International Journal of Applied Power Engineering (IJAPE)
Vol. 14, No. 2, June 2025, pp. 264~274
ISSN: 2252-8792, DOI: 10.11591/ijape.v14.i2.pp264-274

Powering the future of electrical load forecasting using a regression learner in machine learning

Sushama D. Wankhade[1,2], Babasaheb R. Patil[3]
[1]Department of Electrical Engineering, Sardar Patel College of Engineering, Andheri, India
[2]Department of Instrumentation Engineering, Bharati Vidyapeeth College of Engineering, Navi Mumbai, India
[3]Vishwaniketan's Institute of Management Entrepreneurship and Engineering Technology, Kumbhivali, India

Article Info

Article history:

Received Feb 28, 2024
Revised Oct 10, 2024
Accepted Oct 23, 2024

Keywords:

Artificial neural network
Electrical load forecasting
Machine learning
Regression learner
Root mean square error
Support vector machine
Wide neural network

ABSTRACT

The primary intent of the present research was to design and execute an electrical load forecasting system using machine learning (ML) techniques. The implementation of an advanced predictive method, specifically an ML algorithm, helped in accurate load forecasting, which is crucial for efficient power grid management, and optimizing resource allocation. Electricity load fluctuates due to various complex factors, making traditional forecasting methods struggle. This is where ML shines. ML algorithms can learn from historical data, identifying intricate patterns and relationships that influence electricity demand. This allows them to make more accurate predictions than static models. In this work, regression learning models in ML are used with the MATLAB platform. Three years of real-time data from the Wavi substation in India are used. Considering day, date, hour of day, max and min temperature of the day, and voltage and current are taken as input parameters to test fourteen different models of assorted regression algorithms. The performance of these models is evaluated using commonly used metrics, root mean square error (RMSE), mean squared error (MSE), and mean absolute error (MAE), along with a few other parameters. The optimized trained model is then tested with real data to obtain the forecasted load. The correlation between the Actual load and forecasted load is found to be 0.999962.

Corresponding Author:

Sushama D. Wankhade
Department of Instrumentation Engineering, Bharati Vidyapeeth College of Engineering
CBD Belapur, Navi Mumbai, Maharashtra 400614, India
Email: sushamawankhade13@gmail.com

1. INTRODUCTION

Load forecasting is a crucial component of both the planning and operation of power systems. Accurate load prediction is crucial for determining the scheduling of power units, planning for capacity, improving the network, and managing demand from consumers [1]. Given the challenges associated with storing large amounts of electrical energy and the fluctuations in power demand, it is necessary to ensure that the power generation system is able to adapt to changes in load in real-time. Load forecasting is crucial in power infrastructure planning and grid operation. Precise load forecasting helps reduce the discrepancy between electricity supply and demand, thereby enhancing the stability of power systems.

The precision of the forecasting models is crucial in managing the emerging energy generation and consumption. Artificial intelligence (AI)-based methods are being researched and used in a wide range of applications around the world because they are better at handling complex input-output relationships. Machine learning (ML) is revolutionizing electrical load forecasting, paving the way for a more efficient, reliable, and sustainable future for the power grid [2], [3]. Traditionally, predicting electricity demand relied on statistical

models and historical trends. However, the rise of ML has revolutionized the field, offering more accurate and flexible forecasting.

Here's how ML is changing the game:

- Capturing complexity: Electricity demand is influenced by a myriad of factors, like weather, time of day, holidays, and even social events. ML algorithms can unravel these complex relationships and create models that adapt to dynamic conditions [3].
- Increased accuracy: ML models have the ability to acquire knowledge from extensive datasets, which encompass historical load patterns, weather forecasts, and real-time grid information. This leads to more precise predictions, reducing the risk of energy shortages or overproduction [4].
- Enhanced scalability: ML models can handle large datasets efficiently, making them ideal for forecasting at different levels, from individual buildings to entire power grids. This flexibility empowers tailored solutions for diverse scenarios [4].
- Proactive planning: Accurate forecasts enable efficient resource allocation, optimizing power generation and distribution. This translates to cost savings, reduced emissions, and improved grid reliability.
- Improved grid management: Predicting peak demand allows utilities to optimize generation and distribution, reducing costs and enhancing reliability.
- Renewable energy integration: ML can help integrate the variable output of renewable sources like solar and wind into the grid, maximizing their contribution.
- Demand-side management: Predicting peak demand allows utilities to optimize generation and distribution, reducing costs and enhancing reliability. By understanding future load, utilities can incentivize consumers to shift consumption patterns, smoothing demand peaks and reducing stress on the grid.

A diverse set of ML tools is at play, like regression models, which include models like random forests and support vector machines (SVM). These models capture non-linear relationships between multiple input features (weather, time of day) and the electricity load. Deep learning techniques like long short-term memory (LSTM) networks excel at handling temporal data, effectively capturing day-to-day and seasonal patterns in electricity consumption. Hybrid approaches, combining different ML algorithms can leverage their strengths, boosting overall forecasting accuracy.

Regression refers to a set of statistical methods used to analyze the relationship between a dependent variable and one or more independent variables. A regression model can determine if there is a relationship between changes in the dependent variable and changes in one or more of the explanatory variables. Regression methods are commonly employed in electrical load forecasting to predict future electricity usage accurately [5]. Various studies have highlighted the effectiveness of regression models in this domain. For instance, a study utilized linear regression equations to forecast electricity loads, achieving an average forecasting error of 3.86% for active power and 3.77% for apparent power [6]. Additionally, another research paper evaluated 24 regression model-based algorithms for half-hourly load forecasting, with Gaussian process regression models demonstrating the best performance [7]. Furthermore, a meta-regression analysis identified the LSTM approach and neural networks combined with other methods as effective forecasting techniques, emphasizing the importance of model selection in load forecasting [8]. These findings underscore the significance of regression methods in accurately predicting electrical loads, aiding in efficient energy management and resource allocation. Utilized artificial intelligence, neural network, ARIMA models, Bayesian models, and regression models for forecasting and proposed a solution to the problem of selecting three parameters for the support vector regression (SVR) model using a chaotic algorithm to enhance global optimization and prevent falling into local optimization [9], [10]. Gaussian process regression method is recommended for load prediction [11].

This paper focuses on regression learners for electrical load prediction using ML, utilizing real time hourly data from January 2019 and July, 2022 from the 33/11 kV substation at Wavi, India for analysis to compare 14 regression models like linear regression, SVM, and neural networks. The main contributions of the paper include: i) propose load forecasting approach for Wavi substation and ii) demonstrate probabilistic forecasting models. Regression models' performance is evaluated using root mean square error (RMSE), mean squared error (MSE), and mean absolute error (MAE) metrics. A conclusion is drawn by identifying intricate patterns and relationships affecting electricity demand for accurate predictions. The optimized model shows a high correlation between actual and forecasted load.

2. METHODOLOGY

The experimentation is done with regression learners using ML on MATLAB platform. Regression learner is a MATLAB tool that can be used to train different regression models with supervised ML. Initially, the real-time data is acquired from the substation, and, arranged on a daily basis for twenty-four hours. The features selected for data arrangement are date, day of the week, hour of the week, and maximum and minimum

temperature. Data is then tested for different models. In this experiment total of fourteen models are tested. After validating the scheme and parameter optimization of hyperparameters model performance is assessed again. In all fourteen models from five regression families Table 1 are tested here and the results obtained are tabulated as shown in Table 2. The optimized model is then tested for forecasting of the load. The different models used for regression are discussed below. The complete process flow of the work is as shown in Figure 1.

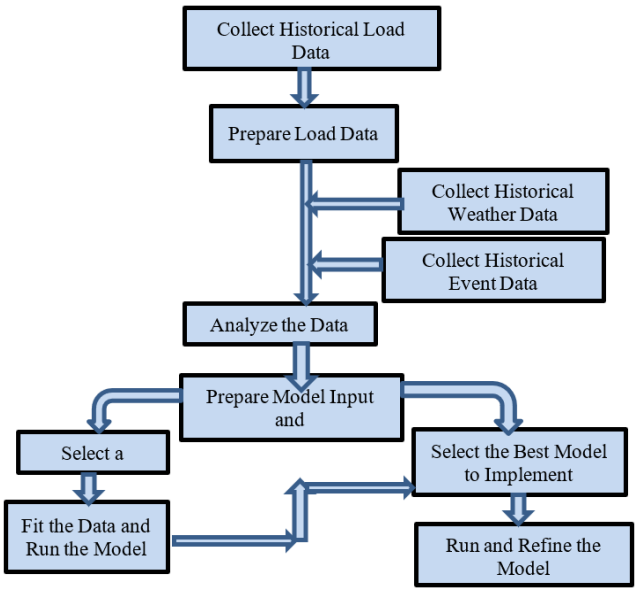

Figure 1. Process flowchart for forecasting

Table 1. Regression models used in the work

Family of regression models	Selected regression model
Linear regression	Linear regression model
Regression trees	Medium tree
	Coarse tree
	Fine tree
Support vector machines	Linear SVM
	Quadratic SVM
	Cubic SVM
	Medium Gaussian SVM
	Coarse Gaussian SVM
Ensemble of trees	Boosted trees
	Bagged trees
Neural networks	Narrow neural network
	Medium neural network
	Wide neural network

2.1. Linear regression model

A linear regression model is a statistical model that elucidates the connection between a dependent variable and one or more independent variables. The dependent variable is alternatively referred to as the response variable. A linear model example is a verbal scenario that can be modelled using a linear equation or vice versa [12], [13]. Linear regression is employed to precisely ascertain the nature and magnitude of the relationship between a dependent variable and a set of independent variables. It facilitates the generation of models for the purpose of making predictions [14].

2.2. SVM regression

SVM regression, also known as SVR, is an ML algorithm utilized for regression analysis. Unlike traditional linear regression methods, this approach seeks to identify a hyperplane that optimally aligns with the data points in a continuous space, rather than fitting a line to the data points. SVMs can use different kernel functions to transform the data into a higher-dimensional space, allowing for non-linear decision boundaries [15]. The accuracy of the test results with the SVM method is better than the linear regression

method. The Kernel trick is the primary component of SVM that is renowned for its significance. A Kernel is a method for calculating the dot product of two vectors, x, and y, in a feature space that is often of very high dimensionality. This is why Kernel functions are occasionally referred to as "generalized dot products". The SVM method can perform a Kernel trick that can overcome the non-linear distribution of data [15], [16].

Table 2. Performance evaluation of different models

Sr.no.	Model name	RMSE	R-squared	MSE	MAE
1	Linear	0.075141	0.99	0.0056462	0.050381
2	Fine tree	0.073444	0.99	0.0053941	0.033671
3	Medium tree	0.076972	0.99	0.0059247	0.035531
4	Coarse tree	0.080925	0.99	0.0065489	0.039339
5	Linear SVM	0.075689	0.99	0.0057289	0.053784
6	Quadratic SVM	0.036556	1.00	0.0013363	0.030937
7	Cubic SVM	0.043361	1.00	0.0018802	0.035833
8	Medium Gaussian SVM	0.049747	0.99	0.0024747	0.035148
9	Coarse Gaussian SVM	0.032607	1.00	0.0010632	0.023245
10	Boosted trees	0.09199	0.98	0.0084621	0.070687
11	Bagged trees	0.095252	0.98	0.009073	0.050177
12	Narrow neural network	0.021983	1.00	0.00048326	0.0054259
13	Medium neural network	0.017198	1.00	0.00029578	0.003339
14	Wide neural network	0.012576	1.00	0.00015816	0.0027143

2.3. Ensemble of trees

- Bagged tree model

We can create a random forest by combining multiple decision trees via a technique called bagging. In this method multiple decision trees are trained on different subsets of training data, randomly sampled with replacement. Each tree undergoes independent training, and the final prediction is derived by averaging the predictions of all the trees [17]. A primary constraint of bagging trees is that it employs the complete feature space during the process of creating splits in the trees. If certain variables within the feature space are indicating specific predictions, there is a possibility of having a cluster of correlated trees, which ultimately leads to an increase in bias and a decrease in variance.

- Boosted tree model

The primary benefit of bagged trees lies in their reliance on multiple decision trees instead of a single one, enabling the utilization of the collective knowledge from numerous models. Decreases variability by taking the average of predictions made by models trained on distinct subsets of data. Efficient for models exhibiting significant variability. Boosting, mitigates bias by iteratively training models that specifically target the errors made by previous models. Suitable for models exhibiting significant bias [18], [19].

2.4. Neural networks

The network learns from input-output data pairs, adjusting its weights and biases to approximate the underlying relationship between the input variables and the target variable [20]. This enables neural networks to perform regression tasks, making them valuable in various predictive and forecasting applications wide neural networks are characterized by having a smaller number of hidden layers (typically 1-2), but a larger number of neurons per layer [21], [22]. Neural networks are an exciting and promising type of ML algorithm that can help us better understandand predict complex patterns and relationships. As the network receives more data, it adjusts its weights and biases to approximate the underlying association between the target variable and the input variables [23]. Neural networks can handle large datasets efficiently, making them suitable for applications with extensive historical load data [24]. This scalability ensures that models can be trained on comprehensive datasets, potentially leading to more accurate forecasts. This makes neural networks useful in various predictive and forecasting applications, as they can perform regression tasks. Wide neural networks have a smaller number of hidden layers (typically 1-2), but a larger number of neurons per layer [25], [26]. A wide neural network model for regression typically involves a neural network architecture with a large number of neurons in its layers. The model can be described mathematically as follows:

For input layer

Let the input features be $x=[x_1, x_2, \ldots, x_n]$

Suppose there are L hidden layers, each with a large number of neurons.

The output layer produces the prediction $\hat{y}$

Then in forward propagation, for first hidden layer l=1

$z^{(1)} = W^{(1)} x + b^{(1)}$

$a^{(1)} = \sigma(z^{(1)})$

For subsequent hidden layers (l=2, 3, …, L)
$z^{(l)}= W^{(l)}a^{(l-1)}+b^{(l)}$
For the output layer
$\hat{y}=W^{(L+1)}a^{(L)}+b^{(L+1)}$

Here, $W^{(l)}$ and $b^{(l)}$ are the weight matrix and bias vector for the l-th layer, respectively, and σ is the activation function.

3. MODEL EVALUATION

In order to assess the suitability of a model, it is essential to have a performance metric that measures how well it fits the data. It is crucial to ascertain the adequacy of a regression model, which involves assessing whether the model accurately predicts the target variables within an acceptable level of accuracy. These metrics can be used for evaluation to measure the accuracy of a regression model. The following metrics are generally employed for model performance evaluation.

- RMSE

This is a frequently employed metric for evaluating the accuracy of predictions by measuring the Euclidean distance between predicted values and true values. It is frequently employed in supervised learning applications due to its reliance on accurate measurements for each predicted data point. RMSE can be represented as (1).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \|y(i) - \widehat{y(i)}\|^2}{N}} \quad (1)$$

Where N represents the size of the dataset, $y(i)$ is the i-th measurement, and $\widehat{y(i)}$ is its correlative prediction. Having a single numerical metric to assess a model's performance is highly advantageous in ML, whether it is for training, cross-validation, or post-deployment monitoring. RMSE is a highly prevalent metric for this purpose. This scoring rule is both comprehensible and consistent with prevalent statistical assumptions.

- R^2

The coefficient of determination, also known as R^2, is a metric used to examine the accuracy of a regression model. It measures the dispersion of the data points around the regression line that has been fitted. Higher R-squared values indicate a smaller discrepancy between the observed data and the fitted values for the same data set. It also depicts the proportion of the variability in the dependent variable that can be accounted for by a linear model. The (2) defines R^2.

$$R^2 = \frac{Variance\ explained\ by\ the\ model}{Total\ variance} \quad (2)$$

The output of this method varies between 0 and 1, with a value of 1 indicating a perfect fit of the regression line to the data. A value of 0.7 indicates that 70% of the data points are within the range of the regression line.

- MAE

In the domain of ML, absolute error denotes the magnitude of the disparity between the forecasted value of an observation and its actual value. The mean absolute error quantifies the average size of errors in a collection of forecasts, irrespective of their direction. It quantifies precision for variables that have a continuous range of values. Typically, a lower MAE specifies better predictive performance of the model. Nonetheless, the correlation between MAE values and the efficacy of a model is contingent upon the characteristics of the data. It is calculated using (3).

$$MAE = \frac{\sum_{t=1}^{n} |Yt - \widehat{Yt}|}{n} \quad (3)$$

Where, Yt is the actual value and $\widehat{Yt}$ is the predicted value and n is the number of measurement point.

- MSE

MSE is the mean of the squared differences between the actual values and the predicted values obtained from the regression model. We are endeavoring to ascertain the disparity between the real value and the projected value, and then compute its square. The formula for the same is provided as (4).

$$E = \frac{1}{2n} * \sum_{i=1}^{n} \left(Y_{pred} - Y\right)^2 \quad (4)$$

Here, E represents the model's error. The variable Y_{pred} represents the model's output for the given data, while Y represents the expected output. n denotes the quantity of data rows that we feed into the model. The purpose of squaring the error is to eliminate any negative values. MSE applies a higher penalty to the error compared to the MAE by squaring it.

4. RESULTS AND DISCUSSION

ML models are adaptive and continuously evolve by assimilating new sample data and experiences. Therefore, the models are capable of discerning the patterns within the data. Initially, fourteen models from five different families were trained.

4.1. Response plots

Response plots for the selected fourteen models are plotted. Figure 2 shows the response plots of thirteen different models during the training of each model. It's a dot graph between the response and record number. Where the response is plotted against the record numbers i.e. 21891 row entries. This includes the date, day hour of day voltage and current at that hour and maximum and minimum temperature at that hour. Blue dots represent actual response whereas yellow dots represent predicted response. Referring Figures 2(a)-2(n) we can see the responses overlapping and some of the responses are singular. This is visible in almost all the models in more or less quantity. In bagged tree and boosted tree model the overlapping is less, so different colored dots are remarkably distinguishable. In other models like linear, fine tree, medium tree and coarse tree models also some points are not overlapping, but the percentage is lesser that the bagged and boosted tree model. There is greater superimposition of points in neural network models. Wide neural network shows highest superimposition as compared to narrow and medium neural networks.

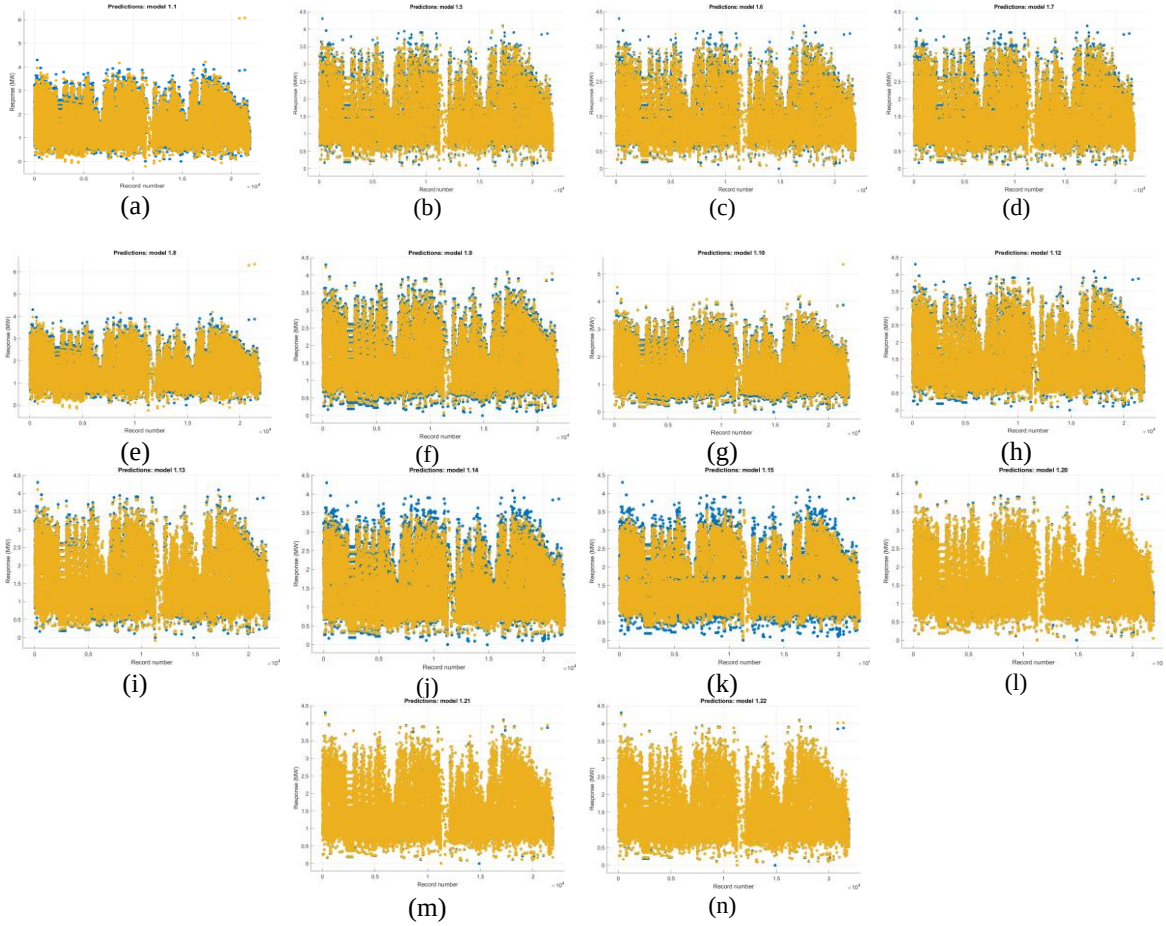

Figure 2. Response plots of various models: (a) linear, (b) fine tree, (c) medium tree, (d) coarse tree, (e) linear SVM, (f) quadratic SVM, (g) cubic SVM, (h) medium Gaussian SVM, (i) coarse Gaussian SVM, (j) boosted trees, (k) bagged trees, (l) narrow neural network, (m) medium neural network, and (n) wide neural network

4.2. Prediction versus actual plot

Figures 3(a)-3(n) display the predicted values against the actual values to assess the performance of the model. The objective of this plot is to assess the predictive accuracy of the regression model across various response values. Typically, a well-constructed model exhibits data points that are distributed in a roughly symmetrical manner around the diagonal line. In an ideal regression model, the predicted response matches the true response, resulting in all data points aligning perfectly along a diagonal line. The error of the prediction for any point is determined by the vertical distance from the line to that point. An effective model exhibits minimal errors, resulting in predictions that are widely dispersed. Referring the Figures 3(a)-3(n), we can see that quadratic SVM shows the best desired spreading of the points. Whereas bagged tree model shows wider spread of the points along the diagonal.

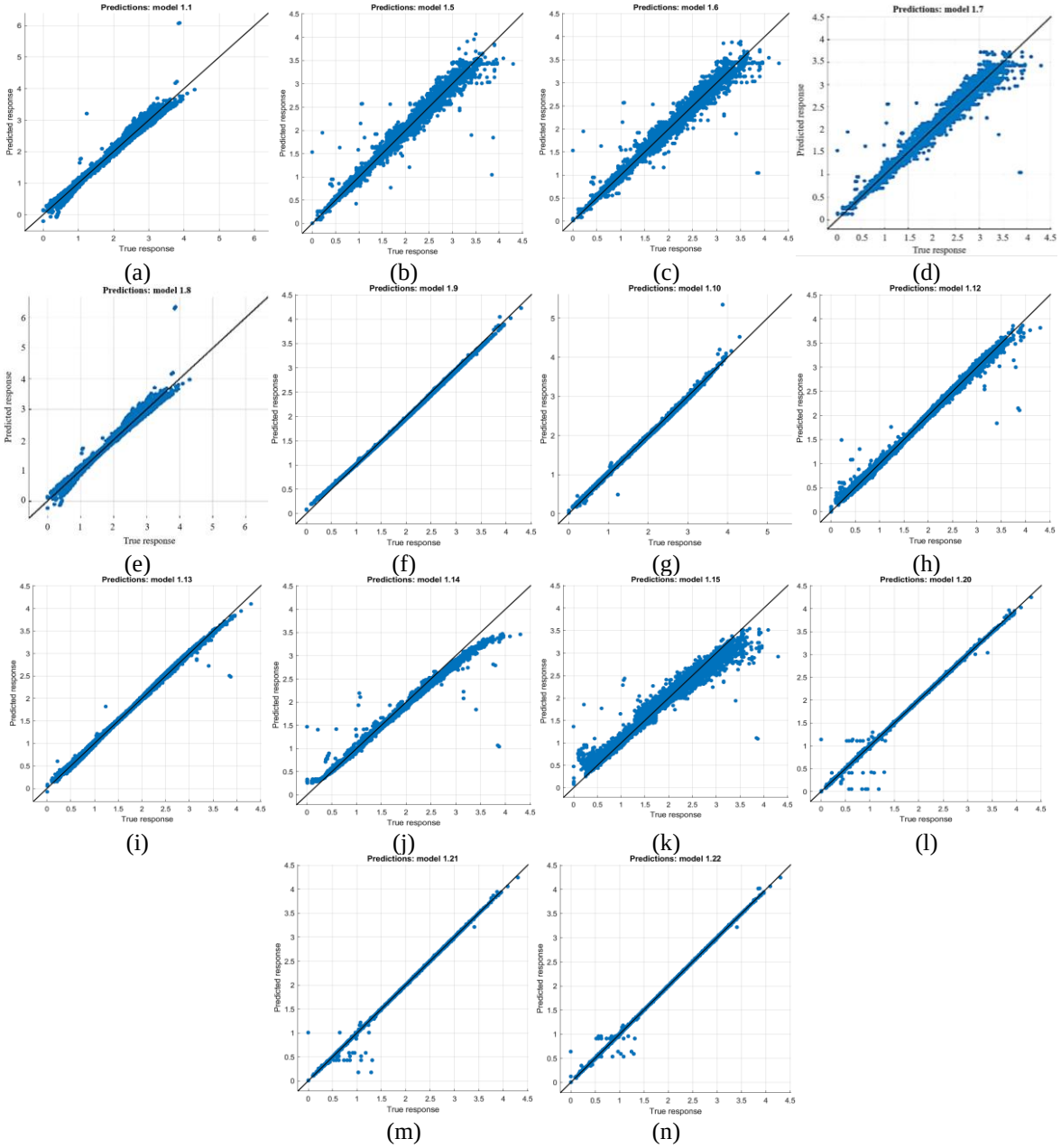

Figure 3. Predictions by various models narrow neural network: (a) linear, (b) fine tree, (c) medium tree, (d) coarse tree, (e) linear SVM, (f) quadratic SVM, (g) cubic SVM, (h) medium Gaussian SVM, (i) coarse Gaussian SVM, (j) boosted trees, (k) bagged trees, (l) narrow neural network, (m) medium neural network, and (n) wide neural network

4.3. Residual plots

Figures 4(a)-4(n) display residual plots for various models. A residual is a quantitative indicator of the vertical distance between a data point and the regression line. Essentially, it refers to the discrepancy between a projected value and the verified real value. Typically, a reliable model exhibits residuals that are distributed in a roughly symmetrical manner around zero. These plots indicate if the model has fully captured the predictive information of the data, resulting in the residuals being affected. A perfect residual plot should exhibit a concentrated cluster of data points in close proximity to the origin while displaying a sparse distribution of points further away from the origin. Additionally, the plot should demonstrate symmetry with respect to the origin. Every regression model inherently possesses a certain degree of error due to the impossibility of achieving 100% accurate predictions. Therefore, a regression model can be defined as: Response = Deterministic + Stochastic.

A model or process is considered stochastic when it incorporates randomness, which means that it can generate varying outputs when provided with identical inputs. In deterministic models, the results are completely determined by the inputs to the model, meaning that if the same inputs are used, the outputs will be the same. Here, the regression model is employed to capture the deterministic component of the model. The equation model should ideally precisely capture the predictive information. The remaining residuals should be entirely stochastic, meaning they are completely random and unpredictable. In our results, the cubic SVM model shows a good residual plot. The performance of all the models trained are evaluated considering different parameters like RMSE, R2, MSE, MAE, prediction speed (observation/sec), and training time required in seconds, which are tabulated in Table 2.

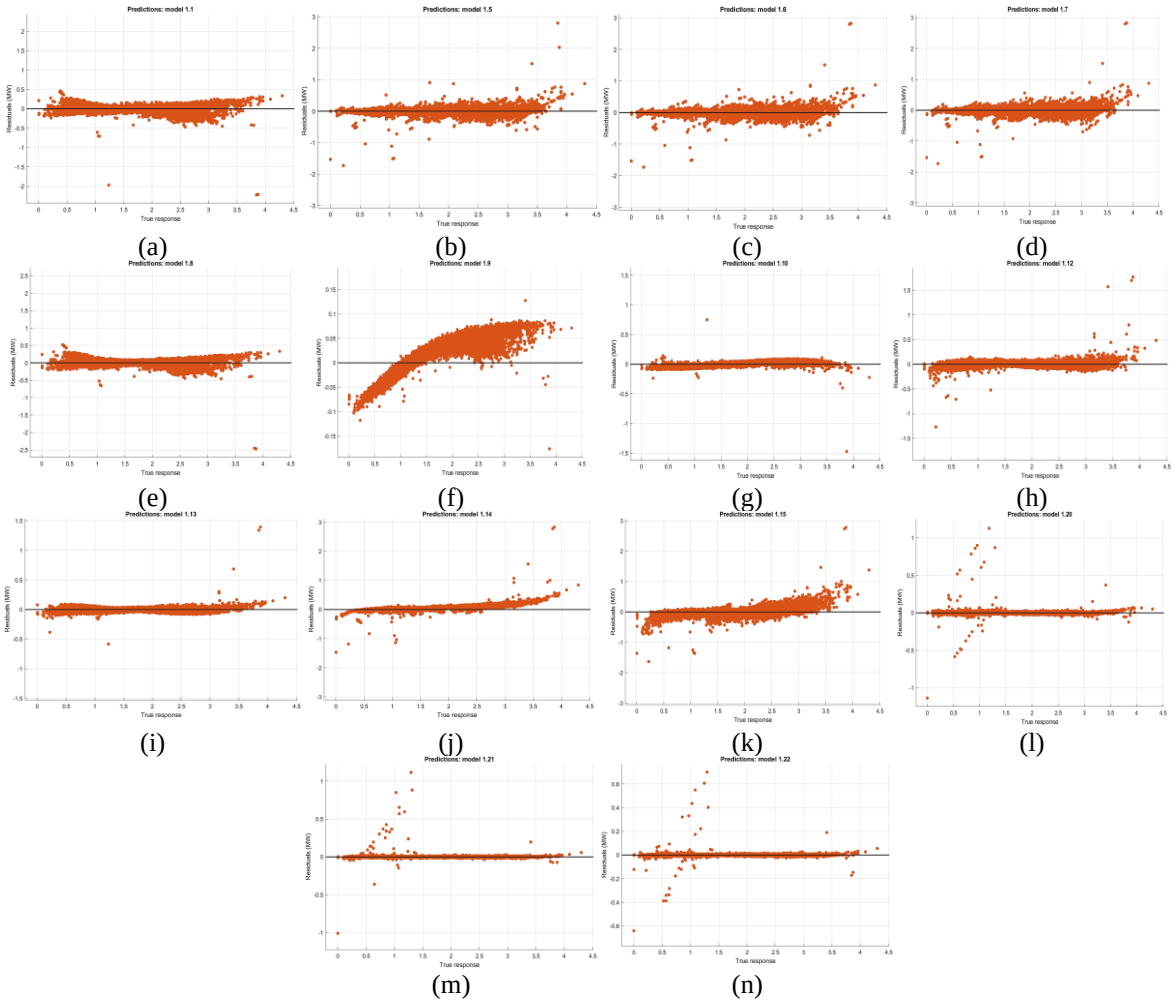

Figure 4. Residual plots for various models: (a) linear, (b) fine tree, (c) medium tree, (d) coarse tree, (e) linear SVM, (f) quadratic SVM, (g) cubic SVM, (h) medium Gaussian SVM, (i) coarse Gaussian SVM, (j) boosted trees, (k) bagged trees, (l) narrow neural network, (m) medium neural network, and (n) wide neural network

Table 2 shows that the wide neural network model has the lowest RMSE value of 0.012576 during validation, with the medium neural network model following closely behind. In both scenarios, the R-squared value is 1. The narrow neural network model has the lowest MSE value of 0.00015816 and a MAE value of 0.0027143. Although this model can handle 35000 data points, it takes the longest training time of 1941.4 seconds. The training time for the coarse tree and medium tree models is significantly short, only 2.158 seconds with a minimum leaf size of 36 and 2.2251 seconds with a minimum leaf size of 12. The quadratic SVM model demonstrates the top performance among SVM models, with RMSE of 0.036556, MSE of 0.0013363, and MAE of 0.030937. The duration of the training process is 41.118 seconds. After analyzing all the models, we discovered that wide neural networks produced the most superior results with minimal effort. After calculating RMSE and MSE values for all models, we concluded that the wide neural network produced the most optimal result with the lowest value. Values for RMSE and MSE. Once the model is selected, it can be tuned for optimized parameters. To avoid overfitting, five-fold cross-validation is performed in this work. The tuning parameters are number of fully connected layers, and the regularization strength (ʎ) value. The higher value of ʎ will result in underfitting of the plot and the lower value shows overfitting of the values. With three fully connected layers, the first layer size, is 100 with the second-and-third-layer size 10. With these parameter settings, the results obtained are shown in Table 3.

Table 3. Performance of WNN after parameter tunning

Parameter setting	RMSE	MSE	MAE	Prediction speed	Training time sec
No. of layers=3, ʎ=0	0.01066	0.00011363	0.003187	30000	2428.6
No. of layers=3, ʎ=0.1	0.11393	0.012981	0.078003	31000	256.84

5. CONCLUSION

Electricity usage is influenced by a range of factors, including weather, time of day, holidays, and social events. ML algorithms can comprehend complex relationships and create adaptable models that respond to altering conditions. In this paper, an effective forecasting approach for the 33/11 kV substation at Wavi, Nasik, India. 14 regression models are evaluated based on different performance indices initially. Among the 14 models studied, a wide neural network model is recommended based on RMSE, MSE, and MAE. Many researchers have explored various regression models for forecasting. But the performance of neural network models in regression remained unexplored many researchers have utilized SVM and GPR techniques. The current study evaluates the precision of SVM, decision trees, and neural networks. The results show that wide neural networks had the best performance, with a regression error of 0.01066, and an MSE of 0.00011363. The study investigates the effectiveness of neural network models using the regression method. This implies that wide neural networks have significant promise for precise electrical load forecasting. This implies that wide neural networks have significant promise for precise electrical load forecasting. As research advances, we expect to see the development of even more sophisticated ML models. Utilizing smart grids and renewable energy sources will enhance forecasting accuracy and drive towards a more resilient and eco-friendly energy future.

While neural networks offer significant advantages for tasks like electrical load forecasting, they also come with several challenges. These challenges can impact their performance, usability, and integration into real-world applications. Large and good quality data, high computational cost, are some of the challenges. Overcoming challenges like Model complexity, interpretability, and training time requires a combination of advanced techniques, robust infrastructure, and interdisciplinary expertise. Continuous research and development in neural network methodologies, along with improvements in data management and computational resources, are essential to mitigate these challenges and fully harness the potential of neural networks.

ACKNOWLEDGEMENTS

Authors thank Mr. S. S. Khandare, Chief Engineer MSEB Nasik Division India for his help in acquiring site data and site visits for this research work.

FUNDING INFORMATION

No funding involved in this project.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Sushama D. Wankhade	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Babasaheb R. Patil		✓				✓				✓	✓	✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Derived data supporting the findings of this study are available from the corresponding author [SDW], on request.

REFERENCES

[1] A. Alhendi, A. S. Al-Sumaiti, M. Marzband, R. Kumar, and A. A. Z. Diab, "Short-term load and price forecasting using artificial neural network with enhanced Markov chain for ISO New England," *Energy Reports*, vol. 9, pp. 4799–4815, 2023, doi: 10.1016/j.egyr.2023.03.116.

[2] M. Cordeiro-Costas, D. Villanueva, P. Eguía-Oller, M. Martínez-Comesaña, and S. Ramos, "Load forecasting with machine learning and deep learning methods," *Applied Sciences (Switzerland)*, vol. 13, no. 13, 2023, doi: 10.3390/app13137933.

[3] J. S. Chou and D. S. Tran, "Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential householders," *Energy*, vol. 165, pp. 709–726, 2018, doi: 10.1016/j.energy.2018.09.144.

[4] R. Wazirali, E. Yaghoubi, M. S. S. Abujazar, R. Ahmad, and A. H. Vakili, "State-of-the-art review on energy and load forecasting in microgrids using artificial neural networks, machine learning, and deep learning techniques," *Electric Power Systems Research*, vol. 225, 2023, doi: 10.1016/j.epsr.2023.109792.

[5] X. Chen, J. Peng, Y. He, B. Zhang, D. Jiang, and P. Lin, "Power load forecasting algorithm based on regression support vector machine," in *Lecture Notes in Electrical Engineering*, vol. 895, 2022, pp. 1268–1274, doi: 10.1007/978-981-19-4775-9_163.

[6] M. W. Nur Azami, I. Wahyudi Farid, C. W. Prianandа, and A. Musthofa, "Electric load forecasting on smart energy meter (SEM) using linear regression," in *2023 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation, ICAMIMIA 2023 - Proceedings*, 2023, pp. 793–798, doi: 10.1109/ICAMIMIA60881.2023.10427910.

[7] S. Gochhait and D. Sharma, "Regression model-based short-term load forecasting for load despatch centre," *Journal of Applied Engineering and Technological Science (JAETS)*, vol. 4, no. 2, pp. 693–710, Jun. 2023, doi: 10.37385/jaets.v4i2.1682.

[8] K. Hopf, H. Hartstang, and T. Staake, "Meta-regression analysis of errors in short-term electricity load forecasting," in *e-Energy 2023 Companion - Proceedings of the 14th ACM International Conference on Future Energy Systems*, 2023, pp. 32–39, doi: 10.1145/3599733.3600248.

[9] K. Akash, V. Arulmozhi, A. Balakrishnan, B. Harisankar, and J. Anudev, "Comparison of load forecasting algorithm for industrial loads," in *2023 International Conference on Control, Communication and Computing, ICCC 2023*, 2023, doi: 10.1109/ICCC57789.2023.10165318.

[10] H. A. El-Sadah and S. H. Z. Al-Thalabi, "Modeling and forecasting using support vector regression and chaotic algorithms/applied study," in *AIP Conference Proceedings*, 2023, p. 050015, doi: 10.1063/5.0119575.

[11] A. Yadav, A. Kumar, R. P. S. Rana, M. Chandrakar, M. Pazoki, and R. A. El Sehiemy, "An efficient monthly load forecasting model using Gaussian process regression," in *2021 IEEE 4th International Conference on Computing, Power and Communication Technologies, GUCON 2021*, 2021, doi: 10.1109/GUCON50781.2021.9574008.

[12] M. Dinesh Reddy and N. Vishali, "Load forecasting using linear regression analysis in time series model for RGUKT, R.K. valley campus HT feeder," *International Journal of Engineering Research and Technology*, vol. V6, no. 05, 2017, doi: 10.17577/ijertv6is050443.

[13] M. Huang, "Theory and implementation of linear regression," in *Proceedings - 2020 International Conference on Computer Vision, Image and Deep Learning, CVIDL 2020*, 2020, pp. 210–217, doi: 10.1109/CVIDL51233.2020.00-99.

[14] D. Maulud and A. M. Abdulazeez, "A review on linear regression comprehensive in machine learning," *Journal of Applied Science and Technology Trends*, vol. 1, no. 2, pp. 140–147, 2020, doi: 10.38094/jastt1457.

[15] B. Gaye, D. Zhang, and A. Wulamu, "Improvement of support vector machine algorithm in big data background," *Mathematical Problems in Engineering*, vol. 2021, 2021, doi: 10.1155/2021/5594899.

[16] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189–215, 2020, doi: 10.1016/j.neucom.2019.10.118.

[17] P. Y. Taser, "Application of bagging and boosting approaches using decision tree-based algorithms in diabetes risk prediction," in *7th International Management Information Systems Conference*, 2021, p. 6, doi: 10.3390/proceedings2021074006.

[18] A. Lasisi, M. O. Sadiq, I. Balogun, A. Tunde-Lawal, and N. Attoh-Okine, "A boosted tree machine learning alternative to predictive evaluation of nondestructive concrete compressive strength," in *Proceedings - 18th IEEE International Conference on Machine Learning and Applications, ICMLA 2019*, 2019, pp. 321–324, doi: 10.1109/ICMLA.2019.00060.
[19] S. González, S. García, J. Del Ser, L. Rokach, and F. Herrera, "A practical tutorial on bagging and boosting based ensembles for machine learning: Algorithms, software tools, performance study, practical perspectives, and opportunities," *Information Fusion*, vol. 64, pp. 205–237, Dec. 2020, doi: 10.1016/j.inffus.2020.07.007.
[20] M. A. Nielsen, *Neural networks and deep learning*. Determination Press, 2015. [Online]. Available: https://books.google.co.id/books?id=STDBswEACAAJ
[21] E. Grossi and M. Buscema, "Introduction to artificial neural networks," *European Journal of Gastroenterology & Hepatology*, vol. 19, no. 12, pp. 1046–1054, Dec. 2007, doi: 10.1097/MEG.0b013e3282f198a0.
[22] K. Tyagi, C. Rane, Harshvardhan, and M. Manry, "Regression analysis," in *Artificial Intelligence and Machine Learning for EDGE Computing*, vol. 382, no. 8520, Elsevier, 2022, pp. 53–63, doi: 10.1016/B978-0-12-824054-0.00007-1.
[23] M. J. A. Shohan, M. O. Faruque, and S. Y. Foo, "Forecasting of electric load using a hybrid LSTM-neural prophet model," *Energies*, vol. 15, no. 6, 2022, doi: 10.3390/en15062158.
[24] S. C. Lim, J. H. Huh, S. H. Hong, C. Y. Park, and J. C. Kim, "Solar power forecasting using CNN-LSTM hybrid model," *Energies*, vol. 15, no. 21, 2022, doi: 10.3390/en15218233.
[25] T. Gao, X. Huo, H. Liu, and H. Gao, "Wide neural networks as Gaussian processes: lessons from deep equilibrium models," in *37th Conference on Neural Information Processing Systems*, 2023. [Online]. Available: http://arxiv.org/abs/2310.10767
[26] J. Xi, O. K. Ersoy, J. Fang, M. Cong, X. Wei, and T. Wu, "Scalable wide neural network: A parallel, incremental learning model using splitting iterative least squares," *IEEE Access*, vol. 9, pp. 50767–50781, 2021, doi: 10.1109/ACCESS.2021.3068880.

BIOGRAPHIES OF AUTHORS

Sushama D. Wankhade is currently pursuing the Ph.D. degree in Electrical Engineering from Sardar Patel College of Engineering, Andheri, and the University of Mumbai, India. She is working as an assistant professor with the Department of Instrumentation Engineering, Bharati Vidyapeeth College of Engineering, Navi Mumbai, India. Her current research interests include renewable energy resources, integration of RE, ESS, and artificial intelligence techniques. She can be contacted at email: sushamawankhade13@gmail.com.

Babasaheb R. Patil received the B.E. degree in Electrical Engineering from, Shivaji University, India in 1989, the M.E. degree in Electrical Engineering (specialization in control systems) from VJTI, University of Mumbai, India, in 1999 and, the Ph.D. degree in Electrical Engineering from Bhabha Atomic Research Centre, University of Mumbai, India, in 2011. Currently he is the Principal of Vishwaniketan's Institute of Management Entrepreneurship and Engineering Technology (ViMEET), Mumbai, India. He has more than 23 years of experience in teaching engineering students and training industrial employees. His fields of interest are biomedical engineering, electrical engineering, and nuclear instrumentation. He can be contacted at email: principal.vimeet@vishwaniketan.edu.in.