

Light Weight Stacked Ensemble Model for Electric Load Forecasting using Fused Health Index

Anantha Krishna Kamath

Department of Computer Science and Design Engineering,
Canara Engineering College, Mangalore, Visvesvaraya Technological University, Karnataka, India.
Corresponding author: akkamath2109@canaraengineering.in

B. L. Rajalakshmi Samaga

Department of Electrical and Electronics Engineering,
Nitte (Deemed to be University), NMAM Institute of Technology (NMAMIT), Nitte, Karnataka, India.
E-mail: raji.samaga@nitte.edu.in

G. K. Dayananda

Department of Electronics and Communication Engineering,
Canara Engineering College, Mangalore, Visvesvaraya Technological University, Karnataka, India.
E-mail: dayananda@canaraengineering.in

(Received on January 7, 2025; Revised on April 9, 2025 & June 4, 2025 & July 29, 2025 & September 8, 2025;
Accepted on September 12, 2025)

Abstract

The operation of electric power system is a continuous process which demands coordination of various entities from power generating plants to distributing substations to render uninterrupted service still sticking to quality power delivery. Electric demand depends on external factors such as temperature, humidity, social activity pattern. Power grids are becoming complex due to integration of renewable energy sources. Thus, there is a need of electric energy forecasting. Efficiency of traditional forecasting approaches are less and existing many learning and ensembled models require high computational resources. To improve accuracy of prediction, a fast and efficient processing model which is suitable for real-time applications, light weight ensemble model is proposed in this paper. This research proposes a novel stacked light weight ensemble model that integrates the prowess of various weak base learners. The final prediction of the model is further improved by using extreme gradient boosting as a meta learner, which evolutionarily learns the predictions from individual learners and gives the final load forecast. Further the temporal nature of the exogeneous variables is preserved by a unique feature fusion technique which estimates the exponentially weighted moving average of the individual variable which are then aggregated. The efficacy of this model is validated by testing it on Panama electricity load forecasting dataset and the results are explored using important regression-based metrics. The analysis shows that the proposed method can vividly forecast the electricity load using the lightweight ensemble model in terms of Root Mean Square Error (RMSE), Mean Bias Error (MBE), Mean Absolute Error (MAE) and R^2 values.

Keywords- Feature fusion, Exponentially weighted moving average, Short term load forecasting, Ensemble model, Meta learner.

1. Introduction

Any country's development directly relies on its power sources, infrastructure, availability and distribution of energy. The transformation of human lives to acclaim great deal of comfort and easiness can be attributed to the improvement in energy generation, dissipation, distribution, and management (Zhang et al., 2018). Hence, the global demand for electricity is ever growing and its prices have fluctuated in the recent past. The huge gap between the global electricity demand and its generation has necessitated efficient planning in distribution and conservation with modern technologies. To augment this, the generation, distribution and transmission of electrical energy is highly complex and is not cost effective. This paved a motivation for several studies and research in the area of electricity forecasting for estimating the electricity demand

to facilitate the electricity generators, suppliers, and distributors to effectively devise plans much ahead and also to create awareness about energy conservation among the electricity users. **Figure 1** portrays the accelerated growth in global electricity consumption from 1980 to 2023 expressed in terawatt-hours (TWh). (Global Electricity Consumption 2023| Statista, n.d.). The graph reveals a continuous upward growth, with consumption rising from 7000 TWh in 1980 to over 25000 TWh in recent years. This rapid growth is driven by multiple factors such as industrial growth, urbanization, technological advancements, population growth and the shift toward electrification.

The effective electric grid management protocols involve appropriate planning about the load demand planning, maintenance schedule for generation, distribution lines and transmission along with effective distribution of load along the supply lines (Shen et al., 2016). Proper planning of electric load involves identification of forecasting intervals or lead time which forms a major criterion in classifying load forecasting into four major categories namely very short-term forecasting, short-term forecasting, medium or moderate term forecasting and long-term forecasting. The characteristic features of each of the forecasting type is described in **Table 1** (Verdejo et al., 2017).

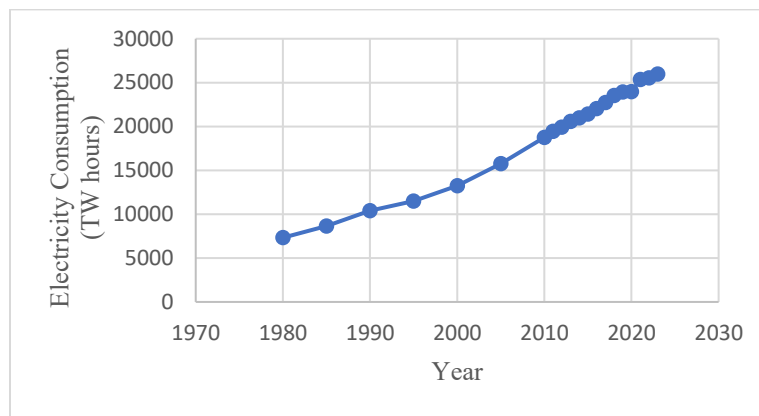


Figure 1. Growth in global electricity consumption.

Table 1. Characteristics of various load forecasting.

Type	Applications	Predicting period
Very short-term	Real time control	1 hour
Short term	Scheduling electricity for Daily life operations	1 hour-7 days (or even months)
Medium term	Purchase of fuel, maintenance scheduling and utility assessments	1 week-1 year
Long term	Development and construction of new generations, strategic power planning, tapping for new alternate power resource, demand side management.	1 -20 years

Forecasting Electricity Energy Demand (EED) is a complicated tasks that involves assimilation of multiple heterogeneous factors as shown in **Figure 2**. Nevertheless, EED forecasting is very complex and cannot be solved using sinister mathematical expressions and formulas. The problem of EED is not confined only in industrial sectors but extends its wings to cater both commercial and domestic needs. This necessitates detailed investigation about the dynamic growth patterns for an envisioning an accurate power system. On the other hand, inaccurate predictions may lead to power shortage, and increased expenditure. Hence a proper balance has to be maintained to supporting the development of the electrical power systems.

It is evident that accurate load forecasting will aid the planning process which will eventually reduce the gap between the global electricity demand and energy production. The process of EED forecasting can be done using many proven scientific methods which may be either computation or statistics based (Kuster et al., 2017). The conventional forecasting techniques can be categorized into correlation, extrapolation, and their hybrid (Almeshaiei & Soltan, 2011). The Extrapolation or trend analysis proceeds by fitting a curve on historical electricity demand data which depict the growth trend. The future value of can be forecasted by estimating the curve function at the desired point. Though this method is very realistic and simple, its results are not very accurate. Alternatively, correlation techniques focused on investigating the system load pertaining to various demographic and economic factors and capture the relation between the demand pattern and the factors influencing the demand. But the bottlenecks in accurate forecasting of multifaceted demographic and economic factors is very tedious and inconsistent.

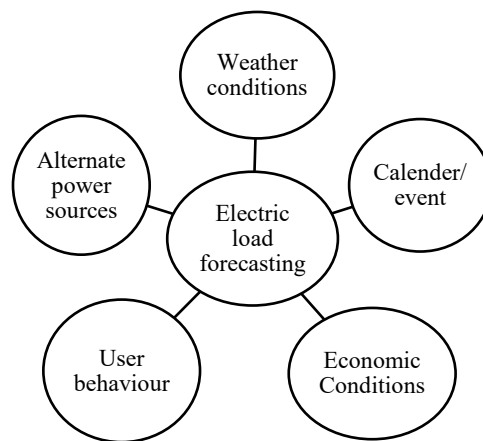


Figure 2. Major influential factors in electric load forecasting.

The recent developments in computing technologies, has introduced new paradigms to develop advanced forecasting models using data-driven Artificial Intelligence (AI) based techniques. Among the various flavors of AI, Machine Learning (ML) and Deep Learning (DL) paragon are gaining momentum because of their ability to render automatic forecasting or prediction with continuous improvement at much reduced human effort and time. These learning strategies tend to observe and learn the historic data and give future predictions with being explicitly programmed. **Figure 3** shows the functioning of ML algorithms.

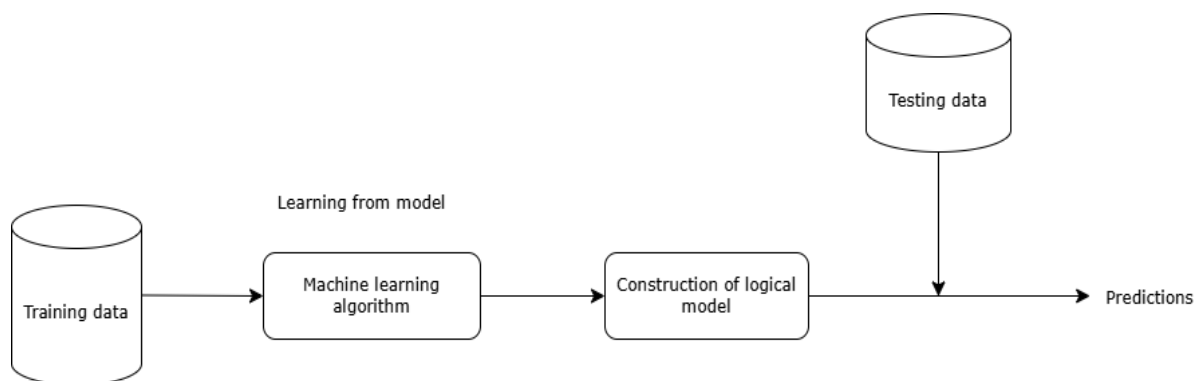


Figure 3. Overview of machine learning algorithm.

ML algorithm learns the patterns and trends from the previous historical data and constructs the model paste on the data. The new data applied to the logical model which gives the prediction results. The ML and DL algorithms are categorized into four major types namely supervised, unsupervised, semi supervised and reinforcement learning approaches. The algorithms like Support Vector Machines (SVM), Simple and multinomial Linear regression, Long Short-Term memory, Decision trees, Random Forests etc. are predominantly used in many applications because of their predictive power by learning intricate trends and patterns from the data space.

From the perspective of EED and electricity load forecasting, a lot of research have been done in analyzing the impact of AI based models in predicting the power loads (Pai & Hong, 2005). As the demand for electricity is not always consistent and is subjected to dynamic changes, these models have been very helpful in arriving at accurate and timely predictions. Reduced operational cost, stabilised power supply, efficient load management, security and safety off the power systems are some of the potential benefits that can be benefitted by deploying ML and DL models (Almashaiei & Soltan, 2011). In the view of developing an efficient electricity load forecasting method, this work divided into three phases as follows:

- i) Phase 1: Construction of fused health index using exogenous variables for electric load forecasting using exponentially weighted moving average. In this phase, the data from various sources are collected which are influencing electricity consumption rate such as temperature, wind speed. Also, robust dataset is prepared by handling missing values, outlier values and categorical data if any using exponentially weighted moving average method.
- ii) Phase 2: Prediction of electricity load demands from the fused health index using stacked light weight heterogeneous ensemble base learners.
- iii) Phase 3: Impact assessment of the proposed approach of electric load forecasting.

The organization of the work is as follows: Section 2 entails a brief literature review on prominent machine learning and deep learning-based works in electricity load forecasting. The proposed methodology described in Section 3 and Section 4 elaborates the experimental setup, hey results of the model and discussions on the results. Section 5 highlights the future research extensions apart from providing the concluding remarks.

2. Literature Survey

Electricity load forecasting long running issue in industrial, commercial and domestic sectors. Both long range and short-range forecasting are queen essential for proper planning of electricity distribution as well as for developing strategies and policies for tapping renewable alternate energy resources.

Solyal (2020) employed ANN, multiple linear regression, adaptive neuro-fuzzy inference system along with SVM to forecast electricity demand in Cyprus. The input parameters that are considered in the study includes temperature, solar, humidity, capita income, population and electricity consumption. The experimental analysis indicates that SVM and ANN able to accurate long term as well as short term predictions. The method of Bayesian Clustering by Dynamics (BCD) along with support vector regression is used for a day ahead forecasting of the electricity (Fan et al., 2008). The BCD classifier create multiple subsets off the time series data which is then handled by regression model to give accurate predictions.

A holistic review on application of DL algorithms in electricity forecasting is done by Yildiz et al. (2017). This work explodes different applications of DL models load forecasting systems, apart from comparing the performance of the DL models. Bouktif et al. (2018) used Long Short-Term Memory (LSTM) model for aggregate load forecasting. The optimal features are chosen using embedded feature selection and wrapper methods which are further fine-tuned by genetic algorithm. Ask the model chooses only the best

optimal parameters it exhibited much lower prediction errors. A hybrid method that integrates entropy-based feature selection with random forest, ANN and fuzzy inductive reasoning is proposed by Jurado et al. (2015). The proposed method was much faster and more reliable than conventional regressive integrated moving average model.

In addition to SVM, RF and LSTM, nonlinear auto regressive exogenous neural network yes used in medium term electricity forecasting by Shirzadi et al. (2021). The performance of this model was assessed by using load demand data from Ontario region in Canada. Model exhibited better efficacy in commercial and domestic power forecasting. Jawad et al. (2020) proposed a least cost electricity load forecasting model what reduced prediction errors. The model focused on learning from weather parameters from Pakistan for making accurate predictions.

Andriopoulos et al. (2020) proposed a Convolution Neural Network (CNN) that explores the statistical properties of the time series electrical load data. This method performs better than the LSTM open examined with publicly available data sets. Seven different autoregression models were employed for electricity load forecasting, whose results are validated using adaptive-network-based fuzzy inference system by Yousaf et al. (2021). The feature selection was made using principal component analysis and the real time testing of the model is done on the electricity load forecasting data obtained from Pakistan. A hybrid parallel LSTM-CNN Network is created by Farsi et al. (2021) for predicting the hourly electricity load in Malaysia (Farsi et al., 2021). The performance of this model is validated on data acquired from two countries namely Germany Malaysia.

Development of ensemble models combines the predictive power off multiple base learners. This is explored by Khwaja et al. (2020) to create ensemble-based ensemble models for electricity load forecasting. This work integrates boosting and bagging methods to train the multiple ANN models to predict the final forecasted load. The results of the proposed indicates remarkable decline in the bias and variance hey than individual ANN model.

A novel hybrid methodology which integrates bidirectional LSTM and CNN electric load is proposed by Sekhar & Dahiya (2023). This method deploys grey wolf optimization to select optimal set of parameters for the hybrid model. The model was able to give better efficacy in hourly load prediction. A detailed review on the data driven electric load forecasting methods is done by Manandhar et al. (2023). The review proceeds by explaining different stages in electric load forecasting the techniques deployed for the prediction process. This book highlights the salient features, advantages and limitations of each work which forms a basis for future research. A Training Data Generator (TDG) along with Look-Back Optimiser (LBO) were deployed by Wang et al. (2023) for enhancing the performance of base ML models (Wang et al., 2023). The TDG select subsidy meaning data which is further optimised by LBO, which considers the accuracy of previous forecast to adjust the framework for accurate future predictions. This this method was found to be efficient in handling imbalanced classes and the comprehensive evaluation shows that the model was much efficient than the other state of art techniques. Different machine learning and deep learning algorithms has been used in papers, to get clarity, this reviewed literature has been systematically compiled and summarized in **Table 2** to provide a clear comparison of methodologies, performance metrics and key findings and limitations of the papers.

Table 2. Characteristics of various load forecasting.

Author(s) & Year	Title	Methodology	Key findings	Limitations
Andriopoulos et al. (2020)	Short term electric load forecasting based on data transformation and statistical ML	Statistical ML + Data Transformation	Improved short-term load forecast accuracy	Focused on short-term only
Bouktif et al. (2018)	Optimal LSTM using GA & Feature Selection	LSTM + Genetic Algorithm	Outperforms traditional ML for load forecasting	Computationally expensive
Farsi et al. (2021)	Parallel LSTM-CNN model	Deep Learning (LSTM-CNN)	Accurate short-term forecasting	Training complexity
Jawad et al. (2020)	Cost effective electricity load forecasting using meteorological parameters	ML + Weather Features	Low-cost, accurate predictions	Dependent on weather data
Jurado et al. (2015)	Hybrid methodologies for electricity load forecasting	Entropy-based Feature Selection + Soft Computing	Improved accuracy with hybrid models	Computational cost
Khwaja et al. (2020)	Joint bagged-boosted ANN	Ensemble ML (Bagging + Boosting)	Improved accuracy with ensemble ANN	Model interpretability
Manandhar et al. (2023)	Urban energy modeling review	Review of ML in Urban Energy	Highlights current challenges	Generalized overview
Sekhar & Dahiya (2023)	Robust framework based on hybrid deep learning approach for short term load forecasting of building electricity demand	Hybrid Deep Learning	Robust for building-level load	Needs fine-tuning
Shirzadi et al. (2021)	Medium-term regional forecasting	ML + DL Models	Effective for regional planning	Data demand
Wang et al. (2023)	Real-time load forecast + anomaly detection	ML + Anomaly Detection	Dual-function system	Real-time processing needs
Yildiz et al. (2017)	Regression & ML models in buildings	Regression + ML Review	Analyzes building energy models	Commercial building focus
Yousaf et al. (2021)	Improved residential load forecasting	ML + Feature Selection + Integration Strategy	Highly accurate residential predictions	Model integration complexity

Some of the research gaps from the intensive literature survey are as follows: 1) Most of the models are tailored for specific applications- residential, commercial, regional or building level- and fail to generalize well across different contexts. 2) Deep learning and hybrid models offer high accuracy but are often computationally expensive and making real-time deployment impractical. 3) Only few models address real-time forecasting needs. However, they lack scalability and have high latency. 4) Many feature selection studies used manual or heuristic-based approach which are not automated or adaptive to changing conditions. 5) Added significant design and implementation complexity in the case of multi-stage framework models. Though the methods discussed in the survey are very effective in predicting the electric load, because of computational complexity and memory constraints, there is still room for improvement.

Electric load forecasting presents a multifaceted challenge due to the inherent variability and nonlinearity in consumption patterns, influenced by various temporal, behavioural and environmental factors. Traditional statistical or single stage machine learning models, struggle to capture the full spectrum of underlying patterns and also fail to generalise well across diverse scenarios. Though the deep learning models are powerful, can be computationally intensive and lacks interpretability. To address these gaps and limitations, this work proposes a stacking ensemble regression framework that strategically combines the predictive capabilities of multiple diverse base learners, followed by a meta learner for final prediction. This ensemble approach is motivated by the need to leverage unique strength of each algorithm to reduce both bias and variance, enhance predictive accuracy and improve robustness against data noise and feature interactions. The modular and scalable nature of stacking enables seamless integration of additional

learners, making the model more adaptable to evolving forecasting needs and real-world deployment scenarios.

3. Lightweight Ensemble Model for Electric Load Forecasting

The electric load consumption data is highly volatile as it depends on many factors, temperature, humidity, season, time of consumption, type of consumers. Though, the traditional ensemble models achieve high accuracy, the complexity and resource requirements are high and less suitable for real-time or resource constrained environments. Thus, to balance accuracy, computational efficiency and practical applicability light weight ensembled model is proposed in this paper. The model development proceeds in three phases namely 1) Data preprocessing and influencing feature identification 2) Feature fusion into fused health index and 3) Development of light weight ensemble model.

3.1 Data Preprocessing and Influencing Feature Identification

Accurate electric load forecasting critically depends on the quality of dataset and relevance of input features. In this work, a set of exogenous variables are collected from Panama electricity load forecasting dataset that reflect environmental and meteorological conditions known to influence electricity demand. Data preprocessing is a step to improve the quality of data, by means of handling missing values and outlier handling. Missing or corrupted values are handled using exponentially weighted moving average interpolation as this ensures temporal consistency while smoothing short-term noise in the data. The Exponentially Weighted Moving Average (EWMA) is widely used in applications where volatility is a primary concern such as electric load. In this method the older observations are assigned lower priority or lower weights which the recent ones hold high weights. The decay of weights happens exponentially, hence the name exponentially weighted moving average. The hyperparameter (α) determined the significance of the observation. This is a computationally efficient way to track time- series data such as electric load. The estimation of EWMA is done according to Equation (1).

$$EWMA_t = \alpha \cdot x_t + (1 - \alpha) \cdot EWMA_{t-1} \quad (1)$$

The input data at current timestamp is indicated as x_t . It is evident that EWMA is a recursive function, where current value relies on the value from previous timestamp. The value of α should lie within $[0,1]$ and the observation is more significant for larger value of α . After estimating the EWMA of each observation at every timestamp, the data preprocessing is done by trimmed mean fusion, which is a flavour of measures of central tendency. This works by curbing the lowest and highest values and then estimating the mean of remaining values. This data preprocessing technique handles the outliers in the data.

To access the influence of each exogenous variables on electric demand univariate Pearson correlation coefficients and feature importance scores are calculated. Calculated Univariate Pearson coefficients between each feature and electricity load demand and feature importance scores from an initial run of tree based extreme gradient boosting model. The features that consistently showed low influence in both methods are removed.

3.2 Feature Fusion into Fused health index (FHI)

Fused health index (FHI) is a composite statistical metric used in condition monitoring and predictive maintenance. It integrates multiple individual condition indicators into a single value that reflects the overall health of the system or component. In this work, FHI is constructed from multiple dependant variables such as temperature, humidity, wind speed and precipitation. Initially calculate $EWMA_i(t)$ for all individual exogenous variable $x_i(t)$ using Equation (1). Later, calculate fused health index at time t as the trimmed mean of smoothed exogenous variables using Equation (2).

$$FHI(t) = \text{TrimmedMean}(\varepsilon(t), p) \quad (2)$$

where, $\varepsilon(t)$ is a vector of EWMA of all exogenous variable, $\{EWMA_1(t), EWMA_2(t), \dots, EWMA_N(t)\}$ and p is the percentage values trimmed from both ends.

3.3 Development of Proposed Model

The key idea behind the development of lightweight ensemble model which combines the promise of multiple base ML models is to achieve better computational efficiency reduced cost. The method of ensemble here comprises of amalgamating diverse models for getting a single prediction value. There are two types of ensembles namely homogeneous and heterogeneous. The former method proceeds by constructing ensemble of same ML model while the later one explores the usage of various models. The models that are used for creating the ensembles are commonly known as base models or weak learners. By employing independent and different type of base learners increases the chance of reducing the generalisation error. It is evident that the ensemble models use multiple ML algorithms each of them with their own advantages and disadvantages. However, the ensemble model will give one single final outcome. These models can be used for both regression as well as classification tasks (Garai & Paul, 2023). The classification ensemble models give their results by applying a majority voting rule. On the other hand, regression ensemble models use various aggregation techniques to give the final result. Alternatively, a meta learner can be used as an aggregator of the results of base learners in both classification as well as regression type of ensemble models. This meta learner is yet another ML model, whose choice has a great influence on the final prediction (Hospedales et al., 2022).

The proposed work focuses on constructing a light weight ensemble model with five different base learners. (1) Linear Regression (LR) – captures linear relationship between independent and dependent variable, fast and interpretable. (2) Classification and Decision Trees (CART) which captures non-linear interactions and variable splits, also interpretable and robust to outliers. (3) Support Vector Machines with Radial Basis Function kernel (SVM-RBF): effective for high dimensional spaces and handles non-linear separability using kernel trick. (4) Light Gradient Boosting Machine (LGBM) which is fast, leaf-wise tree growth model that performs well on structured tabular data with high dimensionality. (5) Extreme Gradient Boosting (XGB) regressor which predicts based on regularized boosting technique. The predictions of these base learner models are subjected to high accuracy, built in regularization machine learning model XGB regressor, which has ability to model complex residual patterns, acts as meta learner in this work. The proposed work uses stacking paradigm to converge the predictions of multiple individual base learners. The XGB meta learner aggregates the individual predictions of the base learners to yield the final forecast of the electricity demand. **Figure 4** shows the architecture diagram of the proposed electricity load forecasting model. Feature fusion identifies most influencing parameters from the raw dataset and ignores the least significant attributes. Each base learner model learns from fused dataset and captures different aspects of relationship between features and target variable. Different base models contribute diverse predictions and best combinations of these outputs acts as input to the meta learner model. From the diversified output from base model as input, meta learner model captures relationship between depending variable and target variable and using relation makes final prediction.

As the work focuses on learning the load patterns and trends by analysing various exogeneous variables, deploying a single ML model would be inefficient in forecasting the electric load. Hence, the proposed work builds an ensemble model whose individual base learners probe the data in a unique and distinct way to uncover new patterns and hidden trends in the data space. Each of the base learner has strong learning capability and are computationally less expensive than devising a powerful layered model.

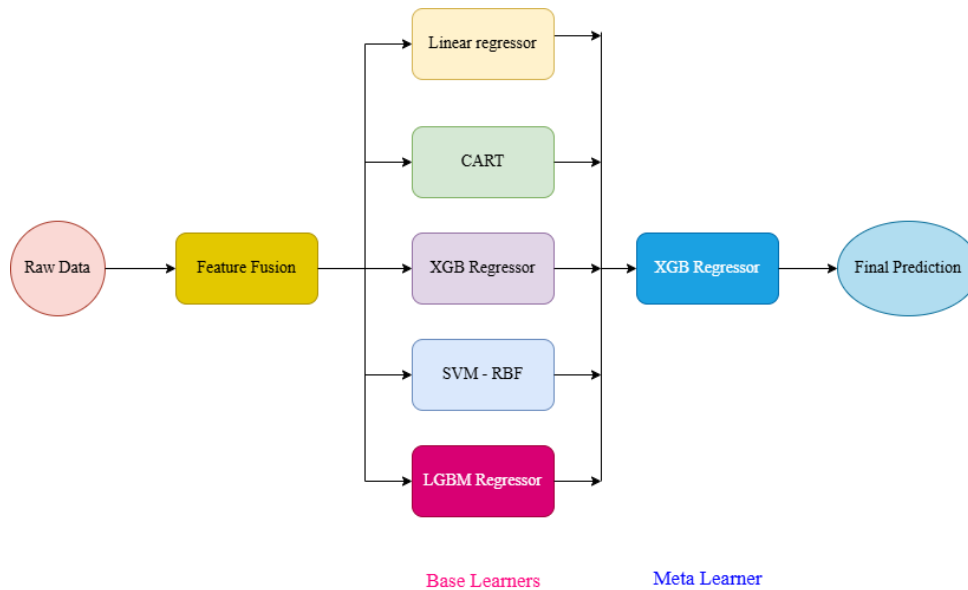


Figure 4. Proposed Light weight ensemble model with meta learner.

3.3.1 Multivariate Linear Regression

Linear regression (LR) predicts results by investigating the independent variables and creates a regression model for dependent variable. The proposed architecture uses multivariate LR, which works on multiple independent variables to give the prediction. The mathematical expression for the multivariate LR is given as Equation (3).

$$\text{Prediction} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (3)$$

The terms bias (β_0) and weights ($X_1, X_2, \dots, X_n$) are learnt by the ML model by analysing the dependent and independent exogeneous variables in electric load forecasting.

3.3.2 Support Vector Machines

The objective of SVM regressor is to constitute an imaginary hyperplane to partition the dimensional space of the data. The knowledge acquired by the algorithm is harnessed to choose the optimal support vectors that best fit the data. SVM supports many types of kernels each with its own pros and cons. The proposed methodology uses RBF kernel where only the nearest support vectors are used for further processing, which makes it light weight. Equation (2) signifies the drawing of the hyper plane with x as input and was the weight. The equation of the hyperplane of the Support Vector Regressor (SVR) is given by Equation (4), which is similar to equation of a straight line (Y). The weights are given as ω and input is mentioned as x .

$$Y = \omega x + b \quad (4)$$

The hyper plane is drawn according to Equations (5) and (6). Here d indicates the distance.

$$\omega x + b = +d \quad (5)$$

$$\omega x + b = -d \quad (6)$$

The hyperplane will be drawn based on the following criteria: $-d < Y - (\omega x + b) < +d$.

3.3.3 Classification and Regression Trees

CART forecast the electric load by bifurcating into decision trees where each node indicates the point of decision. The outcome of the decision is indicated as branches. The growth of the tree is controlled by tree-splitting criteria, which is a hyper parameter (Bittencourt & Clarke, 2003). The tree extends greedily by selecting the best possible split. This work focuses on splitting the tree based on Sum of Squared Error (SSE), which is estimated according to Equation (7).

$$SSE = \sum_{i \in s_1} (y_i - \widehat{y_1})^2 - \sum_{i \in s_2} (y_i - \widehat{y_2})^2 \quad (7)$$

The bifurcation of the dataset is done into two subsets namely s_1 and s_2 . The choice of a record belonging to a particular subset will be governed according to its least value of SSE. The term $\widehat{y_1}$ and $\widehat{y_2}$ are mean of the samples in the set s_1 and s_2 respectively. This method follows greedy recursive partitioning to spawn the tree to output the final prediction.

3.3.4 XGB Regressor

The power of XGB lies in its objective function which imbibes an inherent regularization term. The XGB is a homogeneous ensemble weak decision tree. Equation (8) determine the predicted value of the XGB regressor. The term α_n and R_{n-1} indicates regularization parameter and residual computed from the n^{th} tree respectively. The residual value of the input exogeneous variables (X) is denoted using the term h_n . Equation (9) portrays the differentiable loss function (f_i) (Velthoen et al., 2023).

$$f_n(X) = f_{n-1}(X) + \alpha_n h_n(X, R_{n-1}) \quad (8)$$

$$argmin(\alpha) = \sum_{i=1}^n L(Y_i, f_{i-1}(X_i) + \alpha h_i(X_i, R_{i-1})) \quad (9)$$

3.3.5 Light Gradient Boosting Machines

LGB machines are special kind of ML which are specifically designed for faster convergence at low resource consumption. Like XGB, this is also a Decision tree-based model, where the bifurcation of the tree is done at the leaves, resulting in more deeper trees (Fan et al., 2019).

Thus, the proposed model aggregated individual base learners with a meta-learner. This work creates five distinct ensemble models with different base learners and meta learner to forecast the electricity load.

3.3.6 Training and Testing Phase of the Model

The proposed ensemble model exhibits two different behaviours during the training and testing phase which is depicted in **Figure 5**.

The operation of the model during training phase proceeds by learning the fused features and testing of the model's accuracy is done. If the prediction accuracy is not satisfactory then the model adjusts or modifies its assumptions about various hyper parameters. This is an iterative process which converges when the model outputs satisfiable results with great accuracy. This trained model is then tested with real time data, which yields the electric load as output.

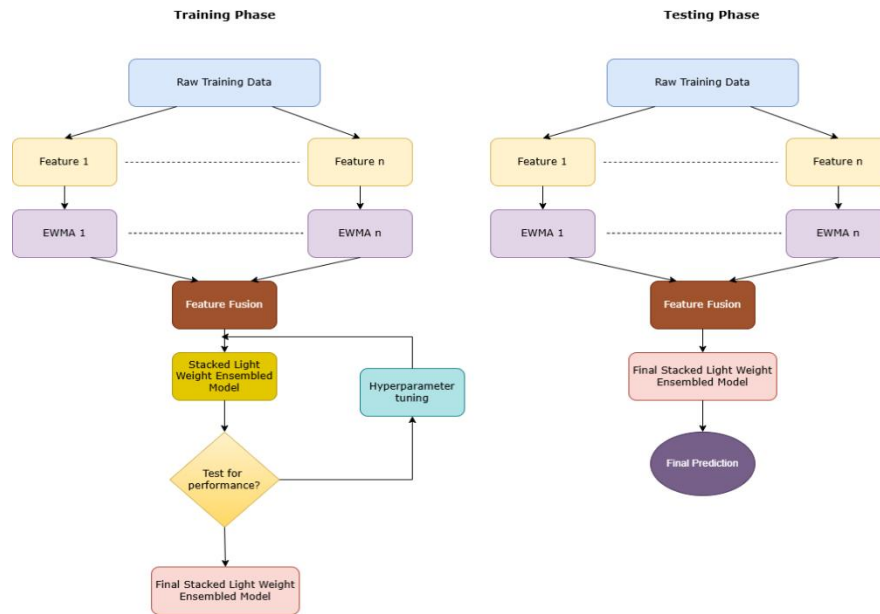


Figure 5. Testing and training behavior of the model.

4. Experimental Set Up and Data Analysis

This section explains the dataset and provides detailed analysis about the data to understand the efficacy of the proposed model in predicting electric load. The work validates the model's performance using the Panama national electricity load data (Huang et al., 2023) and is publicly accessible through electric load forecasting research repositories. This data is obtained at one-hour intervals with a vast data comprising of 17,520 records. The data is measured from 0:00 on January 1st, 2018, till 23:00 on December 31st, 2019 among three areas namely Tocumen, Santiago and David belonging to the Panama City. The dataset includes both target variables and exogenous variables: temperature, relative humidity, wind speed and liquid precipitation from each of the three areas. **Table 3** enumerates the list of features considered for electricity load prediction.

Table 3. Features and their units in the dataset.

Feature with description	Unit	Feature with description	Unit
nat_demand -Electricity load	MWh	TQL_san- Liquid precipitation in Santiago of Panama City	l/m2
T2M_toc -Temperature at a distance of 2 m in Tocumen of Panama City	°C	W2M_san - Speed of wind at a distance of 2 m in Santiago of Panama City	m/s
QV2M_toc -relative humidity at a distance of 2 m in Tocumen of Panama City	%	T2M_dav Temperature at a distance of 2 m in David of Panama City	°C
TQL_toc -Liquid precipitation in Tocumen of Panama City	liters/m2	QV2M_dav - relative humidity at a distance of 2 m in David of Panama City	%
W2M_toc- Speed of wind at a distance of 2 m in Tocumen of Panama City	m/s	TQL_dav Liquid precipitation in David of Panama City	l/m2
T2M_san- Temperature at a distance of 2 m in Santiago of Panama City	°C	Speed of wind at a distance of 2 m in David of Panama City	m/s
QV2M_san - relative humidity at a distance of 2 m in Santiago of Panama City	%		

The Exploratory Data Analysis (EDA) of the above dataset is shown in **Figure 6**. The X axis of the fig indicates the feature values and Y axis is the count of values. This helps us to understand the distribution of data.

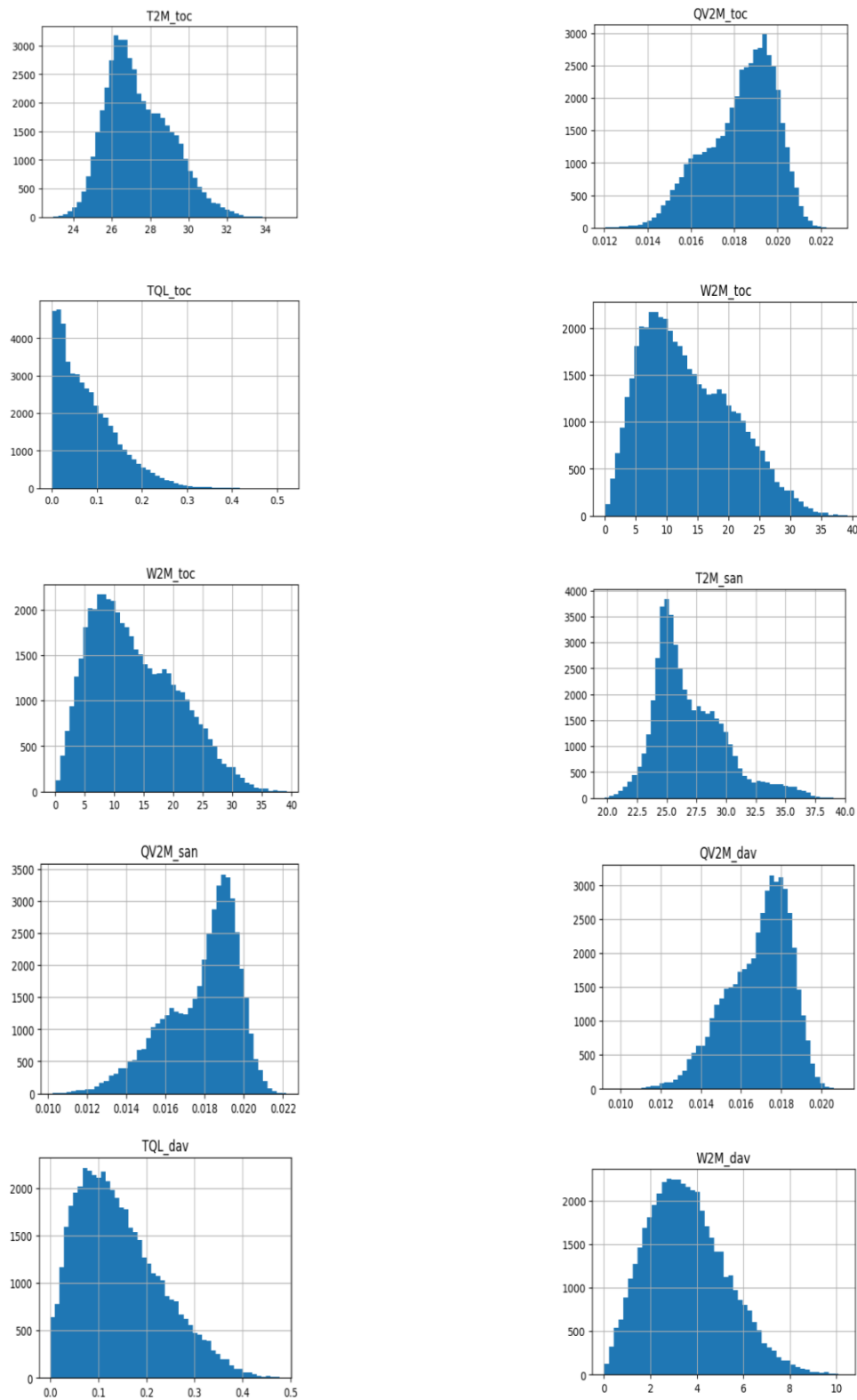


Figure 6. EDA of the dataset using histograms.

The data roughly follows normal distribution except the feature TQL_toc. As there is good variation in the data, the base learners of the data can explore the data space in more robust manner. Also, the various hyperparameters of the base learners are listed in **Table 4**.

Table 4. Hyperparameter setting for the model.

Model	Hyperparameter	Value
LGM machine	No. of leaves	6
	Learning rate	0.01
	No. of estimators	100
	Maximum bin size	10
	Bagging fraction	0.8
	Bagging frequency	4
	Seed	8
	Feature fraction	0.2
XGB regressor	Maximum depth	8
	Number of estimators	500
	Minimum child weight	1000
SVM	Kernel	RBF

5. Results and Discussion

Feature fusion by trimmed mean of the EWMA of the raw electricity load prediction data is performed and the fused value is used for predicting the national load demand. 80% of the data is used for training the model while the remaining 20% is used for testing performance of the model. The location of the data collection does not hold any significance in this study. The prediction efficacy of the model for a sample of test data is presented in **Figure 7** and **Figure 8**.

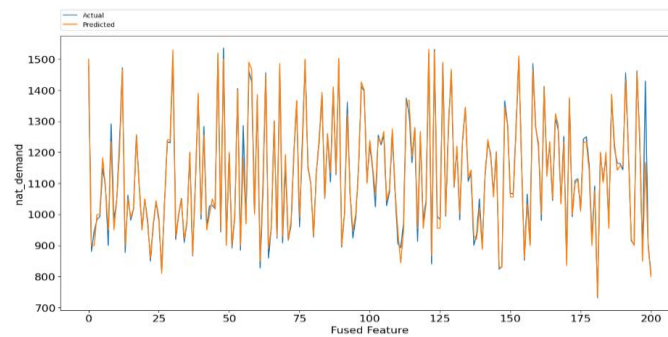


Figure 7. Predicted and actual national electricity load demand of all test data points.

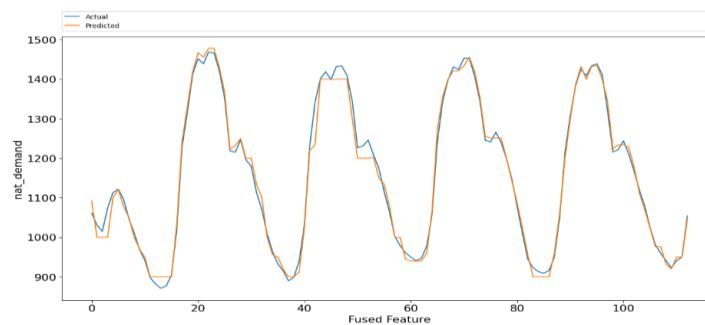


Figure 8. Predicted and actual national electricity load demand of sample data points.

At most of the time the actual and predicted data is in alignment with each other. Also, a very feeble deviation can be sensed in some of the data points in the model. The **Figure 7** shows much of randomness in the data points while the data points in **Figure 8** are more periodic. However, their periodicity and non-periodicity does not hold any relevance to the performance of the proposed model. The data sampling is done to have a closer look on the deviation between the predicted and actual values. However, quantifying the efficacy of the model is very important to draw scientific inferences from the study. Hence, the performance of the proposed model is validated on few of the prediction performance metrics as described below:

5.1 Mean Bias Error (MBE)

This is a measure of the average difference between the testing and training data. Ideally, the best model will achieve an MBE value of zero. Both the positive and negative MBE are said to attain under estimation and overestimation of the values respectively. This metric is best described by the Equation (10).

$$MBE = \sum_{k=1}^k (\log L_p - \log L_a) \quad (10)$$

Here, the L_p signifies the predicted load while L_a indicates actual electricity load demand.

5.2 Root Mean Squared Error (RMSE)

RMSE is the most predominantly used metric in regression-based analysis. This is a direct indication of the normalised residuals for every data point. Equation (11) portrays the expression for computing the RMSE and k data points.

$$RMSE = \sqrt{\frac{\sum_{k=1}^K \|L_p(k) - L_a(k)\|^2}{K}} \quad (11)$$

5.3 R^2 Value

This metric is used in almost all regression-based problems to find the variance ratio of independent variables. This coefficient of determination estimates the square of the correlation. Zero value of R^2 signifies that dependent variable is not predictable from independent variables considered in the study. Hence, increased values of R^2 indicates a positive trend which indicates the right choice of independent variables for model building. The mathematical expression for computing the R^2 values is given in Equation (12).

$$R^2 = 1 - \frac{\sum_{k=1}^K (L_{average} - L_p)^2}{\sum_{k=1}^K (L - L_a)^2} \quad (12)$$

5.4 Mean Absolute Error

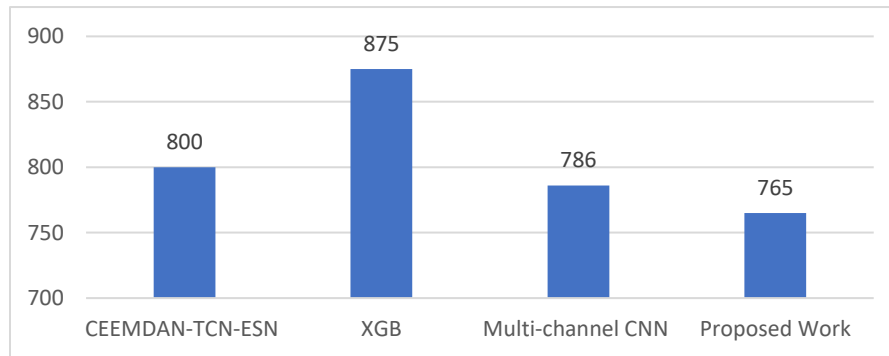
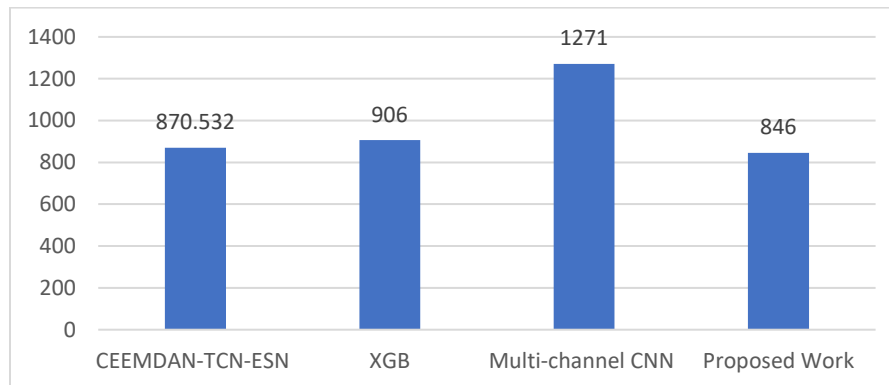
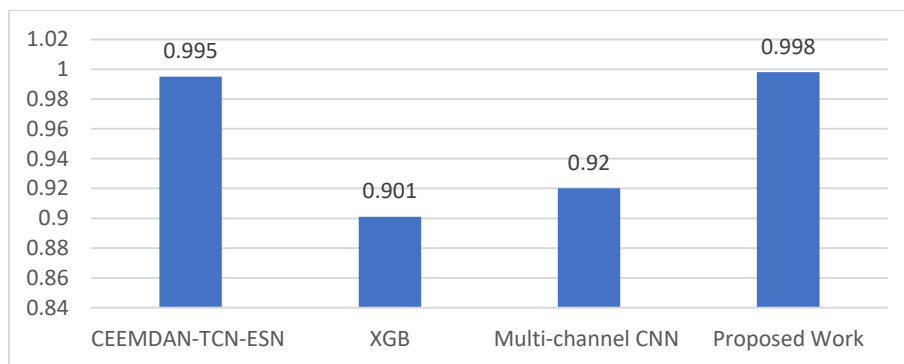
This is an important metric that defines the mean variance among the significant data point and the predicted values. It is an estimate of the mean of mispredictions irrespective of the direction. Equation (13) displays the mathematical expression:

$$MAE = \frac{\sum_{k=1}^K |L_a - L_p|}{K} \quad (13)$$

The proposed methodology is compared with three other states of the art techniques such as CEEMDAN-TCN-ESN model (Huang et al., 2023), XGB model (Madrid & Antonio, 2021) and multi-channel CNN (Ibrahim et al., 2023). **Table 5** summarises the results of the four models using four regression metrics. The graphical comparative analysis is given in **Figure 9** to **Figure 12**.

Table 5. Comparison of various performance metrics.

Method	MBE	RMSE	R ²	MAE
CEEMDAN-TCN-ESN	800	870.532	0.995	595.843
XGB	875	906	0.901	653
Multi-channel CNN	786	1271	0.92	678
Proposed Work	765	846	0.998	576

**Figure 9.** Performance comparison of state of art techniques with proposed work based on MBE.**Figure 10.** Performance comparison of state of art techniques with proposed work based on RMSE.**Figure 11.** Performance comparison of state of art techniques with proposed work on R2_score.

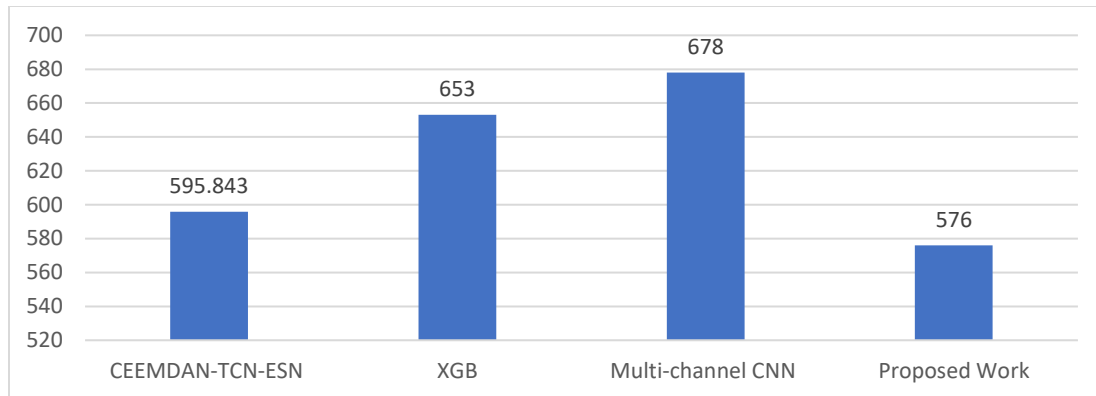


Figure 12. Performance comparison of state of art techniques with proposed work based on MAE.

From **Table 5** and **Figure 9** to **Figure 12**, the proposed model demonstrates superior performance across all metrics. It records minimum MBE (765) and RMSE (846), indicating minimal prediction bias and reduced large error occurrences. The model also achieves the highest R^2 value of 0.998 which suggest 99.8% of model predictions are close to actual data, thus reflecting excellent generalization capability. In addition, the MAE (576) is also lowest compared to other models, which further affirming the model's precision in average prediction. Compared to CEEMDAN-TCN-ESN hybrid model, XGB and CNN- based architectures, the proposed model exhibits enhanced accuracy and robustness. These improvements can be attributed to the diversity of base learners and ability of meta learner to optimally combine their strengths leading to more reliable and consistent forecasts.

6. Conclusions and Future Research

The proposed work presents a novel model that predicts the electric load by exploring the exogeneous variables. Trimmed mean of EMWA is used as a feature fusion technique that act as an integral index for load prediction by the model. The model was developed by ensemble robust light weight base learners, whose predictions are acted upon by the meta learner. XGB algorithm is used as meta learner in this work. The performance of the proposed methodology is validated using Panama electric load forecasting method which is a comprehensive dataset that comprises of electric load data from three regions. Several important research implications can be drawn from the results of the model. The model exhibited superior performance than the compared state of art models in terms of MBE, MAE, RMSE and R^2 error. As future research extension, the deployment of Deep Learning models can be explored. Especially, Recurrent Neural Networks will act as excellent choice for time series forecasting.

Conflict of Interest

The authors confirm that there is no conflict of interest to declare for this publication.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors would like to thank the editor and anonymous reviewers for their comments that help improve the quality of this work.

AI Disclosure

The author(s) declare that no assistance is taken from generative AI to write this article.

References

- Almashaiei, E., & Soltan, H. (2011). A methodology for electric power load forecasting. *Alexandria Engineering Journal*, 50(2), 137-144. <https://doi.org/10.1016/j.aej.2011.01.015>.
- Andriopoulos, N., Magklaras, A., Birbas, A., Papalexopoulos, A., Valouxis, C., Daskalaki, S., Birbas, M., Housos, E., & Papaioannou, G.P. (2020). Short term electric load forecasting based on data transformation and statistical machine learning. *Applied Sciences*, 11(1), 158. <https://doi.org/10.3390/app11010158>.
- Bittencourt, H.R., & Clarke, R.T. (2003). Use of classification and regression trees (CART) to classify remotely-sensed digital images. In *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No. 03CH37477)* (Vol. 6, pp. 3751-3753). IEEE. Toulouse, France. <https://doi.org/10.1109/igarss.2003.1295258>.
- Bouktif, S., Fiaz, A., Ouni, A., & Serhani, M.A. (2018). Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: comparison with machine learning approaches. *Energies*, 11(7), 1636. <https://doi.org/10.3390/en11071636>.
- Fan, J., Ma, X., Wu, L., Zhang, F., Yu, X., & Zeng, W. (2019). Light gradient boosting machine: an efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data. *Agricultural Water Management*, 225, 105758. <https://doi.org/10.1016/j.agwat.2019.105758>.
- Fan, S., Chen, L., & Lee, W.J. (2008). Machine learning based switching model for electricity load forecasting. *Energy Conversion and Management*, 49(6), 1331-1344. <https://doi.org/10.1016/J.enconman.2008.01.008>.
- Farsi, B., Amayri, M., Bouguila, N., & Eicker, U. (2021). On short-term load forecasting using machine learning techniques and a novel parallel deep LSTM-CNN approach. *IEEE Access*, 9, 31191-31212. <https://doi.org/10.1109/access.2021.3060290>.
- Garai, S., & Paul, R.K. (2023). Development of MCS based-ensemble models using CEEMDAN decomposition and machine intelligence. *Intelligent Systems with Applications*, 18, 200202. <https://doi.org/10.1016/J.ISWA.2023.200202>.
- Global electricity consumption 2023| Statista. (n.d.). Retrieved September 5, 2025, from <https://www.statista.com/statistics/280704/world-power-consumption/>.
- Hospedales, T., Antoniou, A., Micaelli, P., & Storkey, A. (2022). Meta-learning in neural networks: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(9), 5149-5169. <https://doi.org/10.1109/tpami.2021.3079209>.
- Huang, J., Zhang, X., & Jiang, X. (2023). Short-term power load forecasting based on the CEEMDAN-TCN-ESN model. *Plos One*, 18(10), e0284604. <https://doi.org/10.1371/journal.pone.0284604>.
- Ibrahim, B., Rabelo, L., Sarmiento, A.T., & Franco, E.G. (2023). A holistic approach to power systems using innovative machine learning and system dynamics. *Energies*, 16(13), 5225. <https://doi.org/10.3390/en16135225>.
- Jawad, M., Nadeem, M.S.A., Shim, S.O., Khan, I.R., Shaheen, A., Habib, N., Hussain, L., & Aziz, W. (2020). Machine learning based cost effective electricity load forecasting model using correlated meteorological parameters. *IEEE Access*, 8, 146847-146864. <https://doi.org/10.1109/access.2020.3014086>.
- Jurado, S., Nebot, À., Mugica, F., & Avellana, N. (2015). Hybrid methodologies for electricity load forecasting: Entropy-based feature selection with machine learning and soft computing techniques. *Energy*, 86, 276-291. <https://doi.org/10.1016/j.energy.2015.04.039>.
- Khwaja, A.S., Anpalagan, A., Naeem, M., & Venkatesh, B. (2020). Joint bagged-boosted artificial neural networks: Using ensemble machine learning to improve short-term electricity load forecasting. *Electric Power Systems Research*, 179, 106080. <https://doi.org/10.1016/j.epsr.2019.106080>.
- Kuster, C., Rezgui, Y., & Mourshed, M. (2017). Electrical load forecasting models: a critical systematic review. *Sustainable Cities and Society*, 35, 257-270. <https://doi.org/10.1016/j.scs.2017.08.009>.

- Madrid, E.A., & Antonio, N. (2021). Short-term electricity load forecasting with machine learning. *Information*, 12(2), 50. <https://doi.org/10.3390/info12020050>.
- Manandhar, P., Rafiq, H., & Ubinas, E.R. (2023). Current status, challenges, and prospects of data-driven urban energy modeling: a review of machine learning methods. *Energy Reports*, 9, 2757-2776. <https://doi.org/10.1016/j.egyr.2023.01.094>.
- Pai, P.F., & Hong, W.C. (2005). Support vector machines with simulated annealing algorithms in electricity load forecasting. *Energy Conversion and Management*, 46(17), 2669-2688. <https://doi.org/10.1016/j.enconman.2005.02.004>.
- Sekhar, C., & Dahiya, R. (2023). Robust framework based on hybrid deep learning approach for short term load forecasting of building electricity demand. *Energy*, 268, 126660. <https://doi.org/10.1016/j.energy.2023.126660>.
- Shen, Z., Wu, X., Guerrero, J.M., & Song, Y. (2016). Model-independent approach for short-term electric load forecasting with guaranteed error convergence. *IET Control Theory & Applications*, 10(12), 1365-1373. <https://doi.org/10.1049/iet-cta.2015.0818>.
- Shirzadi, N., Nizami, A., Khazen, M., & Nik-Bakht, M. (2021). Medium-term regional electricity load forecasting through machine learning and deep learning. *Designs*, 5(2), 27. <https://doi.org/10.3390/designs5020027>.
- Solyali, D. (2020). A comparative analysis of machine learning approaches for short-/long-term electricity load forecasting in Cyprus. *Sustainability*, 12(9), 3612. <https://doi.org/10.3390/su12093612>.
- Velthoen, J., Dombry, C., Cai, J.J., & Engelke, S. (2023). Gradient boosting for extreme quantile regression. *Extremes*, 26(4), 639-667. <https://doi.org/10.1007/s10687-023-00473-x>.
- Verdejo, H., Awerkin, A., Becker, C., & Olguin, G. (2017). Statistic linear parametric techniques for residential electric energy demand forecasting: a review and an implementation to Chile. *Renewable and Sustainable Energy Reviews*, 74, 512-521. <https://doi.org/10.1016/j.rser.2017.01.110>.
- Wang, X., Yao, Z., & Papaefthymiou, M. (2023). A real-time electrical load forecasting and unsupervised anomaly detection framework. *Applied Energy*, 330, 120279. <https://doi.org/10.1016/j.apenergy.2022.120279>.
- Yildiz, B., Bilbao, J.I., & Sproul, A.B. (2017). A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *Renewable and Sustainable Energy Reviews*, 73, 1104-1122. <https://doi.org/10.1016/j.rser.2017.02.023>.
- Yousaf, A., Asif, R.M., Shakir, M., Rehman, A.U., & S. Adrees, M. (2021). An improved residential electricity load forecasting using a machine-learning-based feature selection approach and a proposed integration strategy. *Sustainability*, 13(11), 6199. <https://doi.org/10.3390/su13116199>.
- Zhang, J., Wei, Y.M., Li, D., Tan, Z., & Zhou, J. (2018). Short term electricity load forecasting using a hybrid model. *Energy*, 158, 774-781. <https://doi.org/10.1016/j.energy.2018.06.012>.



Original content of this work is copyright © Ram Arti Publishers. Uses under the Creative Commons Attribution 4.0 International (CC BY 4.0) license at <https://creativecommons.org/licenses/by/4.0/>

Publisher's Note- Ram Arti Publishers remains neutral regarding jurisdictional claims in published maps and institutional affiliations.