Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Integrating Swarm Intelligence and Deep Learning for Enhanced Accuracy in Electrical Load Prediction

Sajawal khan
The University of Faisalabad

Muhammad Zunair Zamir
Changan University China

Maria Saman
Changan University China

Arqam Bashir
The University of Faisalabad

Ahmed Shahzad
The University of Faisalabad

Muhammad Mudassir Riaz
The University of Faisalabad

Hina shabbir
The University of Faisalabad

ABSTRACT

In this modern era, energy became an essential need for human's daily life. This fluctuation can be solved only with accurate load forecasting. In this paper feature engineering has divided into feature extraction and selection parts, this part has been used to select the most important features and reject irrelevant features. To perform the accurate load forecasting a combination of feature engineering, data-mining, and machine learning techniques has been used. In feature engineering combination of Decision Tree, XGBoost and Relief-F techniques have been proposed. Most relevant features forward to machine learning techniques for classification. To improve the accuracy, rate a combination of Particle Swarm Optimization (PSO) and Convolutional Neural Network (CNN) has been implemented for classification. Forecasting also has been performed with the proposed technique and some conventional machine learning techniques such as; Convolutional Neural Network (CNN), Support Vector Machine (SVM), Random Forest (RF), Linear Discriminant Analysis (LDA) and Linear Regression (LR). Comparative analysis has been done with the help of Mean Square Error, Root Mean Square Error, Root Mean Absolute

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Error, F1_Score, and Recall values. The comparison shows that the proposed technique has fewer error rates and higher accuracy rates.

Keywords: Load Forecasting, Feature Engineering, Machine Learning, Convolutional Neural Network, Optimization

Introduction:

Electricity becomes an essential part of every human life now. As the population increasing, the demand for electricity is increasing too [1]. Through forecasting, it becomes easy to know about the future required electrical load. It is necessary to make a balance between generation and consumption of electricity. Smart Grid (SG) is a modern term of conventional grid stations. Through SG electricity load can be managed easily. The NIST (National Institute of Standards and Technology) smart grid model and design is one of the widely used and accepted [22]. It has seven parts such as; Operation, Markets, Generation, Supply, Transmission, Users, and Service-Providers, etc. [23].

Numerous significant choices in SG are taken dependent on load forecasting, for example, distribution of load, load management, and arranging. There are plenty of elements, for example, stickiness, season, wind and temperature, and so forth greatly affect load forecasting. Henceforth, exact LF as for load utilization and consumption isn't a simple undertaking. Distribution systems have important economic implications, and related investments require extensive planning, operation, design, and construction. The complexity & number of reflections that affect system planning initially seem shocking. Significant increases in power capacity, increased distribution voltage, increased automation, and increased control requirements are just a few of these factors. The planning problem is to try to minimize the cost of the sub-transmission system, substation, feeder, horizontal, etc. and the cost of the loss. Besides, distribution system planners do not affect certain factors, but these factors should still be considered when planning a long-term distribution system. In fact, with this set of requirements and constraints, the problem of optimal delivery plans goes beyond human solutions. The traditional grid becomes the old term, now modern grid used that’s called Smart Grid (SG). With the help of artificial intelligence implementation in the smart grid system, it became easy to manage, plan, and distribution of electrical load. Smart Meter (SM) also used in SG to facilitate consumers to make communication with utility. Through Demand Side Management (DSM) of SG, It became very easy for consumers to shift their usage of appliances from the highest peak hours to the lowest peak hours [21]. Management of energy demand, also known as demand-side-management (DSM) or demand-side response (DSR), is a way to change consumer demand for energy in a variety of ways, including financial incentives and changing behavior through education. Typically, the goal of demand-side management is to reduce peak consumer power consumption or to change energy usage to off-peak hours, such as nights or weekends. Peak demand management does not necessarily reduce total energy consumption, but it can be expected to reduce the need for investment networks and/or power plants to meet peak demand. As an example, an energy storage unit is used to store energy during off-peak hours and to discharge during peak hours. DSM's new application is to help

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

grid operators' balance the intermittent generation of wind and solar generators, especially when the time and scale of energy demand does not match renewable energy generation. The US power industry, whether it is consumables or fossil fuels, initially relied heavily on foreign energy imports to generate electricity. During the energy crisis of the 1970s, the federal government hoped to pass the Public Utilities Regulatory Policy Act (PURPA) to reduce dependence on foreign oil, improve energy efficiency and alternative energy sources. The law forces power companies to obtain the cheapest possible power from independent power companies, thereby promoting renewable energy and encouraging them to reduce the power they need to improve energy efficiency and demand management challenges. The term DSM was created during the 1973 energy crisis and the 1979 energy crisis. Many governments have mandated the implementation of various demand management plans. An early example was the 1978 US National Energy Conservation Policy Act, followed by similar measures in California and Wisconsin. Demand-side management was released by the Electric Power Research Institute (EPRI) in the 1980s. Today, DSM technology integrates with power systems through information and communication technologies, and new terms such as integrated demand-side management (IDSM) or smart grid are becoming more and more feasible.

Forecasting is a method to know about the required electrical load in the future. Forecasting has been divided into four parts such as; Very Short Term Electrical Load Forecasting (VSTLF), Short Term Electrical Load Forecasting (STLF), Long Term Electrical Load Forecasting (LTLF) and in the end Medium Term Electrical Load Forecasting (MTLF). VSTLF helps to forecast less than 24-hours ahead, STLF used to know 1 day to one week load, LTLF predicts 1-year to 10-years ahead required load [19].

The latest framework has been used. In this work, the forecasting with the combination of Feature engineering, data mining, and machine learning methods has been performed. The feature engineering part consists of feature selection and feature extraction methods. CNN + PSO machine learning technique has been proposed for classification. Main contributions of this work described below:

Feature selection and extraction method has been used to remove redundant values.
PSO technique has been used to avoid difficulty to tune the parameters of CNN
Result assessed with Mean Square Error, Root Mean Square Error, Root Mean Absolute Error, F1Score, and Recall values
Proposed technique also has been compared with some conventional techniques.

REVIEW OF LITERATURE

Since precise and proficient load forecasting is one of the most significant prerequisites of the power system framework arranging, execution, activity, control, and upkeep, it has pulled in huge measure of specialists and many research gatherings to work in the region. An enormous number of papers and reports have been distributed in the load forecasting space in more than recent decades. Since the exploration work introduced in the theory creates methods for productive and exact transient load forecasting, the writing audit introduced in this part centers around the Short term load forecasting (STLF) and medium-term load forecasting (MTLF) writing accumulated from various sources. At first, an outline of the accessible STLF

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

& MTLF work done as such far is introduced, at that point, the papers managing issues of the variable choice relating to the execution of STLF & MTLF is done, trailed by the survey of the papers dependent on use of explicit procedures for various STLF & MTLF approaches lastly the particular papers that have been legitimately or in a roundabout way utilized in the proposed work of the postulation are evaluated.

In [13], an innovative approach to predict short-term electricity load using a combination of the Wavelet Multiresolution (WM) algorithm and transfer learning techniques. The WM algorithm is utilized to decompose the time series data into multiple components, enhancing the accuracy of load forecasting by capturing essential patterns and trends. Transfer learning is then applied to leverage knowledge from related tasks or domains, improving prediction performance, particularly in scenarios with limited training data. This hybrid model demonstrates superior accuracy and robustness compared to traditional forecasting methods, making it highly valuable for real-time energy management and grid stability.

Authors applied various AI techniques in [1] to evaluate the electrical load forecast. The proposed techniques are random forest (RF), support vector regression (SVR), extreme gradient boosting (XGBoost), multilayer perceptron (MLP), long short-term memory (LSTM), and temporal convolutional network (Conv-1D). The methodology used to enhance the robustness of the best AI method for modelling and forecasting the building's electricity consumption by taking the biasness and variance of the models in consideration. These methods are applied in an American single family home to access their effectiveness. The 10-fold cross-validation technique is utilized to evaluate the comparison performance by examining their bias and variance. The accuracy of the models regarding the future prediction can be guided through the two distinct sets, called the validation and test sets. Result represented that among all AI techniques LSTM gives best result by showing the minimum amount of dispersion in both the test (nMBE, nRMSE) and validation sets (nMBE, nRMSE). In [2], investigate nine distinct hybrids between clustering and recurrent neural networks. The test cases are five commercial buildings spread across five distinct locales in Bangkok, Thailand; Hyderabad, India; Virginia, USA; New York, USA; and Massachusetts, USA. It divides the buildings into five different building types: academic, research laboratory, office, school, and grocery store. The load forecasting findings show that, in comparison to the existing state-of-the-art results for both hour-ahead and 24-ahead load forecasting, the deep learning algorithms implemented in this study give 20–45% improvement in load forecasting performance. According to sensitivity analysis, it is discovered that: hybrid deep learning algorithms can produce satisfactory hour-ahead load prediction with as little as one month of data; 15-min resolution data, if available, improves hour-ahead load forecasting by 30%; and the developed methods are resilient against errors in weather forecasting. For the short-term building-level electrical load forecasting tasks, the forecasting results across all five buildings support the robustness of the proposed deep learning architecture. An extensive analysis has been done in [3] for the current DL-based methods, these methods are designed for electric power load forecasting as well as power forecasting of solar panels and wind turbines. Additionally, it also covers the datasets that were used to test and train the various DL-based prediction models to help the future researchers choose the right datasets for their work. Authors claimed in his research

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

that not a single poll examines production and load side forecasting together. Despite their importance in DL-based forecasting techniques, the datasets utilized for forecasting have not been taken into account in prior surveys.

The long short-term memory (LSTM) layer models has been applied in [4] (short term forecasting load), the model associations between previously seen data by receiving historical data as input. The fully-connected layer's outputs, which are used as input for prediction data, such as weather information for forecasting days, are combined with the LSTM layer's outputs. The suggested forecasting method's ideal parameters are chosen after multiple experiments. The accurate load forecasting that the suggested technology provides is anticipated to help to stable power system operation. The [5] created approach hinges on the merging of long short-term memory (LSTM) networks and convolutional neural networks (CNNs). The technique is used to forecast electrical load in the short term for the power system in Bangladesh. Furthermore, by contrasting the forecasting errors with those of various other methods—such as the extreme gradient boosting algorithm, radial basis function network, and long short-term memory network—the efficacy of the suggested technique is confirmed. It is discovered that the suggested approach improves short-term load forecasting's precision and accuracy. In [7], the authors used a deep learning approach that trains a support vector regression (SVR) model to predict the total power load for the next day (short term) after first refining features using stacked denoising auto-encoders (SDAs) on historical electricity load data and associated temperature factors. The abstract features that SADs retrieved from the original electricity load data have been shown to explain and anticipate the load trend more precisely with reduced errors, which is the heterogeneous deep model's most significant contribution. The suggested model is assessed through comparison with both standard SVR and artificial neural network (ANN) models. The outcomes of the experiments corroborate the enhanced performance of the proposed model.

In [6], the research presents a short-term power system load forecasting model based on Bi-directional Long Short-Term Memory (BiLSTM), which is based on an analysis of historical load data from four industries and a regional power grid. This article first selects features for mid-term load forecasting using a random forest and the Pearson correlation coefficient. Next, a hybrid deep learning model built on random forest and BiLSTM is created. A hybrid deep learning-based model for mid-term power system load forecasting is created once the model's parameters have been optimized. Lastly, the models that serve as benchmarks are chosen for comparison studies. The suggested model's MAPE of 2.36%, as indicated by the experimental findings, is higher than that of the benchmark models. This demonstrates shows the suggested hybrid model can successfully raise the power system load forecasting accuracy. The authors analyze and compared LSTM with two conventional methods, namely Exponential Smoothing and ARIMA in [8]. The accuracy of load projections made using the three previously described methodologies is examined and validated using actual data from 12 houses over a period of three months (medium term load forecasting). The authors proved that because of LSTM models are better at remembering huge amounts of data, time series-based prediction uses them as comparison of two other conventional models. In [9], a novel approach for short-term load forecasting is presented that makes use of one-dimensional CNNs based on

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Video Pixel Networks (VPNs), in which the gating mechanism of the VPNs' Multiplicative Units is somewhat altered. In particular, real-world energy load data is applied to the suggested one-dimensional CNNs, LSTM, and GRU variants for 1-hour-ahead and 24-hour-ahead prediction tasks, which are the primary concerns for the electricity supplier corporations for short-term load forecasting. Ten ensemble predictions from each approach were tested in order to determine the significance of the performance differences in the analyses. The suggested one-dimensional CNN model produced the best overall result with 2.21% mean absolute percentage error for 24-hour ahead predictions, based on the comparative analyses' findings. However, there was no discernible difference between the strategies, even though the suggested one-dimensional CNN method produced the greatest results, with a mean absolute percentage error for 1-hour ahead of time of about 1%.

The authors claimed in [10] that in order to capture uncertainty and reflect the range of electrical load fluctuation, they proposes a deep learning strategy based on K-nearest neighbors to solve the above two concerns. By measuring the distance between the training and testing datasets, the K-nearest neighbors approach is first used to find elements of previous electrical load time series that are similar to the future values. Then, for multi-objective optimization, the second generation of the non-dominated sorting genetic algorithm is used to determine the maximum forecasting accuracy and the smallest category number of K-nearest neighbors. The prediction intervals are obtained using modified non-parameter kernel density estimation, which is based on the deep belief network's forecasting outcomes. We use five datasets gathered from Australia to test the efficacy of the suggested methodology. Experimental results verify that the suggested interval forecasting model not only increases forecasting accuracy and efficiency but also streamlines the forecasting process of deep learning approaches, which can be very useful as a reference for future research through a series of comparisons with other cutting-edge models. The tree-based learning algorithms were used in [11] to construct wind power forecasting models extending six months into the future. The authors designed three models to examine the performance and accuracy of the model. Each model differentiate with other by input data, sample time and wind speed. The outcomes of the simulation demonstrated that tree-based learning algorithms may be effectively used for prospective wind power forecasting at various elevations and locations, in addition to long-term wind power forecasting. In [12], the authors combine different models and proposed a hybrid model to examine the prediction accuracy of wind speed. Firstly, k-NN imputation technique is applied in first step, then, integrates the features of multiple data decomposition techniques—EMD, EEMD, and CEEMDAN—with a BiDLSTM Neural Network to denoise the signal. In this study, they examines several wind speed datasets for five Gujarati cities—Chandori, Bhogat, Gandhinagar, Charanka, and Surat—obtained from the NIWE data site in order to assess the forecasting technique's accuracy (India). The empirical result shows that the suggested hybrid strategy based on k–NN–CEEMDAN–BiDLSTM works better than other hybrid strategies found in the literature in terms of prediction accuracy. The authors of [14] suggested from the overview of plenty of papers that Support Vector Machine (SVM) and Artificial Neural Network (ANN) produced exceptional performance as compared to other data-driven tools such as Linear Regression (LR)

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

and Autoregressive Integrated Moving Average (ARIMA). In [15], authors explained the problems facing in smart building energy management (SBEM) and it concluded that deep reinforcement learning (DRL) has special place to resolve the mentioned problem and marked the DRL the efficient model for energy management. The intention of this [16] is to create a model of the electrical energy demand of buildings, which could then be used in conjunction with storage devices to lower power costs. The authors designed multilayer perceptron neural network models, and mathematical modeling is used to achieve energy use optimization. This strategy is implemented on a Bangkok, Thailand, governmental office building. The authors in [17] identify the best LF approach for a given set of SG applications. The results suggest that, in comparison to other approaches, AI-based LF techniques utilizing machine learning (ML) and neural network (NN) models have demonstrated the highest prediction performance, with better overall root mean squared (RMS) and mean absolute percentage error (MAPE) values. The temporal fusion transformer (TFT) architecture and the transfer learning (TL) idea has been combined in [18] to present a revolutionary technique. The authors proved that through this novel method lessens the requirement for massive volumes of data for every new building by transferring information to new users from a generic pre-trained model based on the history data of the community's buildings. The results demonstrate that by lowering RMSE by over 11%, and by including the TL technique reduces CV_RMSE, SMAPE, and WQLoss by over 40%, improving the forecast of load demand of buildings with low historical data availability.

In [20], the authors suggested a method that fits the building baseload, peak load, and occupancy rate with many weather variable factors using fuzzy logic methods and artificial neural networks (ANNs). The proposed working and non-working hour based regression approach reduces the average RMSE by 35%, while the ANN with fuzzy hours based approach reduces the average RMSE by 42%, comparing with the traditional power demand prediction method.

Proposed and Implementation

A detailed discussion about existing and proposed techniques has been presented in this chapter. For this research work, two techniques CNN and PSO have been combined to make the accuracy rate more accurate. Further, a comparison has been made with some existing techniques such as Logistic Regression (LR), Support Vector Machine (SVM), Convolutional Neural Network (CNN), Random Forest (RF) and Linear Discriminant Analysis (LDA). The proposed framework has also been discussed in this chapter.

AdaBoost:

AdaBoost has been used as a feature selector technique. It has been scaled the features and showed the importance of features in the 0 to 1 scale. Weights have been improved between the nearest hit and nearest miss because it can make noise and cause of redundant features if the weights are not balanced. Adaboost is the short form of adopted boosting. It is generally used to convert weak classifiers to the strong classifier. An AdaBoost flow diagram is shown in Figure 2.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

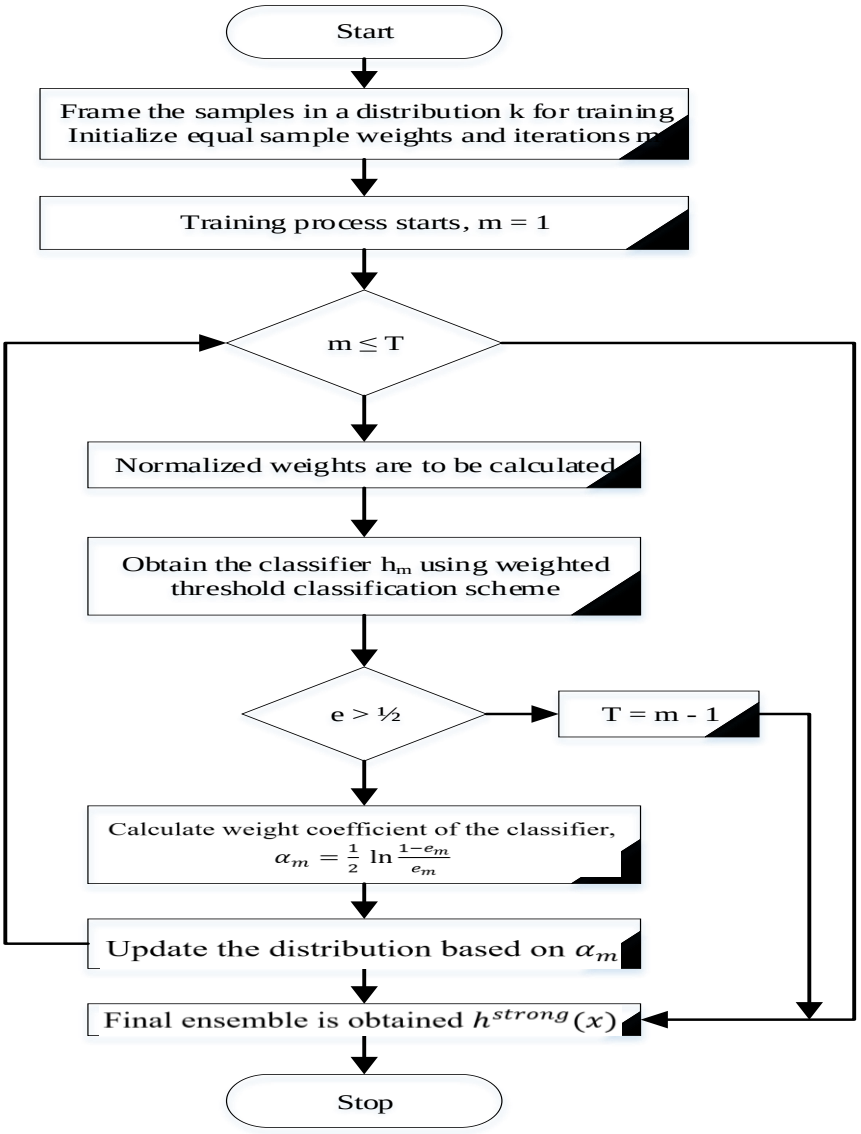


Figure 2: The working flow of Adaboost

Extreme Gradient Boosting:

A decision tree is a machine learning technique that acts as the tree to train data and make useful information. In this research work, Decision Tree is used to scale the features between 0 to 0.8, the highest features have the highest values. It took simple datasets to train them and make a prediction using the DT algorithm. In every iteration, it removes the deficiencies of previous trees and makes the next tree is more useful & helpful for prediction. In figure 3 Decision tree has been shown with an example of buying a new car. If the car is road tested then go for further process otherwise no, in the next process checked for mileage if have a good mileage then go for last otherwise will escape.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

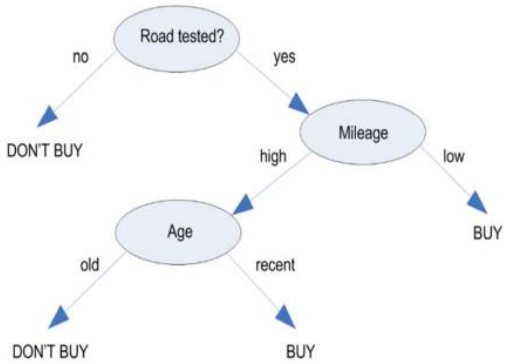


Figure 3: Decision Tree for buying a Car

Relief-F:
Relief-F is used to select the best features and eliminate weak features. RFE requires a predefined number of attributes to keep, anyway it is frequently not known ahead of time 33 what number of attributes are legitimate. A threshold value set for relief-f it will accept or reject features based on this threshold value.

Support Vector Machine:
This technique is part of machine learning. Researchers used it a lot as a classification technique, so, it is used it as a comparison technique. SVM used for two purposes: classification and Regression. SVR used for regression and SVS used for classification. The SVC is used for the purpose of comparison with the linear kernel, gamma, and C values. The value of gamma is 20 and the value of C 10.

Proposed Enhanced Convolutional Technique:
In this research work; a combination of convolutional neural network and particle swarm optimization (PSO) technique has been proposed. PSO is a hyper-tuning algorithm that can be used to tune the parameters of machine learning techniques dynamically. As compared to other tuning algorithms PSO is easy to implement and just needed a few parameters adjustment to implement it. It selects random parameters so it is faster than other techniques. The parameters of enhanced CNN with sequence wise layers is shown in Table 1.

Table 1: ECNN Layers and parameters

Layer (type)	**Output Shape**	**Param #**
con1d_2 (Conv1D)	(None, 11,32)	128
Max_pooling1d_2 (MaxPooling1)	(None, 6, 32)	0
Dropout_3 (Dropout)	(None, 6, 32)	0
Flatten_2 (Flatten)	(None, 192)	0

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Dense_3 (Dense)	(None, 2)	386
Dropout_4 (Dropout)	(None, 2)	0
Dense_4 (Dense)	(None, 1)	3

In figure 4, sequence wise layers of proposed CNN with parameters has been discussed. CNN is the part of deep learning techniques and it is very beneficial to increase the accuracy rate. In deep learning techniques, researchers can increase the layers to improve prediction rates. In my proposed CNN one convolutional layer, one max-pooling layer, one flattens, two drops, and two dense layers have been used.

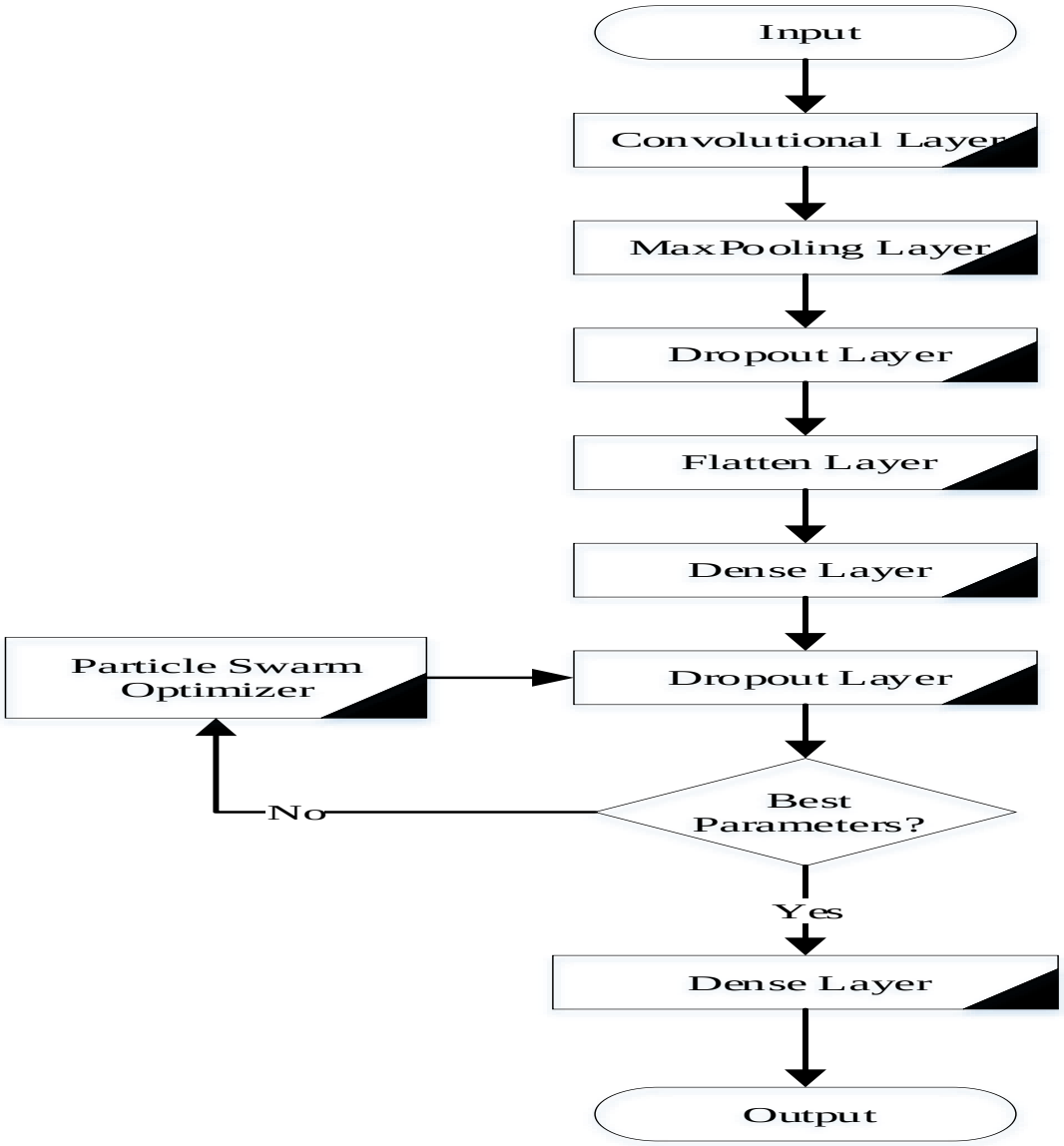


Figure 4: Proposed ECNN

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Proposed System Model:

The proposed system has three parts; Preprocessing, feature engineering, and classification. Data has been arranged in preprocessing step. Feature selection and Feature extraction is the part of feature engineering and classification is the prediction of required electrical load in the future. The proposed model diagram is shown in Figure 5.

Preprocessing

Feature Engineering

Classification

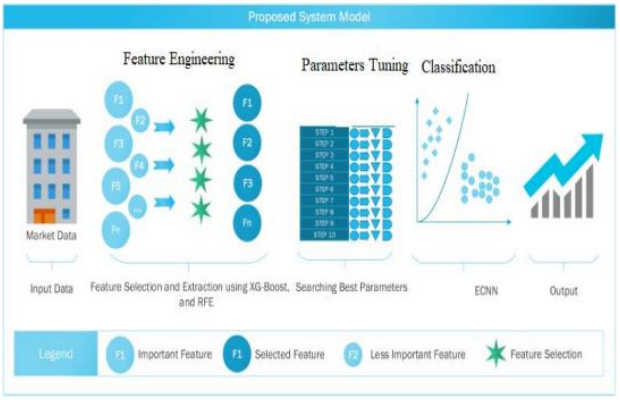


Figure 5: Proposed System Model

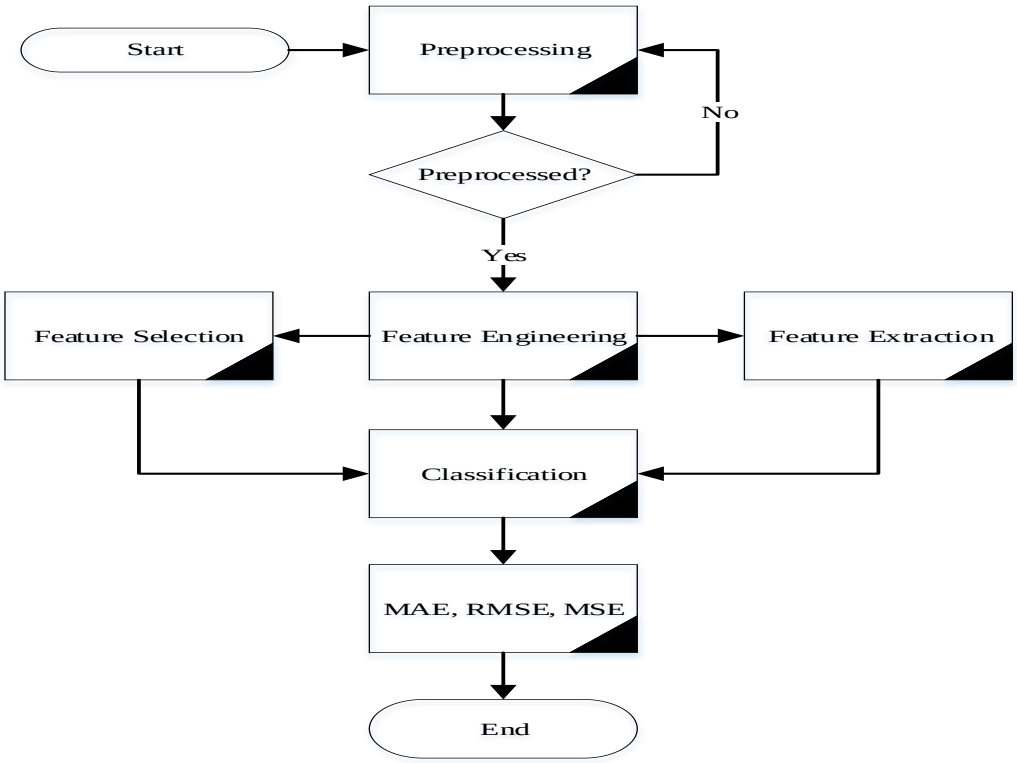


Figure 6: Flowchart

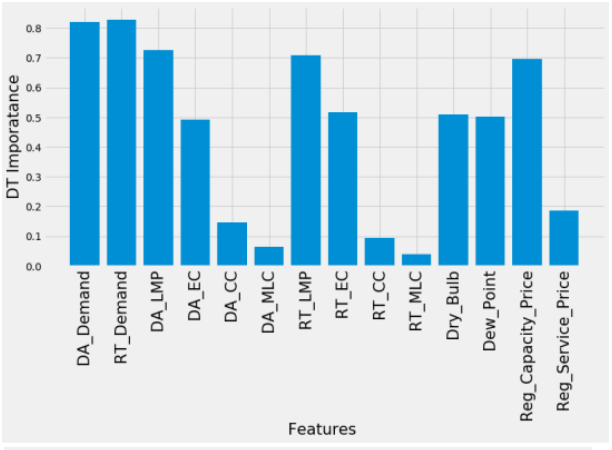
Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Preprocessing: Pre-processing has been performed with sorting the data in month wise. Dataset consisted of fifteen number of attributes. The small dataset consists of 1096 number of features and a large dataset consisted of 2922 number of features. Dataset has consisted of numerical values. 25 % of data used for testing and 75 % data has been divided for training purposes.

Feature Engineering: As discussed before chapters, feature engineering divided into two methods. Feature Selection (FS) and Feature Extraction (FE). In FS DT and AdaBoost used to scale the attributes and in FE method Relief-F used to convert the attributes into two dimensions. In FE a threshold value has been selected for Relief-F, above this threshold value, all attributes selected for extraction and other is rejected. After completing Feature engineering formula is set to show the selected and rejected features. The results of this research work are discussed in the form of graphs and tables. In the feature selection, we find the importance of features. Feature selection with DT and AdaBoost is shown in Figures 7 and 8. In feature extraction, Relief-F is used to convert them into two dimensions. After feature extraction, the 0.5 threshold value is set. Important features are greater or equal than the threshold value. Unimportant features are removed by using both AdaBoost, DT, and Relief-F techniques, which have less value than the threshold value.



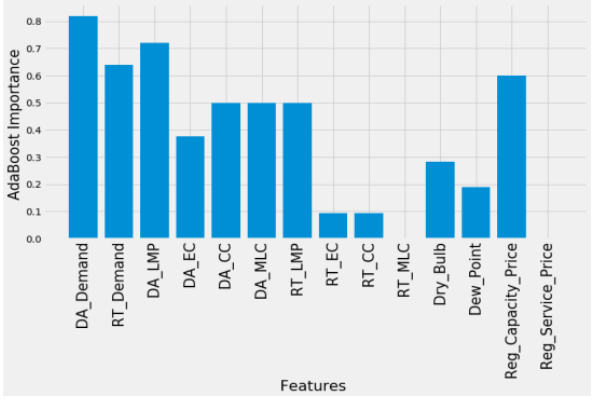


Figure 7: DT Importance & Figure 8: AdaBoost Importance

In above-given figures shown the importance of features calculated by AdaBoost and DT w.r.t 3-Years and 8-Years datasets The scale between 0.0 and 0.8 used to show

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

importance. Attributes of higher levels have higher importance and the lower level has low importance. This importance is very helpful for the further elimination of nonimportant features. Further in FE Relief-F has been used to convert the values of the attributes in numeric form. Relief-F feature importance values have been displayed in Table 2.

Table 2: Feature Priorities calculated by Relief-F

Features	**ADABoost_Feature_Importance**
DA.Demand	1
RT.Demand	0
DA.LMP	6
DA.EC	7
DA.CC	11
DA.MLC	2
RT.LMP	3
RT.EC	10
RT.CC	12
RT.MLC	13
Dry.Bulb	5
Dew.Point	9
Reg_Capacity_Price	4
Reg_Service_Price	8

Table 2 telling about importance of features by Relied-F. In FE, complete features have been converted into importance with numeric values. After completing the feature importance calculation by Relief-F, DT & AdaBoost, threshold values, and a formula set to select and reject over-all features. After completing the feature engineering the formula is set to (RelF_Priorites[i] > 6 and ADBo_importances[i] < 0.6 and ETC_importances[i] < 0.6) values has been used to eliminate the features for final selection. After the 59 execution of the feature selection & extraction framework, the final selected and the rejected feature is shown in Table 3.

Table 3: Selected and rejected features after the feature engineering method

Selected Features RFE	**Rejected Features by RFE**
DA.Demand	DA.EC
RT.Demand	DA.CC
DA.LMP	RT.EC
DA.MLC	RT.CC
RT.LMP	RT.MLC
Dry.Bulb	Dew.Point
Reg_Capacity_Price	Reg_Service_Price

Evaluation and Validation Parameters:

Four error metrics are used to show the results. They all calculate errors and show the accuracy rates concerning errors. Minimum error rates mean the highest accuracy

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

rates. Further, these error rates also subtract into 100 to get accuracy rates. MSE, RMSE, MAE, and MAPE used to represent the error values.
Equations of these metrics given below.

$$MAPE = \frac{1}{T}\sum_{n-1}^{T}|(S - Y_S)|$$

$$RMSE = \sqrt{\frac{1}{T}\sum_{tm-1}^{TM}(Av - Fv)^2}$$

$$MSE = \frac{1}{T}\sum_{Tm-1}^{Tm}|(Av - Fv)^2|$$

$$MAE = \frac{\sum_{n-1}^{N}|(Fv - Av)|}{N}$$

RESULTS AND DISCUSSIONS

All results have been concluded in this section. The development has been performed on two datasets. One is a small dataset and the second is the large dataset. The main object of this study was making a hybrid of CNN + PSO for improving the accuracy rate of load forecasting. For the development of this research, 6 generation HP laptop has been used. The system has 8 Gb of RAM and window 10 init. Spyder 4 software in Anaconda IDE has been used for development purposes. Coding is done in Python language.

Classification:

In this thesis, classification has been used to forecast the electrical power load. One month ahead forecasting has been performed in this research work. Different techniques have been selected to compare with the proposed technique. The comparison with proposed techniques has been shown in graphical form. In figure 9 showing the comparison of the original load and predicted load by all the considered Model.

Figure 9 (a): LinReg Prediction

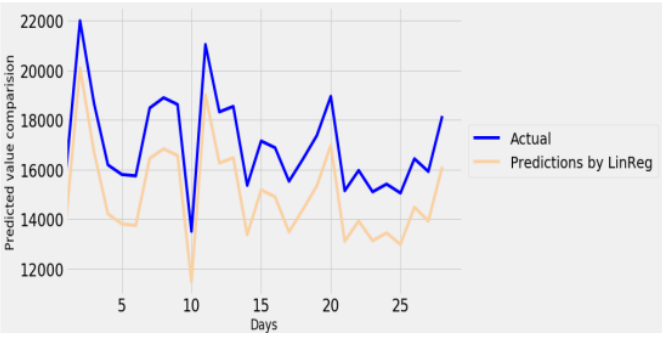


Figure 9 (b): RF Prediction

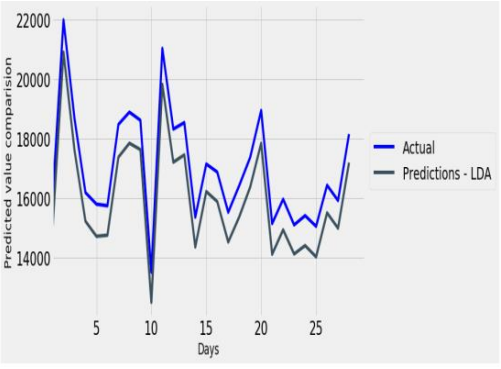
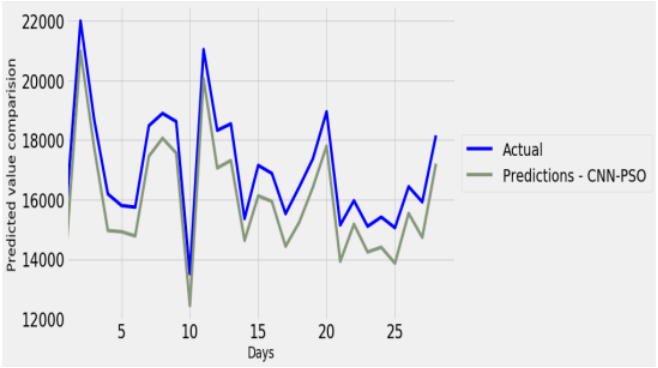


Figure 9 (c): CNN-PSO Prediction

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about



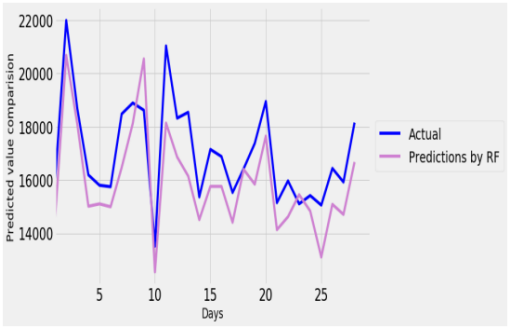
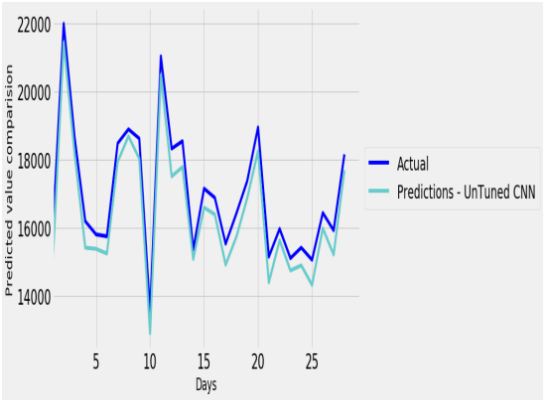


Figure 9 (d): LDA Prediction Figure 9 (e): UnTuned CNN's Prediction

Figure 9 (f): SVM Prediction



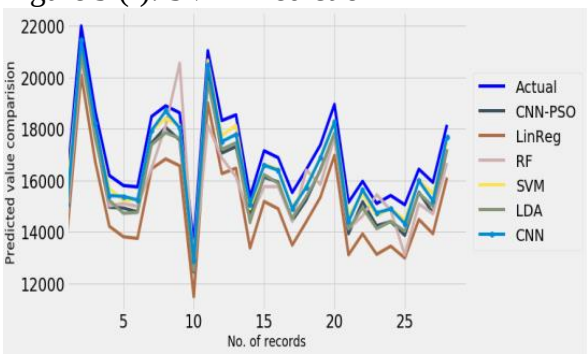
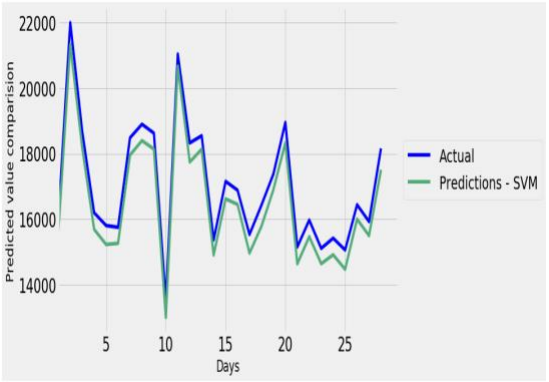


Figure 10: Comparison of Proposed Technique

In figure 9, it is displayed that enhanced techniques have a very similar pattern as the actual load. Conventional CNN, LinReg, RF, LDA & SVM have a lot of variation with the actual load.



Performance Evaluation:

For the assessment of forecasting techniques, four error matrices have been used in this thesis i.e. MAE, F1-Score, MAPE, Precision, RMSE, Recall, and MSE. Figure 4.10, shown the error rates of used techniques, respect to the above-described matrices Further, error rate values are displayed in numeric form in Table 4.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

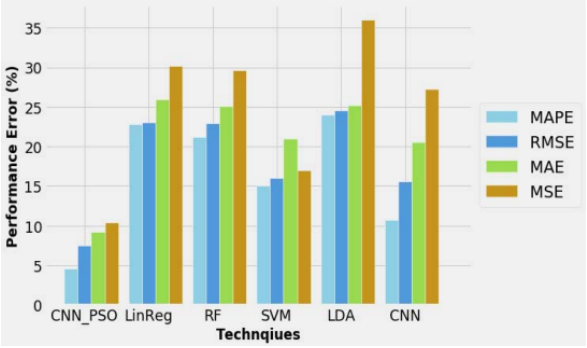


Figure 11: Error Rates

Table 4: Performance Evaluation Table

Model	MSE	MAE	RMSE	R^2-Score
Linear Regression	811.29	22.456	28.483	0.48
Decision Tree	355.64	13.232	18.858	0.77
Lasso Model	811.45	22.463	28.486	0.48
LightGBM	273.52	12.085	16.538	0.82
Catboost	105.27	6.986	10.260	0.93
Random Forest	34.479	3.4861	5.8720	0.97
XGBoost	205.84	10.226	14.347	0.87
RF-XGB Ensemble	24.105	3.1235	4.9098	0.98
SVM	17	21	16	0.87
LDA	36	25.23	24.56	0.81
CNN	27.23	20.56	15.54	0.88
ECNN Model (Proposed)	10.43	9.23	7.45	0.96

Conclusion

This thesis is based on short-term forecasting. The framework that is used in this thesis consists of two methods which are feature engineering and classification. Through feature engineering technique calculated the importance of features and selected the best features with eliminating irrelevant features. Hybrid of DT, AdaBoost, and Relief-F used in the feature engineering method CNN used to perform classification, further, the parameters of CNN are tuned with the PSO technique. The results of the proposed technique also have been compared with some conventional machine learning techniques. According to the results, the proposed technique performed best than all other techniques.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Acknowledge
The Authors declare no conflict of interest.

Reference:

Zhonge Su, Guoqiang Zheng, et al, "An IDBO-optimized CNN-BiLSTM model for load forecasting in regional integrated energy systems" Computers and Electrical Engineering, 2025.

Shimaa A. Ahmed, Entisar H. Khalifa, et al, "Real Time Electrical Load Prediction and Management through Deep Learning and Reinforcement Learning Techniques" Engineering, Technology & Applied Science Research, 2025.

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, p. 109792, 2023.

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, vol. 13, no. 13, p. 7933, 2023.

S. H. Rafi, S. R. Deeba, & E. Hossain, "A short-term load forecasting method using integrated CNN and LSTM network" IEEE Access, vol. 9, pp. 32436-32448, 2021.

H. Xu, G. Fan, G. Kuang, & Y. Song, "Construction and application of short-term and mid-term power system load forecasting model based on hybrid deep learning." IEEE Access, vol. 11, pp. 37494-37507, 2023.

N. Tsalikidis, A. Mystakidis, C. Tjortjis, P. Koukaras, and D. Ioannidis, "Energy load forecasting: One-step ahead hybrid model utilizing ensembling," Computing, vol. 106, no. 1, pp. 241-273, 2024.

R. Mubashar, M. J. Awan, M. Ahsan, A. Yasin, & V. P. Singh, "Efficient residential load forecasting using deep learning approach." International Journal of Computer Applications in Technology, vol. 68(3), pp. 205-214, 2022.

I. Yazici, O. F. Beyca, & D. Delen, "Deep-learning-based short-term electricity load forecasting: A real case application." Engineering Applications of Artificial Intelligence, vol. 109, pp. 104645, 2022.

[10] Y. Dong, X. Ma, & T. Fu, "Electrical load forecasting: A deep learning approach based on K-nearest neighbors." Applied Soft Computing, 99, 106900, 2021

J. Yanmei et al., "Enhanced neighborhood node graph neural networks for load forecasting in smart grid," International Journal of Machine Learning and Cybernetics, vol. 15, no. 1, pp. 129-148, 2024.

K. Sareen, B. K. Panigrahi, T. Shikhola, R. Sharma, "An imputation and decomposition algorithms based integrated approach with bidirectional LSTM neural network for wind speed prediction." Energy, vol. 278, pp. 127799, 2023.

N. Wei et al., "Short-term load forecasting based on WM algorithm and transfer learning model," Applied Energy, vol. 353, p. 122087, 2024.

R. Olu-Ajayi, H. Alaka, H, Owolabi, L. Akanbi, S. Ganiyu, "Data-Driven Tools for Building Energy Consumption Prediction: A Review." Energies, vol. 16, pp 2574, 2023.

Online ISSN 3007-3197
Print ISSN 3007-3189

L. Zhang and D. Jánošík, "Enhanced short-term load forecasting with hybrid machine learning models: CatBoost and XGBoost approaches," Expert Systems with Applications, vol. 241, p. 122686, 2024.

N. A. Kumar, R. Daniel, and P. K. Pasam, "A novel electrical load forecasting model using a deep learning approach," in The Internet of Energy, Apple Academic Press, pp. 67-89, 2024.

H. Habbak, M. Mahmoud, K. Metwally, M. M. Fouda, M. I. Ibrahem, "Load Forecasting Techniques and Their Applications in Smart Grids." Energies, vol. 16, pp. 1480, 2023.

M. López-Santos, S. Díaz-García, X. García-Santiago, A. Ogando-Martínez, F. Echevarría-Camarero, G. Blázquez-Gil, P. CarrascoOrtega, "Deep Learning and transfer learning techniques applied to short-term load forecasting of data-poor buildings in local energy communities." Energy and Buildings, vol. 292, pp. 113164, 2023.

Y. K. Ahranjani, M. Beiraghi, and R. Ghanizadeh, "Short time load forecasting for Urmia city using the novel CNN-LTSM deep learning structure," Electrical Engineering, pp. 1-12, 2024.

M. M. Asiri, G. Aldehim, F. Alotaibi, M. M. Alnfiai, M. Assiri, and A. Mahmud, "Short-Term Load Forecasting in Smart Grids using Hybrid Deep Learning," IEEE Access, 2024.

Fan, G. F., Han, Y. Y., Li, J. W., Peng, L. L., Yeh, Y. H., & Hong, W. C. (2024). A hybrid model for deep learning short-term power load forecasting based on feature extraction statistics techniques. Expert Systems with Applications, 238, 122012,

Y. Eren and İ. Küçükdemiral, "A comprehensive review on deep learning approaches for short-term load forecasting," Renewable and Sustainable Energy Reviews, vol. 189, p. 114031, 2024.

NIST Framework 4.0, 2022. Roadmap for smart grid interoperability standards [online] National Institute of Standards and Technology. Available at: https://

M. Cordeiro-Costas, D. Villanueva, P. Eguía-Oller, M. Martínez-Comesaña, & S. Ramos, "Load forecasting with machine learning and deep learning methods." Applied Sciences, vol. 13, pp. 7933, 2023.

B. Mehdizadeh Khorrami, A. Soleimani, A. Pinnarelli, G. Brusco, and P. Vizza, "Forecasting heating and cooling loads in residential buildings using machine learning: A comparative study of techniques and influential indicators," Asian Journal of Civil Engineering, vol. 25, no. 2, pp. 1163-1177, 2024.