proved that advanced techniques should impart even more accuracy in forecasts. This collection of seasonal and external data has been so recurrent that authors from [7], applied deep learning in energy forecasting according to seasonal variations. This area was taken further by Hasanah et al. [8] in their research on holiday-based electric load forecasting, thus stating the need to consider local occasions when predicting future trends. Chapagain et al. [6] gave an approach to short-term forecasting with specific energy requirements in Thailand, thus stressing the need for area-specific models. The application of ensemble techniques in improving forecasting models has gained much importance. Al-bayaty et al. [9] compared various ML models for city-scale energy demand forecasting, and their studies proved that combining models yields better accuracy than single models. Likewise, Naveenkumar et al. [10] conducted their work on energy generation forecasting with the use of deep learning and showed that model stacking can give better results. In the past few years, ensemble learning techniques and deep learning models have acted as a catalyst for development in energy demand forecasting. These developments emphasize the need to use multiple models in order to improve the accuracy of forecasting. This research introduces EStack Fusion, an ensemble method that uses XGBoost, Gradient Boosting, and LSTM to improve the accuracy of forecasting.

Materials And Methods

Forecasting energy demand using a rich diversity of both meteorological and historical energy consumption data is the essence of the study. The main steps of the methodology would include the following:

Data collection

The dataset upon which this study is based has been put together from collecting information from several reliable sources, defining key features that would contribute to energy demand and meteorological factors associated with it. The most primary variable, Energy Demand (mu) is obtained from Kerala State Electricity Board (KSEB) [11] dealing with energy consumption in Kerala. The feature "Month" is included considering the variation of demand based on the seasons, where it would reveal the cyclic nature of energy demand at different periods of the year. The feature "Holiday", retrieved from Kerala Government's holiday calendar [12], gives whether a day is a holiday or not, allowing one more factor that illustrates holiday effects on energy consumption. Meteorological parameters covered are average temperature (Avg Temp °C) and average humidity (Avg Humidity %). All these are taken from the Weather and Climate [13] platform that provides reliable climate data. Consolidating all these diverse data sources, this dataset provides a comprehensive view of the factors that mold energy consumption patterns and furnish a strong base for an accurate energy demand forecast.

Data preprocessing

Data preprocessing mainly involves imputing missing values and removing outliers using the Interquartile Range method, one-hot encoding for the Month, and StandardScaler normalization encoding numerical features to fit into the same as the whole set.

Data splitting

This dataset was used for creating an 80-20 training-testing dataset, with 80% used as training data and 20% used as testing data. In this way, the model learns from the data it trains on, keeping a separate test set away for model evaluation.

Model development

The selected models proved to predict the effectiveness of energy consumption with a specific combination of aspects such as seasonal variation, meteorological condition, and holidays from features. We deploy three base models here, namely, XGBoost, Gradient Boosting, and LSTM, and then develop an ensemble model, called EStack Fusion, in order to improve the forecasting accuracy.

Existing Models

XGBoost (Extreme Gradient Boosting): XGBoost is a great and coded version of the ensemble learning algorithm for gradient boosting that sequentially creates decision trees, where each current tree attempts to fix the mistakes made by the last, adding a model. The model is trained on historical energy demand data, which includes temperature, humidity, and holidays, for the future prediction of energy demand. Although effective, it is sensitive to noise and outliers, and thus performance could be seriously degraded. Furthermore, it does not cognize well temporal dependency in time series data, compared to typical reference models for modeling sequential tasks like LSTM. That can limit the performance of the model in energy demand, having strong temporal features.

Gradient Boosting: Just like XGBoost, Gradient Boosting constructs a collection of decision trees serially with the primary aim of eliminating previous errors. It has been trained on the same dataset to compare XGBoost performance in predicting energy demand. Although it is very efficient at regression, gradient boosting also exposes excessive overfitting, especially when the number of trees grows or when the data are noisy. In addition to that, it also fails with time-dependency in energy demand like XGBoost, which may certainly nullify its efficacy when predicting time-bound tasks.

Long Short-Term Memory (LSTM): It is the best model when it comes to time series forecasting among all recurrent neural network models, as it resolves long-term dependency issues in sequential data. For the case study, temporal features, including past energy consumption and weather-related parameters, were used to forecast energy consumption using this type of model. However, it is expensive in terms of computation as it requires huge training times, and the cost especially increases as the size of the data becomes bigger. In addition, LSTM models are prone to overfitting if not tuned properly. Moreover, these models also do have performance drawbacks if the data they are fed do not show adequate variability or are inconsistent in time patterns.

Proposed Model

EStack Fusion (Stacking Model): The proposed EStack Fusion Model is a stacking-based ensemble method that uses predictions made by various base models, including XGBoost, Gradient Boosting, and LSTM. Under this stacking method, predictions made by every base model serve as input to a meta-model for generating prediction output. This technique helps the model take the strength of every individual base model and use it in improving overall prediction accuracy of results. In fact, using independent error reductions and the aggregation of their forecasts, EStack Fusion provides a more reliable and accurate solution to the prediction of energy demand. Figure *1* illustrate the architecture of proposed EStack Fusion model.

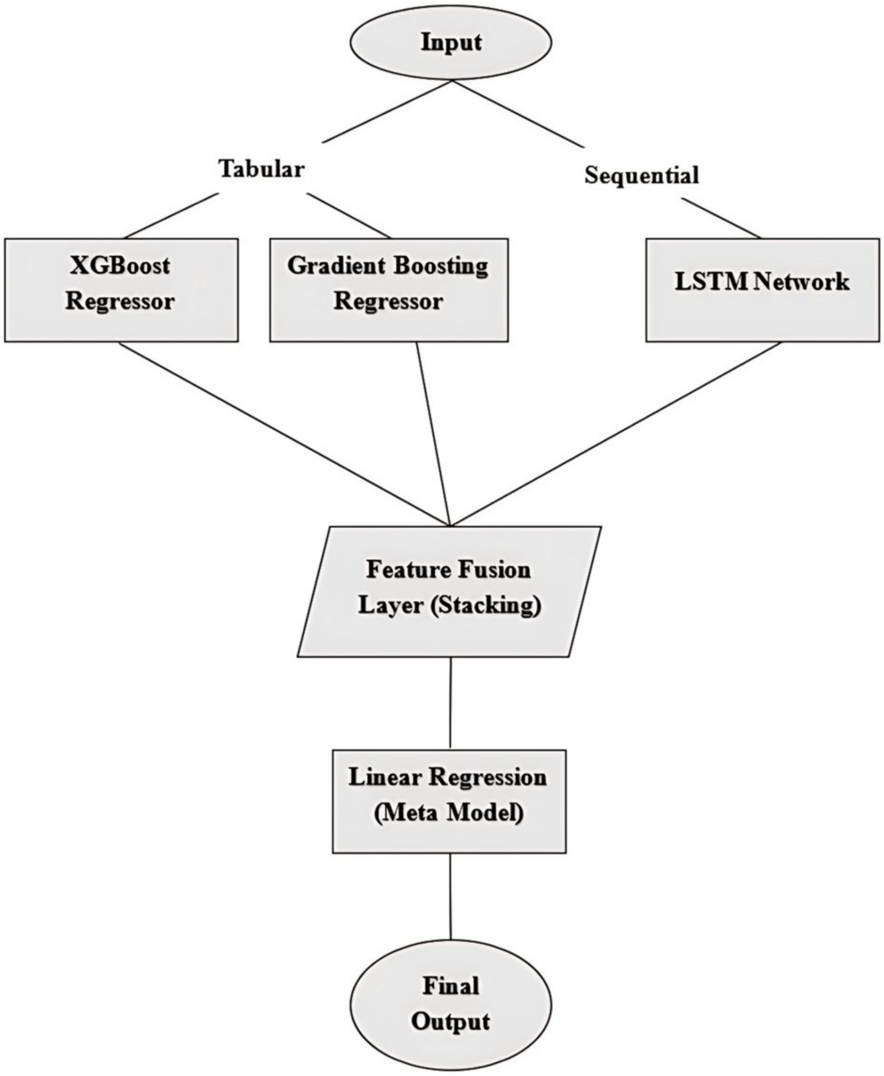

FIGURE 1: Architecture of EStack Fusion

LSTM, Long Short-Term Memory

Model Evaluation

K-Fold Cross-Validation was implemented to keep model performance sound, considering that the dataset was categorized into five folds (k = 5) to avert potential overfitting and thereby confer validation of model generalization. Training and testing sets were iteratively carved out of the entire dataset for each fold, so that the models could be trained on varying data subsets. The performance of the models and that of the ensemble were evaluated in terms of Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R^2) score, averaged over all folds. These metrics provided a sense of the accuracy and stability of the models.

Scatter plots of actual versus predicted values were also plotted to undertake visual assessment of model fit, making sure that the ensemble could also cope well with complex demand variations.

Results

For the evaluation of the proposed framework for energy demand prediction, the performance of individual base models XGBoost, Gradient Boosting, LSTM and the ensemble EStack Fusion Model was evaluated through various measures such as MSE, MAE, R^2. The evaluation results are listed in Table *1*.

Model	MSE	MAE	R^2
LSTM	11.3682	2.6060	0.6815
XGBoost	6.7523	1.3179	0.8108
Gradient Boosting	5.8845	1.3796	0.8351
EStack Fusion	5.7411	1.3601	0.8391

TABLE 1: Model performance results

LSTM, Long Short-Term Memory; MSE, Mean Squared Error; MAE, Mean Absolute Error

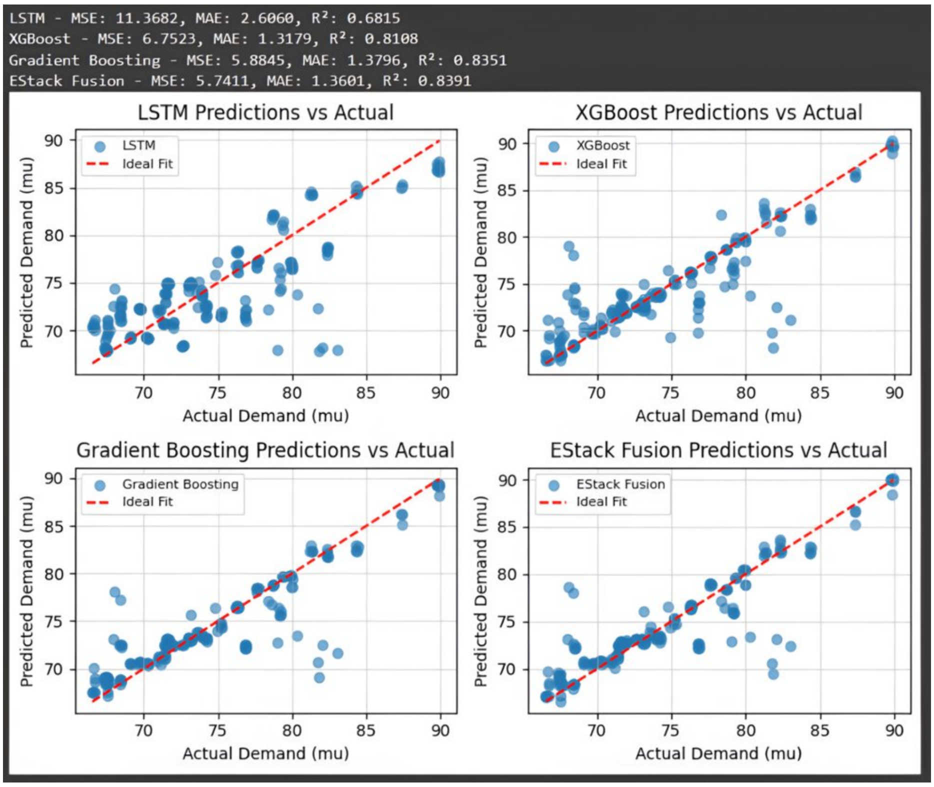

FIGURE 2: Results

LSTM, Long Short-Term Memory; MSE, Mean Squared Error; MAE, Mean Absolute Error

The comparison of predicted and actual demand confirms these observations. Figure 2 shows the results, which demonstrate that the EStack Fusion model cluster points closest to the ideal-fit line, demonstrating its precision concerning the variation of the energy demand. Gradient Boosting took second place, outperforming XGBoost and LSTM, which were both widened out and had less accurate outputs.

EStack Fusion has made its strong mark, but there is still scope for improvement. With each new addition of economic indices, historical usage patterns, and new weather data there seem further refinements possible in the predictions. This may represent possibilities for other ensemble methods or hybrid models for further future developments in energy demand forecasting.

Discussion

The EStack Fusion model performed best among all models, reporting the least MSE value of 5.7411, MAE of 1.3601, and R^2 of 0.8391, thus demonstrating its unparalleled predictive power. This model exhibits non-linear relationships and time dependencies in the energy demand data by combining Gradient Boosting and LSTM.

In comparison, the model having the lowest performance was LSTM, with an R^2 of 0.6815, suggesting its general failure to identify basic patterns. XGBoost and Gradient Boosting performed quite well by themselves with R^2 values of 0.8108 and 0.8351, respectively. However, the accuracy of these algorithms has not yet been matched by EStack Fusion, thus presenting the strength of the ensemble model contribution in complex forecasting problems. The comparison of predicted and actual demand confirms these observations. The EStack Fusion model clusters points closest to the ideal-fit line, demonstrating its precision concerning the variation of the energy demand. Gradient Boosting took second place, outperforming XGBoost and LSTM, which were both widened out and had less accurate outputs.

EStack Fusion has made its strong mark, but there is still scope for improvement. With each new addition of economic indices, historical usage patterns, and new weather data there seem further refinements possible in the predictions. This may represent possibilities for other ensemble methods or hybrid models for further future developments in energy demand forecasting

Conclusions

This study proposed an energy demand forecasting method by implementing an EStackFusion model, which is a simple stacked ensemble of LSTM, XGBoost, and Gradient Boosting. Results show that EStack Fusion provides a significant improvement compared to individual models by recording metrics like overall MSE, MAE, and R^2.The model captures non-linear dependencies and temporal variations effectively and is therefore applicable in smart grid management, demand response systems, and renewable energy forecasting. Yet, the computational effectiveness and scaling are the main problems. It is necessary to investigate the possibility of the real-time operation improvement by the selection of the best parameters, processing of the models, or parallelization. At the same time, the possibilities to increase flexibility and adaptability in distributed energy systems with the help of hybrid AI technologies should be analyzed, which will further improve the accuracy, speed, and practicality of EStack Fusion for real-time forecasting.

Appendices

Appendix A: dataset information

For this work, we created a unique dataset specifically tailored to the experiment. The dataset includes details extracted from the following sources:

=> Kerala State Electricity Board (KSEB) Title: Energy Demand Data for Kerala ,Available:https://www.kseb.in

=> Weather and Climate Title: Average Temperature and Humidity Data for Kerala ,Available:https://weatherandclimate.com

=> Kerala Government Title: Holiday Calendar for Kerala ,Available:https://kerala.gov.in

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Jyothy Joseph, Amal C R

Acquisition, analysis, or interpretation of data: Jyothy Joseph, Amal C R

Critical review of the manuscript for important intellectual content: Jyothy Joseph

Supervision: Jyothy Joseph

Drafting of the manuscript: Amal C R

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

We acknowledge Al-Ameen College for giving us the opportunity to work on this project. The support and guidance from our institution have been invaluable in shaping our efforts and we would also like to extend our sincere thanks to the Kerala Government for providing access to crucial data, which played a key role in our work. In particular, we referred to: Kerala State Electricity Board (KSEB): Energy demand data for Kerala (https://www.kseb.in) Weather and Climate: Average temperature and humidity data for Kerala (https://weatherandclimate.com) Kerala Government: Official holiday calendar for Kerala (https://kerala.gov.in) This project would not have been possible without these valuable resources. We deeply appreciate the efforts of everyone who contributed, directly or indirectly, to its success.

References

1. Piyal AP, Paul A, Ahmed S, Rahman KF, Mohsin ASM: Energy demand forecasting using machine learning perspective Bangladesh. 2023 IEEE IAS Global Conference on Renewable Energy and Hydrogen Technologies (GlobConHT), Male, Maldives. 2023, 1-5. 10.1109/globconht56829.2023.10087679
2. Singh B, Shah OA, Arora S: Seasonal electrical load forecasting using machine learning techniques and meteorological variables. International Journal of Robotics and Control Systems. 2024, 4:1092-108. 10.31763/ijrcs.v4i3.1446
3. Mubarak H, Sanjari J, Bai F, Stegen S, Abdellatif A: Enhanced short-term reactive energy demand forecasting by employing seasonal decomposition and multi-model approach. 2023 IEEE International Conference on Energy Technologies for Future Grids (ETFG), Wollongong, Australia. 2023, 1-6. 10.1109/etfg55873.2023.10407864
4. Ali H, Dedov A, Volkova A: Exploring heat demand forecasting in district heating networks using random parameter linear regression model. Environmental and Climate Technologies. 2024, 28:670-85. 10.2478/rtuect-2024-0052
5. AL-Musaylh MS, Deo RC, Adamowski JF, Li Y: Short-term electricity demand forecasting using machine learning methods enriched with ground-based climate and ECMWF reanalysis atmospheric predictors in southeast Queensland, Australia. Renewable and Sustainable Energy Reviews. 2019, 113:109293. 10.1016/j.rser.2019.109293
6. Chapagain K, Gurung S, Kulthanavit P, Kittipiyakul S: Short-term electricity demand forecasting using deep neural networks: an analysis for Thai data. Applied System Innovation. 2023, 6:100. 10.3390/asi6060100
7. Taleb I, Guerard G, Fauberteau F, Nguyen N: A flexible deep learning method for energy forecasting. Energies. 2022, 15:3926. 10.3390/en15113926
8. Hasanah P, Sari CA, Dinata SAW: Electricity load forecasting in east Kalimantan on religious holidays using SARIMA. Mathematical Journal of Modelling and Forecasting. 2023, 1:54-64. 10.24036/mjmf.v1i2.15
9. Al-bayaty H, Mohammed T, Ghareeb A, Wang W: City scale energy demand forecasting using machine learning based models: a comparative study. DATA '19: Proceedings of the Second International Conference on Data Science, E-Learning and Information Systems. 2019, 1-9. 10.1145/3368691.3368716
10. Naveenkumar S, Rawat P, Ganesh A, Ramakrishnan A, Sankaranarayanan S, Khairuddin ASM: Energy generation forecasting based on seasonality using deep learning. Proceedings of International Technical Postgraduate Conference 2022. 2022, 28-34.
11. Kerala State Electricity Board (KSEB). (2023). https://kseb.in/dashboard/view/consumption.
12. Weather and Climate, 'Average Temperature and Humidity Data for Kerala'. (2023). https://weatherandclimate.com.
13. Kerala Government, 'Holiday Calendar for Kerala'. (2023). https://kerala.gov.in.