



Research article

Artificial Intelligence based accurately load forecasting system to forecast short and medium-term load demands

Faisal Mehmood Butt^{1,2}, Lal Hussain^{3,4,*}, Anzar Mahmood¹ and Kashif Javed Lone³

¹ Department of Electrical Engineering, Mirpur University of Science & Technology, Mirpur 10250, Azad Kashmir, Pakistan

² Department of Electrical Engineering, University of Azad Jammu and Kashmir, Chehla Campus, Muzaffarabad 13100, Azad Kashmir, Pakistan

³ Department of Computer Science & IT, University of Azad Jammu and Kashmir, King Abdullah Campus, Muzaffarabad 13100, Azad Kashmir, Pakistan

⁴ Department of Computer Science & IT, University of Azad Jammu and Kashmir, Neelum Campus, Athmuqam 13230, Azad Kashmir, Pakistan

* **Correspondence:** Email: lall_hussain2008@live.com.

Abstract: An efficient management and better scheduling by the power companies are of great significance for accurate electrical load forecasting. There exists a high level of uncertainties in the load time series, which is challenging to make the accurate short-term load forecast (STLF), medium-term load forecast (MTLF), and long-term load forecast (LTLF). To extract the local trends and to capture the same patterns of short, and medium forecasting time series, we proposed long short-term memory (LSTM), Multilayer perceptron, and convolutional neural network (CNN) to learn the relationship in the time series. These models are proposed to improve the forecasting accuracy. The models were tested based on the real-world case by conducting detailed experiments to validate their stability and practicality. The performance was measured in terms of squared error, Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE). To predict the next 24 hours ahead load forecasting, the lowest prediction error was obtained using LSTM with R2 (0.5160), MLP with MAPE (4.97), MAE (104.33) and RMSE (133.92). To predict the next 72 hours ahead of load forecasting, the lowest prediction error was obtained using LSTM with R2 (0.7153), MPL with MAPE (7.04), MAE (125.92), RMSE (188.33). Likewise, to predict the next one week ahead load forecasting, the lowest error was obtained using CNN with R2 (0.7616), MLP with MAPE (6.162), MAE (103.156), RMSE (150.81). Moreover, to predict the next one-month load forecasting, the lowest prediction error was obtained using CNN with R2 (0.820), MLP with MAPE (5.18), LSTM with MAE (75.12) and RMSE (109.197). The results reveal that proposed methods achieved better and stable performance for predicting the short, and medium-term load forecasting. The findings of the STLF indicate that the proposed model can be better implemented for local system planning and dispatch, while it will be more efficient for MTLF in better scheduling and maintenance operations.

Keywords: Short-term load forecast; long short-term memory networks; convolutional neural networks; deep neural networks; artificial intelligence

1. Introduction

In the presence of load forecasting service, the thing which is important and achievable to meet the upcoming load demands so that power can be obviated to happen. The occurrence of faults due to unnecessary loading, other electrical failures such as blackout etc. can be refrained if properly required precautionary measures to be taken to manage the load demands. However, load demands are kept on continually increasing day by day. It fluctuates a lot on an hourly basis [1], so short-term load demands are highly desired to be automated for proper forecasting. The load forecasting is basically a stochastic problem instead of being random. One cannot find ‘certain’ in this phenomenon, because forecasters have to deal with stochasticity and haphazardness. The outcome of this process is expected to be in complicated form, for instance, a forecast extending below this or that context, a prediction level, and a little quantile of interest. Forecasting is considered crucial and high priority in the power sector. The previously available couple of global power forecasting fixtures, known as GEFCom2012 as well as GEFCom2014, grabbed the attention of many academia data scientists as well as various industries along with resolved numerous forecasting hurdles faced within the power sectors including probabilistic forecasting [2] and hierarchical forecasting [3].

There are three approaches most widely used in load forecasting such as shorter-term load forecasting (STLF), medium-term load forecasting (MTLF) and long-term load forecasting (LTLF) [1]. The STLF has been explored extensively as well as modeled at the generation capacity, distribution strategy and in transmission process. Mostly, the problem arises at the distribution stage because the distribution utilities are linked with the generation capacity and its scheduling and system peaking. These utilities do not have direct relation to short term (ST) load demand of customers himself. The ST load forecasting becomes a central problem for the functioning and transmission of energy networks to interrupt crucial outcomes related to failures in flash and energy systems. That is imperative towards the economic strengths of energy networks and the essence of sending off along with assembling start-up-shutting down schemes that play a crucial part in the automatically controlled power networks. Accurately precise power load forecast does not merely aid consumers to go for far suitable power utilization plan of action, but also decreases electric cost expenditure, meanwhile enhancing equipment utilization and hence minimizing the cost of production and upgrading the economic welfare. It is also conducive to optimize the reversion of the power system, enhancing the efficiency of power supply, and attaining the goal of energy conservation along with emission reduction.

Proper and efficient electric load forecasting can save electricity and make distribution easier and efficient by utilizing minimum energy resources. The cost of the electricity may increase if electric load forecasting is not accurate. Furthermore, when load forecasting is not proper and accurate then it may waste energy resources. Contrary to that, if load forecasting is done properly and accurately, it will help to manage the future electric load demands by efficiently load planning [4].

1.1. Factors affecting Electrical Load

Mostly, in a large power system, the usage schemes of tools might show distinction from each other and a single electrical device being randomly used. There are large fluctuations in discrete

loads, but on summing up these loads into bigger facility load, emerging pattern could be predicted arithmetically [5]. Here are the following factors which can affect electric load including weather, time factors, economic factors, and random effects.

1.1.1. Weather

Weather is one of the most influencing factors from the above-listed factors. Weather details become extremely essential elements in the load forecast scenario. Usually, load forecasting models are built and tested, taking into considerations the actual reading of weather. Nevertheless, outline operational models prerequisite maneuvering weather forecasts, along with related forecast errors and miscalculations. The weather forecast obviously escort to the degradation and breakdown in the model performance. Weather forecasting, for instance, the speed of wind, climatic temperature along with cloud covering performs a massive role in STLF by modifying the graph. The characteristics of these elements are mirrored in the load requirements even though some are influenced far more than others.

Weather have significant effects on ST electrical load forecasting of power network [6]. Weather sensitive loads like heating, air conditioning devices and ventilating have a great effects on small-scale industrial energy structures. Weather elements that could influence hourly power load forecasting are barometric pressure, precipitation, speed of wind, solar irradiance along with humidity. During maximum humidness days, the cooling apparatus can function for prolonged duty patterns to get rid of unneeded condensation from the conditioned air. Precipitation holds the capability to minimize the temperature of air that will result in the reduction of cooling load [5].

1.1.2. Time Factors

There are further three factors that can influence the electric loads: holidays, day to day weekly cycle, and seasonal outputs. The day to day power load forecast for the off day is almost the same as that of what is found on weekends. During holidays, the vastness of power loads is lesser as compared to working days. Day to day weekly cycle is power load forecast schemes which are periodical all through every day and week. While seasoning outcomes describe for long-term modifications in the enhancement and patterns of weather [7].

1.1.3. Economic Factors

Economic factors depend upon the investment that facilitates the basic structure by establishing the labs and new constructions that would add load to the power system.

1.1.4. Random Effects

All other random conflicts except the three earlier mentioned are included in random effects which can disturb energy load depiction. These disturbances include essential loads that have not plan and the absence of employees which makes the prediction difficult [5].

Load forecasting has gained more importance in smart energy management systems in recent few years. The number of users continually increasing for load forecasting on daily basis. Mostly, STLF is utilized to manage load forwarding and to control energy transfer schedule for thirty minutes to the entire day. Therefore, any betterment gained around precision and management of STLF escorts to reduce expenses of electrical management system along with the enhance efficiency of the energy network [8].

According to Gross and Galiana [5], STLF plays an important part in the creation of feasible, secure, dependable, economic as well as consistent working techniques for the electrical management network. STLF provides information about the latest prediction of weather, recent load forecasting strategy, and random behavior of the power system [8]. The load forecast is getting more interesting and attracting the attention of researchers in the current few years because of its importance and increasing trend in microgrids, smart grids, and renewable sources of energy. Different strategies have been implemented for load forecast as well as management, such as auto-regressive integrated moving average models, seasonal auto-regressive, fractionally integrated moving average, regression, Kalman filtering etc. Al-AI-Hamadi and Soliman [9] presented a load management model to meet the time-varying demand of users on an hourly period. There was another approach called the Kalman filter that was implemented to calculate the optimal load forecast on an hourly basis [8]. Song et al. [10] introduced another technique based on the fuzzy linear regression method built on the load data.

In recent years, numerous other techniques were applied to enhance the accuracy and efficiency of load forecast management of power systems, such as fuzzy logic and artificial intelligence (AI). These effective solutions built projected on the technologies of AI to solve load forecasting issues. AI-based intelligent systems have gained more importance, becoming more widespread nowadays and being developed on large scale. Because of their explanation capability, flexibility, and symbolic reasoning, they are being deployed worldwide in several applications. Ranaweera et al. presented another technique that is based on fuzzy rules developed by using a learning type algorithm to incorporate load management and historical weather data [11]. He et al. [12] proposed a novel method for the quantification of unpredictability stringed to power load and to obtain future load information, after that neural structure has been used in order to construct quantile regression framework in the development of probabilistic techniques for forecasting.

Recently, artificial intelligence (AI) based methods have been utilized for loads forecasting including smart grid and buildings [13], next day load demands [14], load forecasting in distributed systems [15] and autoregressive ANN with exogenous vector inputs [16] etc. Moreover, other AI base methods including fuzzy logistic techniques [17] specialist networks structures [18], Bayesian neural system [19] and support vector apparatus [20] are widely used in handling number of forecasting issues. Even though extensive research has been carried out, an error-free as well as precise STLF continue in becoming a challenge because it is having a load data which is non-stationary in addition to carrying durable dependencies related to forecasting horizons. For this reason, the long-short term memory (LSTM) is applied [21], which is a peculiar kind of recurrent neural network (RNN) structure [22] to resolve the STLF issue. LSTM functions effectively around long-time horizon forecasting as compared to the rest of the artificial intelligence techniques which is routed through the previous load statistics which governed outcome as well as connections across the series of time.

The researchers [23] suggested a hybrid technique established on wavelet transform as well Bayesian neural networks (BNN) to get the load characteristics for training of the BNN model. In this technique, an average sum of BNN outputs is used for predicting load for a specific day. Fan et al. [24] introduced another hybrid design established on the junction of a combination of Bi-Square Kernel (BSK) regression framework and phase space reconstruction (PSR) method. In this approach, load statistics were regenerated via using the PSR method to obtain the developmental modes of the documented load statistics to enhance prediction reliability [24]. The researchers [25] proposed another fuzzy logic controller-based hourly for load forecasting depending upon different conditional variables, e.g., random disturbances, load historical statistics, time as well as climate etc. Finally, the

developed hybrid model was successfully established by employing a combination of evolutionary algorithms and neural networks.

In another study, Metaxiotis et al. [26] provides a comparison of models based on CNNs and AI, where CNN models show a distinctive performance in load forecasting prediction. Fukushima K [27] presented CNNs in very simple form. LeCun et al proposed the current form on CNNs with more advanced concepts, there have been many further extensions with improvements, such as batch normalization and max-pooling layer [28] etc. More specifically, some strategies are devised for CNN to facilitate structures, for instance, the addition of pooling surface in the design [29]. In short, CNN is being considered as a potential candidate for load forecasting implementations, so far as the control of over-fitting is concerned. On the other side, Fuzzy time series (FTS) were employed in pattern acquisition-based techniques within numerous implementations which include load forecasting. In 2005, Yu [30] proposed a novel kind of weighted FTS for the forecast of the stock market. In 2009, Wang et al. proposed another technique using FTS that was implemented in stock index forecasting and temperature forecasting [31]. In summary, many other relevant studies show that FTS has been most widely used in load forecasting management systems by many other researchers such as hybrid dynamic and fuzzy time series model for mid-term power load forecasting [32], a new linguistic out-sample approach of fuzzy time series for daily load forecasting [33] and imperialist competitive algorithm combined with refined high-order weighted fuzzy time series (RHWFTS–ICA) for short term load forecasting [34].

The AI-based machine learning and deep convolutional neural networks methods have also successfully been used in many other areas of complex physiological signals and image processing problems. The applications of AI based methods include seizure detection using entropy-based methods [35] and machine learning methods [36], congestive heart failure by extracting multimodal features and employing machine learning techniques [37], Alzheimer detection via machine learning [38], brain tumor detection based on hybrid features and machine learning techniques [39], arrhythmia detection [40], lung cancer detection by extracting refined fuzzy entropy [41], and prostate cancer based on deep learning [42] and machine learning [43] methods.

Deep learning methods such as CNN mostly improve the prediction performance using big data and has improved the traditional computer vision tasks such as image classification etc. Recently, CNN is used for both imaging and non-imaging data. There are many applications of 1D-CNN for time series data including electricity load forecasting [44], electricity load forecasting for each day of week [45], hydrological time-series prediction [46], short-term load forecasting of Romanina power system [47] and short-term wind power forecasting [48] etc.

Figure 1 shows the schematic diagram of the electric load forecasting system for Short-term and medium-term load forecasting. Short-term load forecasting (STLF) is used for the planning of the power systems ranging from 01 hours up to one week. In this case, we computed the STLF for the next 24 hours, 72 hours, and one week. The medium-term load forecasting (MTLF) is used to plan maintenance etc. ranging from one day to a few months. In this case, we computed the next load forecast of one day (24 hours), 72 hours, one week, two weeks, and one month. After applying the data preprocessing, we optimized and initialized the robust neural network and deep learning models such as multilayer perceptron (MPL), LSTM and CNN. The performance was computed on the test set based on the standard performance error metrics such as R-squared, MAPE, MAE, MSE, and RMSE. Finally, the load forecasting for STLF and MTLF ahead was computed and performance was evaluated in terms of errors for actual and predicted load demands.

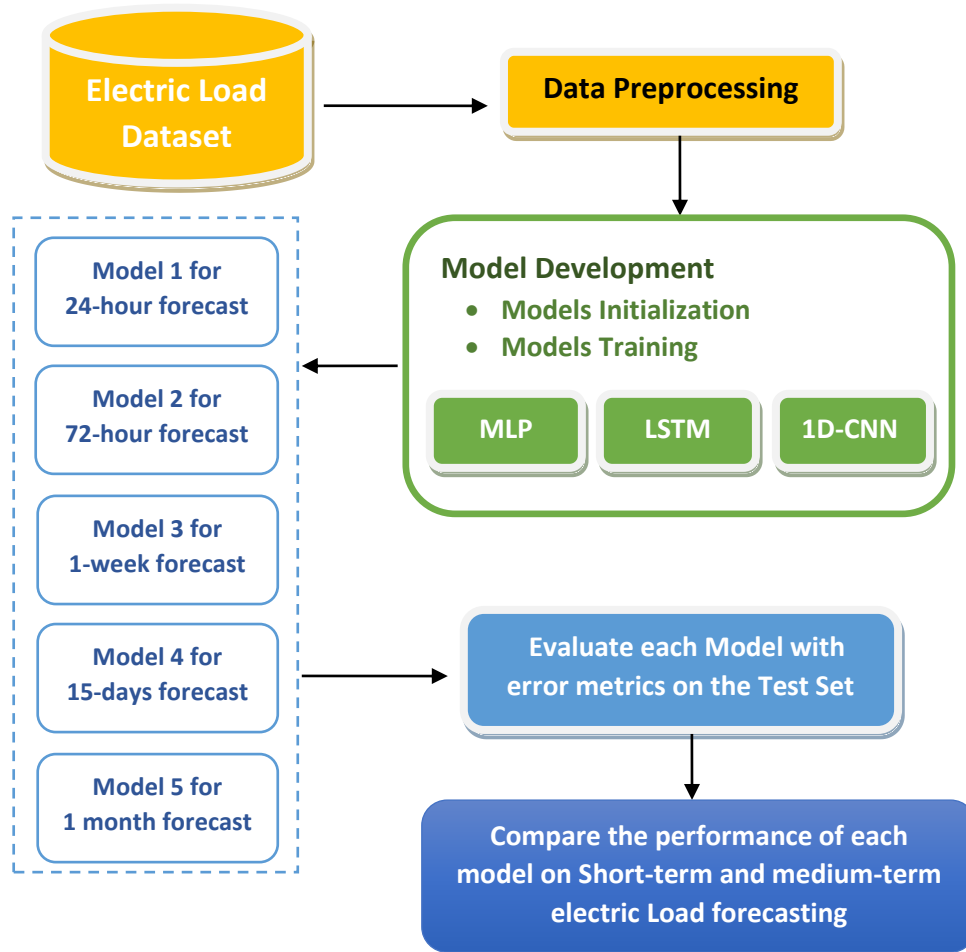


Figure 1. Schematic Diagram of Electric Load forecasting System for Short-term and medium-term load forecasting.

In this study, we optimized and employed robust machine learning and deep learning-based methods to predict the load forecasting demands for STLF and MTLF. For multilayer perceptron (MLP), we optimized the function by changing the number of neurons in the hidden layer, this number must be high enough to model the problem and not too much high to avoid overfitting. The iterative backpropagation algorithm was used for this purpose. Moreover, to solve the problem of minimizing the cost function concerning connection weights, the gradient descent algorithm is used in conjunction with the backpropagation algorithm.

1.2. Fine tuning of parameters

Following methods were used to fine-tune the neural network parameters:

1.2.1. LSTM

We created an LSTM model with one LSTM layer of 48 neurons and "Relu" as an activation function. We also added a dense layers which contains 24 neurons and the last layer, which also acts as the output layer, contains 1 neuron. Finally, we compiled our model using optimizer = "ADAM" and train it for 100 epochs with a batch size of 24.

1.2.2. Conv1D

Firstly, for Conv1D we defined 48 filters and a kernel size of 2 with "RELU" as an activation function. In order to reduce the complexity of the output and prevent over fitting of the data we used Max pooling layer after a CNN layer with size of 2. Moreover, a dense layers were added contained 24 neurons and the final output layer, contains 1 neuron. The model was compiled with "ADAM" optimizer and then fit for 100 epochs, and a batch size of 24 samples is used.

1.2.3. MLP

The used model contained a single layer with 48 neuron and "RELU" as an activation function. Similar to LSTM and Conv1D a dense layers were also added with 24 neurons and a final layer for output contain a single neuron respectively. Lastly, the model was fit using the efficient ADAM optimization algorithm and the mean squared error loss function for 100 epochs

2. Material and methods

2.1. Dataset

The data was taken from Al-Khwarizmi Institute of Technology, The University of Punjab, Lahore, from one of its project regarding the electricity hourly load demands of the complete year 2008 and July to December 2009 of one grid as used in [4]. The data was taken from a feeder to fulfill the maximum load requirements for this study. The load time series was generated for 24 hours ahead, 72 hours ahead, one week ahead, and one month ahead for predicting the short-term and medium-term load demands.

2.2. Prediction Models

2.2.1. Multilayer Perceptron (MLP)

Paul Werbos developed the MLP in 1974, which generalizes simple perception in the non-linear approach by using the logistics function

$$F(x) = \tanh(x) \quad (2.1)$$

It has become one of the most popular neural networks conceived for supervised learning. The Multilayer Perceptron constitutes of 3 layers, an input layer, an intermediate layer as well output layer which can be formed by at least one layer, the information is transmitted in one direction, emerging out of the input layer towards the output layer. By an adjustment iteration set comparing outputs and inputs, the MLP adjusts the weights of neural connections; to find an optimal weight structure through the gradient backpropagation method. The network generally converges to a state where the calculation error is low [49].

Local learning procedure is used to train MLP [50]. For this purpose, few specimens are picked from the neighbourhood of any selected position \mathbf{x}^* to train MLP. The neighborhood means the group of k-nearest neighbors in the scheme of training set Φ of the selected point. The framework is trained to learn and positioned perfectly to the target function about the selected point \mathbf{x}^* be the proficient for this query scheme only. Local complication is lesser as compared to the global complexity of the target function. Therefore, elementary MLP design with least concealed neurons can be used that can learn fast. In [50] researcher reported that using a single-neuron carrying sigmoid acceleration function delivered better outcomes than other systems working along with numerous neurons in the concealed surface. MLP can be implemented with

a backpropagation algorithm and is considered as one of the most popular and commonly used networks. The backpropagation training algorithm is a monitored acquisition algorithm and has so far been implemented on a large scale for the prediction problems such as Pan evaporation (EP) prediction [51] and prediction of annual gross electricity demands based on socio economic indicators and climate conditions [52]. The figure of closest neighbours (that is learning position) becomes the only hyper-parameter which is modulated in the local learning programming (LLP) method. The MLP absorbs utilizing the Levenberg-Marquardt algorithm along with Bayesian regularization [53], which reduces the amalgamation of squared miscalculations along with net weights to keep away from overfitting. The optimization of MLP is figured out in the algorithm given below:

Figure 2 reflects the general architecture and working of the MLP algorithm, which contain the input time series of different normalized loads demands as actual time series, hidden neurons and activation functions, learning and finally error metrics are computed to predicted the difference between actual and predicted load demands.

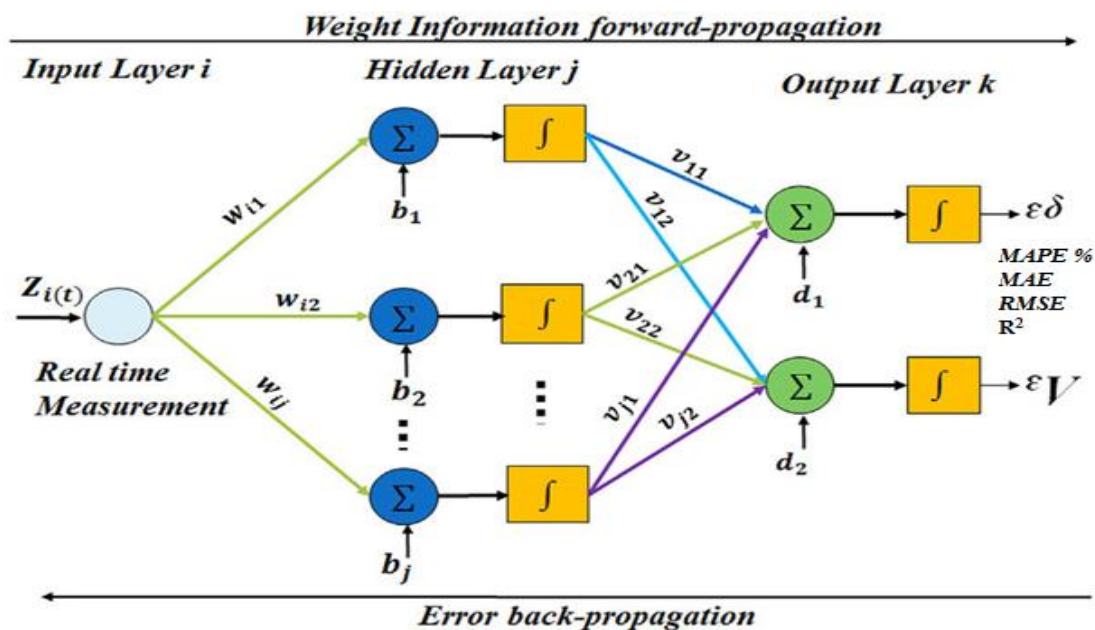


Figure 2. Multilayer perceptron (MPL) architecture.

Multilayer Perceptron Algorithm:

- 1: Start
- 2: Find the set Θ of l nearest neighbours of query pattern \mathbf{x}^* in Φ .
- 3: Do for each $i \in \Xi$
 - 3.1: Create the set $\Phi' = \Phi / (x_i, y_i)$
 - 3.2: Do for $k = k_{min}$ to k_{max}
 - 3.2.1: Find the set Φ'' of k nearest neighbors of x_i in Φ'
 - 3.2.2: Learn MLP on Φ''
 - 3.2.3: Test MLP on x_i
 - 3.3: Select the best value of k for x_i
- 4: Calculate the optimal value of nearest neighbours k_{opt} as the mean value of k selected in 3.3
- 5: Find the set Θ' of k_{opt} nearest neighbors of \mathbf{x}^* in Φ
- 6: Learn MLP on Φ'
- 7: Test MLP on \mathbf{x}^*

8: Stop

2.2.2. Deep Learning Methods

Deep learning is a sub-domain of Artificial intelligence (AI) works in similar strategies as machine learning (ML) and artificial neural networks (ANN) perform. The AI is functioning in similar pattern to human. Like a human brain, ANN take information to process this information using a group of neurons which form layers. These neurons transfer information to other neurons where some information is sent back to the previous layer and finally processed information is sent to the output layer in the form of classification or regression. Deep learning extract features and data automatically and methods improve the prediction of complex problems [54].

2.2.2.1. Long Short – Term Memory (LSTM)

LSTM is an artificial recurrent neural network (RNN) architecture used in the field of deep learning [55]. LSTM is considered a popular and expert model used in the forecasting of time series that can efficiently deal with both long term and short-term data dependencies. LSTM was designed and motivated to get control of the issues related to disappearing gradients in RNN architecture, specifically while dealing with the long-term data dependency, which lead to the short and LTM neural network. Basically, the LSTM framework adds the input barrier, forgetting gate, and output gateway towards the neurons in regressive neural network architecture. This newly developed approach can efficiently manage the problem of vanishing gradient [56]. This addition gets LSTM structure more appropriate for long-termed data dependency problems. LSTM has the ability of learning long-short term memory from any given input sequence and that is why it is being widely employed in the time series prediction [57].

LSTM methods from RNNs have been used in many applications such as speech and language modelling, speech recognition and classification of neurocognitive performance by Greff et al., in [58]. Moreover, P. Nagabushanam et al. [59] employed LSTM for classification improvement of EEG signals. Senyurek et al., [60] employed LSTM for puffs detection in smoking by obtaining temporal dynamics sensors. LSTM improved recognition of gait in a neuron degenerated disease as compared to old recurrent neural network (RNN) methods [59]. The machine learning algorithms faces the problem of gradient learning. LSTM on contrast solve this problem as it is based on the appropriate gradient-based learning algorithm and solve error backflow problems. Moreover, LSTM algorithm is also more appropriate when there is noise or incompressible input sequence without losing short time lag capabilities. LSTM is also more efficient in fast and adaptive learning than other machine learning algorithms. It is capable to solve very long-time lag tasks and complex problems which cannot be solved using conventional machine learning problems.

The hidden layers of LSTM are linear, but self-loop memory blocks allow the gradient to flow through large sequences. LSTM comprised of recurrent blocks termed as memory blocks, each block contain recurrent memory cells and three multiplicative units namely input, output and forget gates [61]. These cells allow memory blocks to store and access information for a long time to solve the vanishing problems [62]. LSTM was originally comprised on input and output gate, the forget gate was included by [63] to improve the functionality by resetting memory cells. Figure 3 reflects the general architecture of the LSTM model.

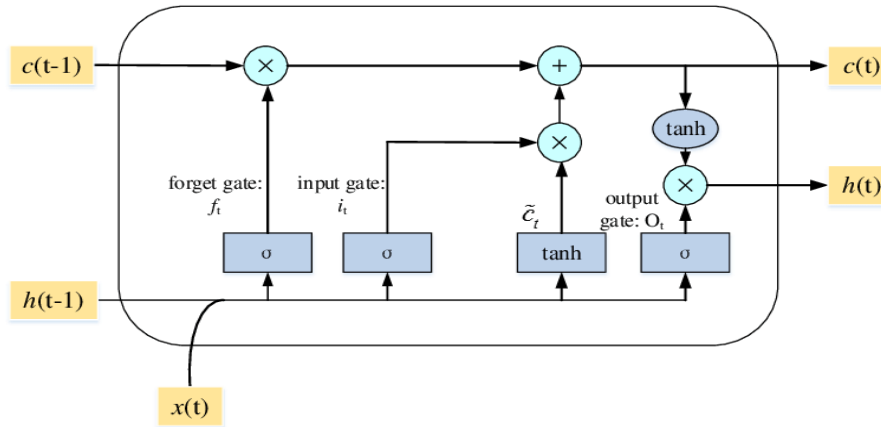


Figure 3. Architecture of LSTM Model.

Memory cell of LSTM is considered as its major innovation, which is used as a hoarder to store the state particulars. In the initial step, the forget gate is used to get rid of the unnecessary information. After that, a sigmoid operation is applied to measure the accelerate the forget state f_t .

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2.2)$$

The second step is used to know which new data is required to get saved within the cell condition. Another sigmoid layer, known as the “input gate layer”, is used to get the updated information. Then, a \tanh function is used to create a vector \tilde{c}_t of novel values that ought to be updated embarked on upcoming state.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2.3)$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (2.4)$$

In the second step, the old cell state c_{t-1} is refurbished to new cell state c_t . To delete the information from the old cell we can multiply c_{t-1} by f_t . Then add $i_t * \tilde{c}_t$. The new candidate values for updating represented by

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \quad (2.5)$$

In the last step, the output is needed to be decided. The step consists of a couple of further steps: the sigmoid function is used as an output barrier to strain the cell state. Further, the obtained cell state is passed across $\tanh(\cdot)$, the obtained output o_t is multiplied for the calculation of desired information.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (2.6)$$

$$h_t = o_t * \tanh(c_t) \quad (2.7)$$

In equations (5)–(10), W_i, W_f, W_c, W_o shows the weight matrices and the bias vectors are represented by the b_i, b_f, b_c, b_o vectors.

2.2.2.2. Convolutional Neural Network (CNN)

CNN is also a class of ANN which become dominant in many applications including computer vision, signal and image processing, electricity load forecasting etc. CNN is designed to automatically and adaptively learn spatial hierarchies of features through backpropagation by using multiple building blocks, such as convolution layers, pooling layers, and fully connected layers. In the field of

computer vision, artificial neural networks improved the results of various computer vision problems as compare to several classical techniques by utilizing and convolutional neural network is the most common type of artificial neural networks used for computer vision tasks such as image.

The scope of CNN is very wider and it has numerous applications, for example, in the health care domain as saving lives is a top priority in healthcare so CNN can be used for health risk assessment. In radiology CNN can be used for efficient classification of certain diseases, Segmentation of organs or anatomical structures, to detect abnormalities, by computed tomography (CT) or MRI images. Drug discovery is another major healthcare field with the extensive use of CNNs. There are many other applications of CNN in computer vision such as object location, object detection & image fragmentation [64], recognition of face [65], video grouping, in-depth assessment as well as image inscribing [66] classification. It is worth mentioning here, that the scope of application of CNN is not only limited to image processing tasks. CNN has many other application as well such as in natural language processing [67], day-ahead building-level load forecasts [68], anomaly detection [69], time series stream forecasting [70].

At that point, this thought acquired by specialists in modern zones there this innovation gets well known. In 1998, Lecun et al. [71] proposed plan which is improved structure than [72] engineering of the present Convolution Neural Network resembles LeCun plan. Convolution Neural Network gets well known by winning the ILSVRC2012 rivalry [73].

In the current years, many other advanced strategies have been implementer serving the purpose of load forecasting, for example, Artificial Intelligence as well as fuzzy logic. Savvy arrangements, considering AI advancements, to understand short term load forecasting are turning out to be increasingly more far-reaching these days. AI-based frameworks are created and sent out globally in numerous implementations, essentially due to their emblematic insightfulness, adaptability, and explain-country capacities. For instance, Ranaweera et al. suggested a technique that utilized fuzzy guidelines to consolidate recorded climate and load information. The historical data is a source from which these fuzzy principles were acquired utilizing a learning-type calculation [11]. He et al. [12] introduced architecture to evaluate the vulnerability associated with electrical load as well as gaining far more data upcoming subsequent load, afterward neural system has been utilized in order to remake quantile regression archetype aimed to develop probabilistically done forecasting technique. In one of the studies, a hybrid forecast technique for the wavelet change, neural system, and a set-up of evolutionary algorithm was suggested [74] and utilized for STLF. In a comparative methodology, Ghofrani et al. [23] presented a mix of Bayesian Neural Network (BNN) and a wavelet transform (WT) to create definite load attributes for BNN preparation. Viewing the design, a weighted total from the Bayesian Neural Network yields were made in application to anticipate the load desired for a particular day. However, in one of the other hybrid models, an STLF design suggested by [24] conjoins Phase Space Reproduction calculation with bi-square kernel (BSK) design of regression. In this particular structure, load information can be imitated by a phase space reconstruction algorithm to determine developmental patterns around historical loads details as well as data along with the implanted significance highlights data to expand the unwavering quality of the forecast. The BSK design, then again, link to the spatial structures between the points of regressions and the point located at their neighbour to get the guidelines of rotatory rules and unsettling influence in each measurement. The suggested model including multi-dimensional relapse was effectively established and utilized for STLF [24]. Mamlook et al. [25] utilized fuzzy rationale controller-based hourly to predict the effect of different restrictive boundaries, e.g., time, atmosphere, historical load sum of data and arbitrary unsettling influences, overload forecasting regarding fuzzy sets through the age

procedure. In this examination, WT deteriorated the time arrangement into its segments, and afterward every part was forecasted by a mix of neural system along with evolutionary algorithm.

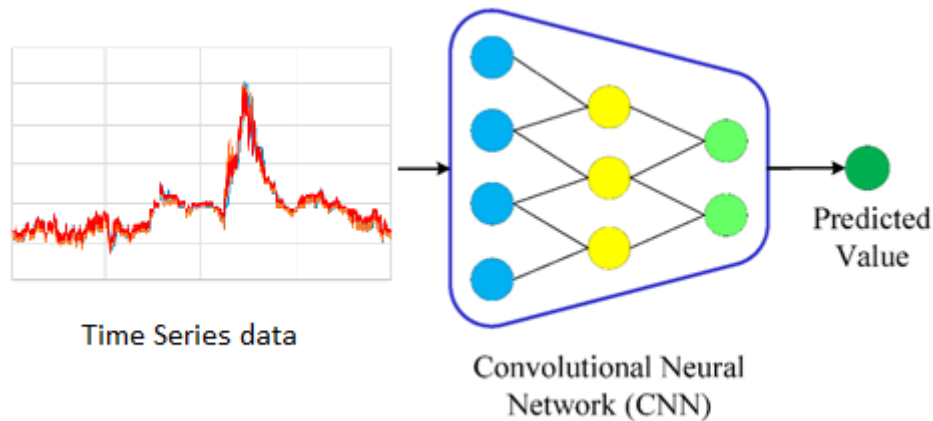


Figure 4. CNN Architecture for load forecasting.

Figure 4 indicates the architecture of the CNN model. A 1D load forecasting time series were used as input to the CNN model, it was processed in different layers of CNN models and finally, the output was in terms of error between actual load demand and predicted values.

2.3. Performance Evaluation Measures

The quality of the predictor was examined by quantitatively measuring the accuracy in terms of root mean squared error (RMSE), coefficient of determination (R^2), mean square error (MSE), and mean absolute error (MAE). The following renowned error prediction metrics detailed in [75] are used:

2.3.1. Root Mean Squared Error (RMSE)

To examine the quality of a predictor, we need metrics to quantitatively measure its accuracy. In the current study, a quantity called RMSE was introduced for such a purpose, as defined by:

$$RMSE = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.8)$$

Where x_i and y_i denote the measured and predicted values of the i -th sample, and 'n' denote the total number of samples of the training dataset. The smaller value of RMSE denotes the better set of selected descriptors.

2.3.2. Co-efficient of determination (R^2)

R^2 can be computed using the following function:

$$R^2 = \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - \bar{y})^2} \quad (2.9)$$

Here \bar{y} denote the average values of all the samples.

2.3.3. Mean Square Error (MSE)

MSE can be mathematically computed as follow

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2 \quad (2.10)$$

The MSE of an estimator measures the average of the squares of errors or deviations. MSE also denotes the second moment of error that incorporates both variance and bias of an estimator.

2.3.4. Mean Absolute Error (MAE)

MAE is the measure of difference between two consecutive variables, for example, variable y and x denote the predicted and observed values, then MAE can be calculated as:

$$\text{MAE} = \sum_{i=1}^n \frac{(y_i - x_i)}{n} \quad (2.11)$$

2.3.5. Mean Absolute Percentage Error (MAPE)

The MAPE is computed using the following formula.

$$\text{MAPE} = \frac{100}{n} \sum_{i=1}^n \frac{(y_i - x_i)}{n} \quad (2.12)$$

3. Results

In this study, we applied machine learning and deep learning methods such as MLP, LSTM, and CNN on load forecasting time series data from January 01, 2008 to December 31, 2009. The performance was evaluated in terms of R², MAPE, MAE, MSE, and RMSE. We computed the next 24 hours, 72 hours, one week, and one-month forecasting using the proposed methods

The distance between actual and predicted values is computed. If the difference between the observed and predicted values is smaller and unbiased, it indicates that the model best fits the data. Statistically, the goodness of fitness is also measured using residual plots which can reveal unwanted residual patterns that indicate biased results more effectively than numbers. R-squared is a statistical measure that indicates how close the data are to be fitted. It is also called as coefficient of determination.

Table 1 reflects the prediction of the next 24 hours ahead load forecasting. Based on the R-squared error method, the LSTM gives the highest prediction with R² (0.5160) followed by CNN yield R² (0.5462) and MLP yield R² (0.6217). Based on MAPE, the highest next 24 hours ahead prediction was obtained using MLP with MAPE (4.97) followed by LSTM yield MAPE (5.17) and CNN yield MAPE (5.62). Accordingly, the prediction performance in terms of MAE, the highest 24hours ahead prediction was obtained using MPL yield MAE (104.33) followed by LSTM with MAE (109.2) and CNN with MAE (115.62). The highest prediction in terms of MSE was obtained using MLP with MSE (17936.03) followed by CNN with MSE (21515.55) and LSTM with MSE (22947.59). Likewise, the next 24 hours ahead load forecasting prediction was obtained using MLP with RMSE (133.92) followed by CNN with RMSE (146.68) and LSTM with RMSE (151.48).

Table 1. Prediction of the next 24 hours ahead load forecasting.

| Method | R ² | MAPE | MAE | MSE | RMSE |
|--------|----------------|------|--------|----------|--------|
| MLP | 0.6217 | 4.97 | 104.33 | 17936.03 | 133.92 |
| LSTM | 0.5160 | 5.17 | 109.2 | 22947.59 | 151.48 |
| CNN | 0.5462 | 5.62 | 115.62 | 21515.55 | 146.68 |

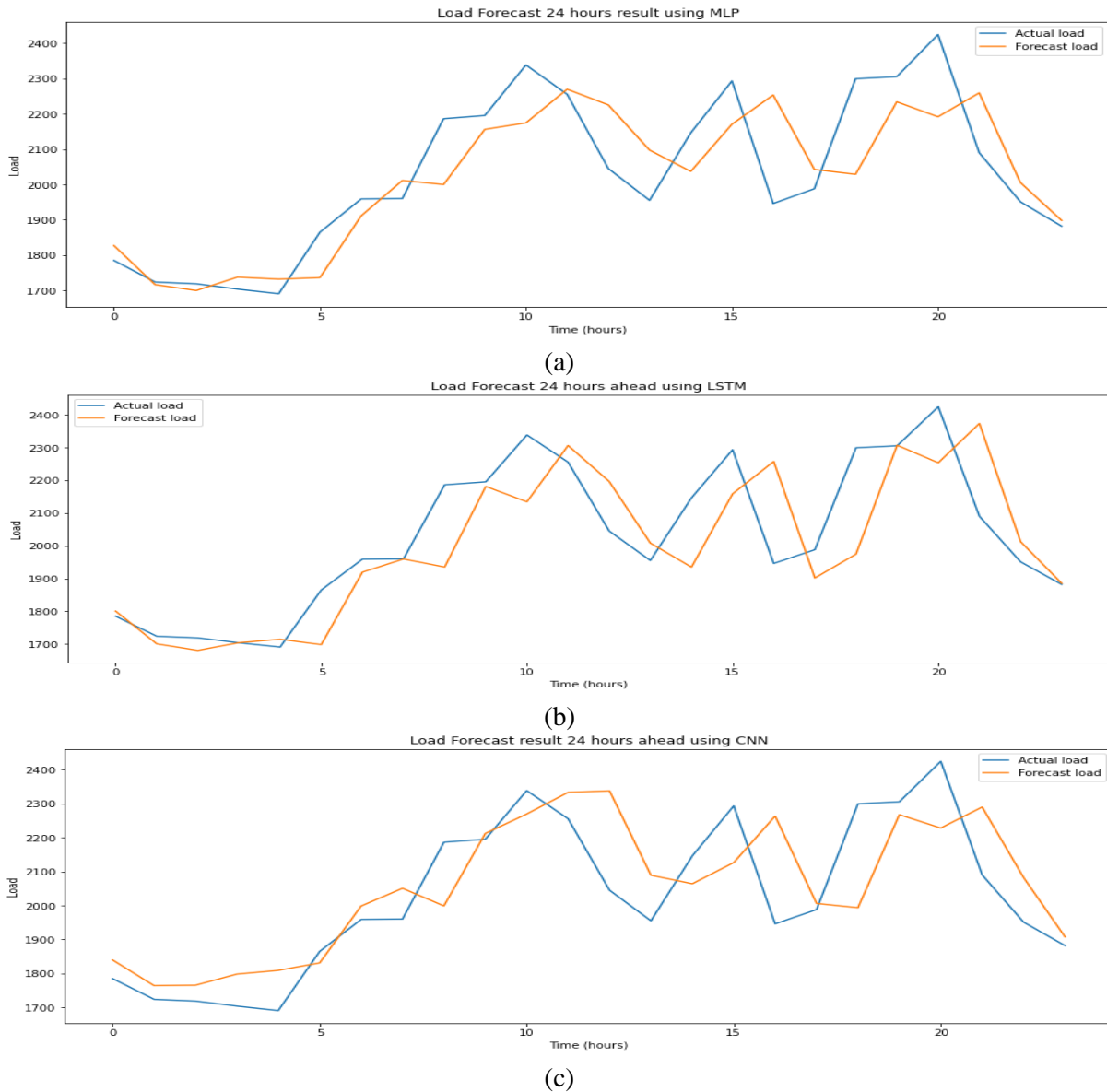
**Figure 5.** 24 hours ahead load forecasting using a) MLP, b) LSTM, c) CNN.

Figure 5 shows the results of 24-hours ahead load forecasts obtained by three different methods (MLP, LSTM, and CNN). From the extracted curves, the MLP is closest to the actual load curve followed by LSTM and CNN. The corresponding error values are obtained as reflected in Table 1.

Table 2 reflects the prediction of the next 72 hours ahead load forecasting. Based on the R-squared error method, the LSTM gives the highest prediction with R² (0.7153) followed by CNN yield R² (0.7176) and MLP yield R² (0.7588). Based on MAPE, the highest next 72 hours ahead prediction was obtained using MLP with MAPE (7.04) followed by LSTM yield MAPE (8.07) and

CNN yield MAPE (7.44). Accordingly, the prediction performance in terms of MAE, the highest 72 hours ahead prediction was obtained using MLP yield MAE (125.92) followed by CNN with MAE (140.21) and LSTM with MAE (144.84). The highest prediction in terms of MSE was obtained using MLP with MSE (35393.93) followed by CNN with MSE (41451.11) and LSTM with MSE (41786.91). Likewise, the next 72 hours ahead load forecasting prediction was obtained using MLP with RMSE (188.13) followed by CNN with RMSE (203.59) and LSTM with RMSE (204.42).

Table 2. Prediction of the next 72 hours ahead load forecasting.

| Method | R^2 | MAPE | MAE | MSE | RMSE |
|--------|--------|------|--------|----------|--------|
| MLP | 0.7588 | 7.04 | 125.92 | 35393.93 | 188.13 |
| LSTM | 0.7153 | 8.07 | 144.84 | 41786.91 | 204.42 |
| CNN | 0.7176 | 7.44 | 140.21 | 41451.11 | 203.59 |

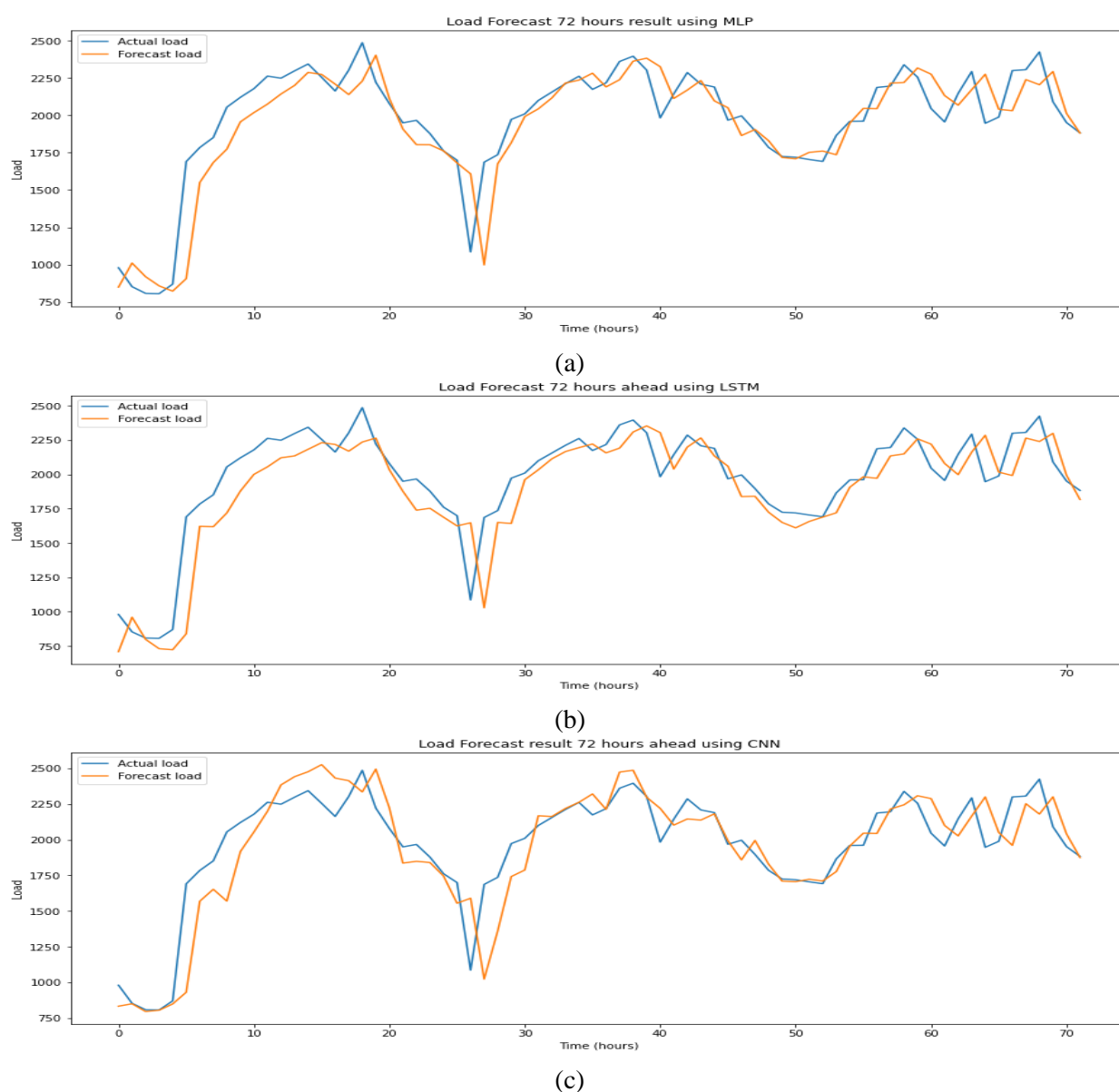


Figure 6. 72 hours ahead load forecasting using a) MLP, b) LSTM, c) CNN.

Figure 6 shows the results of 72-hours ahead load forecasts obtained by three different methods (MLP, LSTM, and CNN). From the extracted curves, it can be seen that the LSTM and CNN are closest to the actual load curve followed by MLP. The corresponding error values are obtained as reflected in Table 2.

Table 3. Prediction of the next one week ahead load forecasting.

| Method | R^2 | MAPE | MAE | MSE | RMSE |
|--------|--------|-------|---------|----------|--------|
| MLP | 0.8879 | 6.162 | 103.156 | 22746.21 | 150.81 |
| LSTM | 0.8814 | 6.74 | 107.13 | 24060.60 | 155.11 |
| CNN | 0.7616 | 9.79 | 158.27 | 48390.42 | 219.97 |

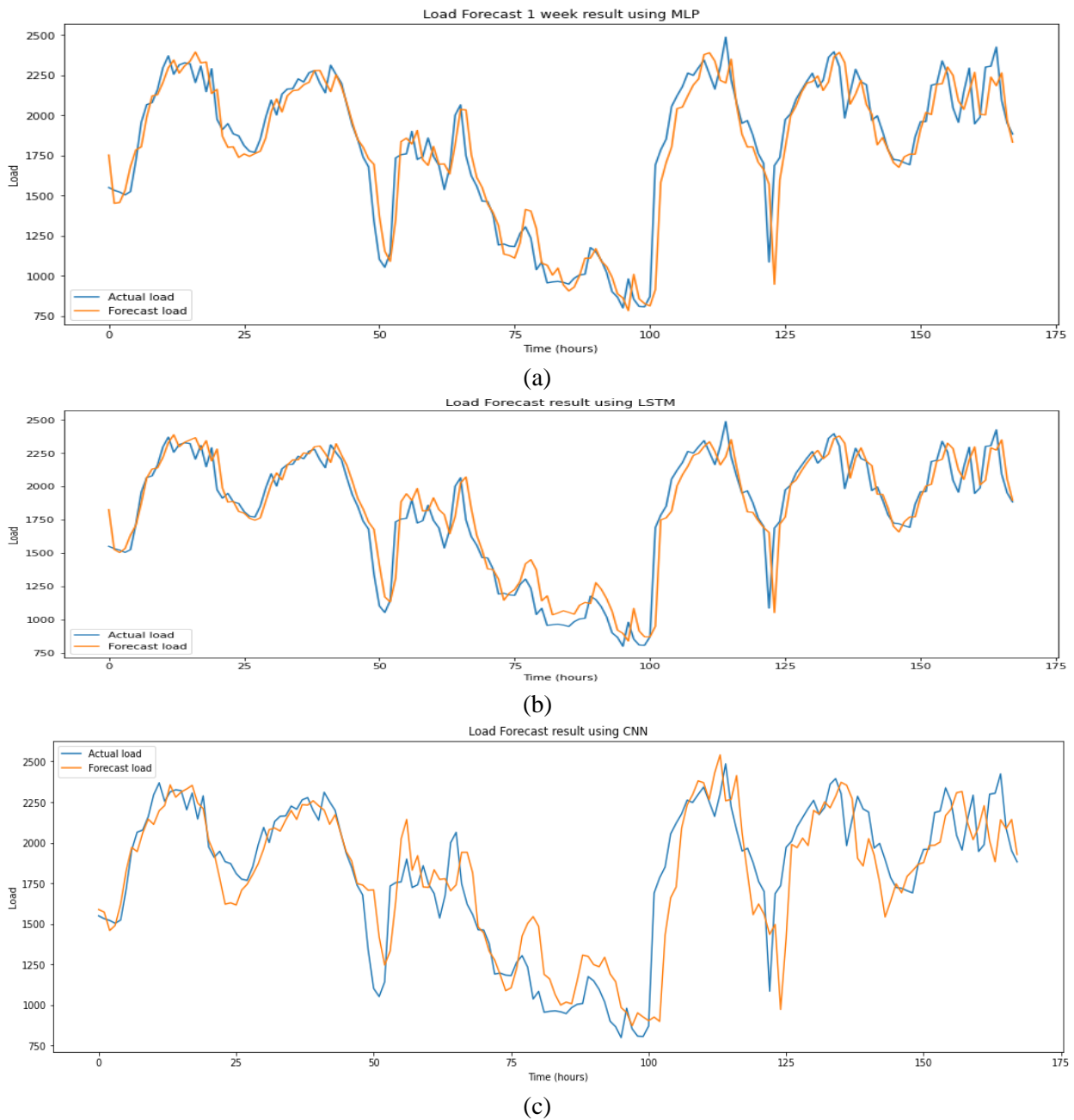


Figure 7. One week ahead load forecasting using a) MLP, b) LSTM, c) CNN.

Table 3 reflects the prediction of the next one week ahead load forecasting. Based on the R-squared error method, the CNN gives the highest prediction with R2 (0.7616) followed by) and LSTM yield R2 (0.8814) and MLP with R-square (0.8879). Based on MAPE, the highest next one-week ahead prediction was obtained using MLP with MAPE (6.162) followed by LSTM yield MAPE (6.74) and CNN yield MAPE (6.79). Accordingly, the prediction performance in terms of MAE, the highest one-week ahead prediction was obtained using MPL yield MAE (103.156) followed by LSTM with MAE (107.13) and CNN with MAE (158.27). The highest prediction in terms of MSE was obtained using MLP with MSE (22746.21) followed by LSTM with MSE (24060.60) and CNN with MSE (48390.42). Likewise, the next one week ahead load forecasting prediction was obtained using MLP with RMSE (150.81) followed by LSTM with RMSE (155.11) and LSTM with RMSE (219.97).

Figure 7 shows the results of one week ahead load forecasts obtained by three different methods (MLP, LSTM, and CNN). From the extracted curves, it can be seen that the LSTM and CNN are closest to the actual load curve followed by MLP. The corresponding error values are obtained as reflected in Table 3.

Table 4 reflects the prediction of the next 15 days ahead load forecasting. Based on the R-squared error method, the CNN gives the highest prediction with R2 (0.76) followed by) and MLP yield R2 (0.87) and LSTM with R-square (0.89). Based on MAPE, the highest next two-weeks ahead prediction was obtained using LSTM with MAPE (5.44) followed by MLP yield MAPE (5.54) and CNN yield MAPE (8.67). Accordingly, the prediction performance in terms of MAE, the highest two-weeks ahead prediction was obtained using LSTM yield MAE (93.67) followed by MLP with MAE (99.50) and CNN with MAE (141.89). The highest prediction in terms of MSE was obtained using LSTM with MSE (17395.14) followed by MLP with MSE (19963.230) and CNN with MSE (36938.71). Likewise, the next two-weeks ahead load forecasting prediction was obtained using LSTM with RMSE (131.89) followed by MLP with RMSE (141.29) and CNN with RMSE (192.19).

Table 3 reflects the prediction of the next one week ahead load forecasting. Based on the R-squared error method, the CNN gives the highest prediction with R2 (0.7616) followed by) and LSTM yield R2 (0.8814) and MLP with R-square (0.8879). Based on MAPE, the highest next one-week ahead prediction was obtained using MLP with MAPE (6.162) followed by LSTM yield MAPE (6.74) and CNN yield MAPE (6.79). Accordingly, the prediction performance in terms of MAE, the highest one-week ahead prediction was obtained using MPL yield MAE (103.156) followed by LSTM with MAE (107.13) and CNN with MAE (158.27). The highest prediction in terms of MSE was obtained using MLP with MSE (22746.21) followed by LSTM with MSE (24060.60) and CNN with MSE (48390.42). Likewise, the next one week ahead load forecasting prediction was obtained using MLP with RMSE (150.81) followed by LSTM with RMSE (155.11) and LSTM with RMSE (219.97).

Table 4. Prediction of the next 15 days ahead load forecasting.

| Method | R ² | MAPE | MAE | MSE | RMSE |
|--------|----------------|------|--------|----------|--------|
| MLP | 0.87 | 5.54 | 99.50 | 19963.23 | 141.29 |
| LSTM | 0.89 | 5.44 | 93.67 | 17395.14 | 131.89 |
| CNN | 0.76 | 8.67 | 141.89 | 36938.71 | 192.19 |

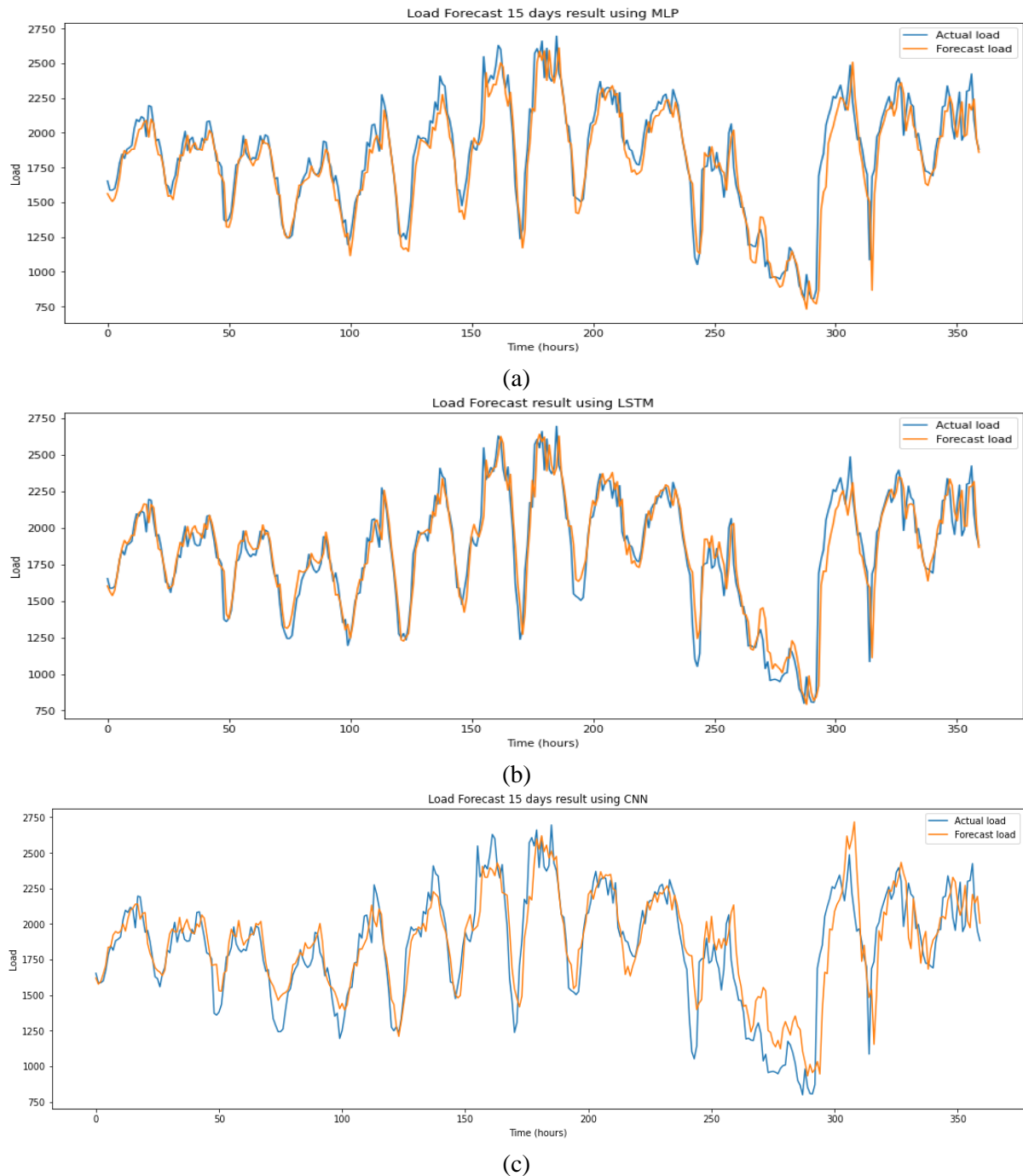


Figure 8. One two weeks ahead load forecasting using a) MLP, b) LSTM, c) CNN.

Figure 8 shows the results of two weeks ahead load forecasts obtained by three different methods (MLP, LSTM, and CNN). From the extracted curves, it can be seen that the LSTM and CNN are closest to the actual load curve followed by MLP. The corresponding error values are obtained as reflected in Table 4.

Table 5 reflects the prediction of the next one month ahead load forecasting. Based on the R-squared error method, the CNN gives the highest prediction with R^2 (0.90) followed by) and MLP yield R^2 (0.92) and MLP with R -square (0.92). Based on MAPE, the highest next one-month ahead prediction was obtained using LSTM with MAPE (4.38) followed by MLP yield MAPE (4.23) and CNN yield MAPE (5.1). Accordingly, the prediction performance in terms of MAE, the highest one-

month ahead prediction was obtained using LSTM yield MAE (75.12) followed by MLP with MAE (71.26) and CNN with MAE (84.95). The highest prediction in terms of MSE was obtained using LSTM with MSE (11258.74) followed by MLP with MSE (11017.13) and CNN with MSE (14584.13). Likewise, the next one month ahead load forecasting prediction was obtained using LSTM with RMSE (106.10) followed by MLP with RMSE (104.96) and LSTM with RMSE (120.76).

Table 5. Prediction of the next one month ahead load forecasting.

| Method | R^2 | MAPE | MAE | MSE | RMSE |
|--------|-------|------|-------|----------|--------|
| MLP | 0.92 | 4.23 | 71.26 | 11017.13 | 104.96 |
| LSTM | 0.92 | 4.38 | 75.12 | 11258.74 | 106.10 |
| CNN | 0.90 | 5.10 | 84.95 | 14584.13 | 120.76 |

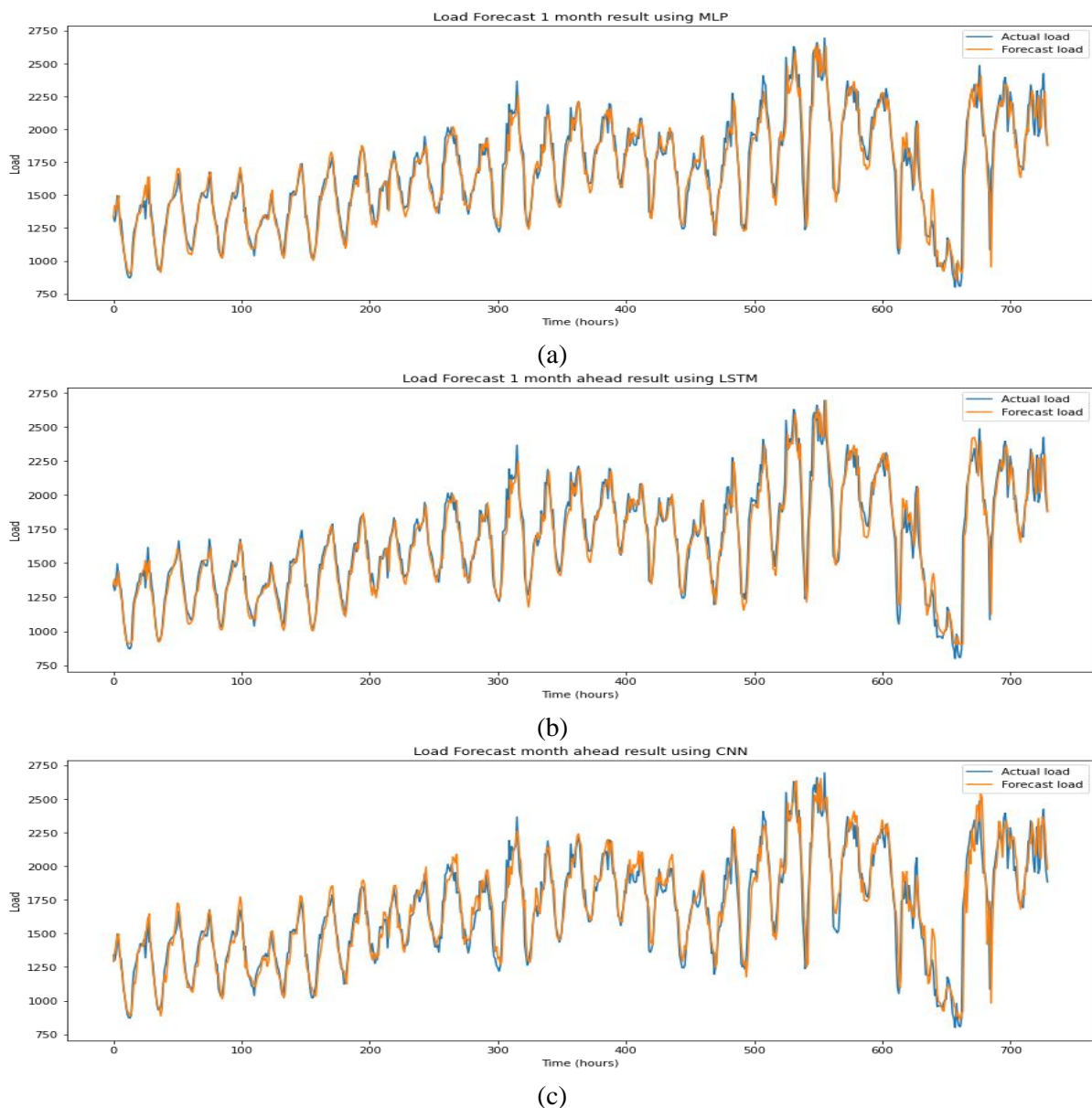


Figure 9. One month ahead load forecasting using a) MLP, b) LSTM, c) CNN.

Figure 9 shows the results of one month ahead load forecasts obtained by three different methods (MLP, LSTM, and CNN). From the extracted curves, it can be seen that the LSTM and CNN are closest to the actual load curve followed by MLP. The corresponding error values are obtained as reflected in Table 5.

4. Discussions

Accurate load forecasting can help alleviate the impact of renewable energy access to the network, facilitate the power plants to arrange unit maintenance, and encourage the power broker companies to develop a reasonable electricity price plan. Many activities within the power system such as the maintenance scheduling of generators, renewable energy integration, and even the investment of power plants and power grids depend on the load forecasting. In the electricity market, the regulators monitor the activities based upon the forecasting load and power generators. Customers and power brokers decide their action strategies.

Convolution Neural Network (CNN) is extensively used in forecasting. CNN are capable of apprehending pattern characteristics along with scale-invariant characteristics when the close by statistics has solid relationships with one another [76]. The design of locally set course of load informational data in close by hours could be extricated by CNN. In [77], another load forecasting design that utilizes CNN infrastructure is given and made a comparison to the rest of the neural systems. The outcomes reveal that MAPE along with CV-RMSE of suggested calculation is 9.77% and 11.66% that are the tiniest number across all the designs. The examinations demonstrate the fact about CNN infrastructure is essential regarding load forecasting and concealed characteristics could possibly be taken out through the planned 1D convolution surfaces. In light of the above description, LSTM as well CNN are equally shown to give high exactness forecast in STLF because of their advantageous feature to catch concealed characteristics. Along these lines, it is required to build up a hybrid neural systematic structure that can catch and incorporate such different invisible traits to give effective execution. More pointedly, it comprises of three sections: The LSTM module, the CNN module as well as featured fusion module. The LSTM module can get familiar with the valuable data for quite a while by the overlook gate as well as memory cell with the CNN module is used to remove schemes of nearby patterns and similar design which shows up in various areas. The featured fusion combination module is utilized to incorporate these unseen essentials and make the ultimate forecast. The suggested CNN-LSTM framework was created and implemented to foresee a real word electric load series of time. Also, a few strategies were actualized to be contrasted with our suggested model. To demonstrate the legitimacy of the suggested model, the CNN and LSTM modules were trailed separately. Moreover, the data record was separated into a few segments to test the strength of the suggested model. In outline, this paper proposes a profound learning structure that can successfully catch and incorporate the concealed characteristics of LSTM as well as the CNN model to accomplish far better precision and strength.

In this study, we computed ahead forecasting of one day, one week, two weeks and one month by applying MLP, LSTM and CNN. The computational performance was computed in terms of R^2 , MAPE, MAE, MSE and RMSE. We optimized the parameters of these algorithms in order to get the improved performance to obtained the ahead STLF and MTLF as reflected in the Table 6. The results yielded from our study with previous findings in terms of MPE for one day, one week, two weeks and one month reveals that our proposed models with parameters optimization gives the highest ahead detection performance.

Table 6. Comparison of MAPE results with previous studies.

| Method | Day-ahead | 72 hrs ahead | Week-ahead | 15 days ahead | Month ahead |
|-----------------|-----------|--------------|------------|---------------|-------------|
| RDNN [78] | 5.66% | - | 7.55% | - | - |
| SOM-SVM [79] | 6.05% | - | 8.03% | - | - |
| Capula DBN [80] | 5.63% | - | 7.26% | - | - |
| LSTM [81] | - | - | 9.34% | - | - |
| ANN [82] | - | 12.96% | - | - | - |
| NP-ARMA [83] | - | - | - | - | 4.83% |
| CNN [84] | - | - | - | - | 5.08% |
| LSTM [85] | - | - | - | - | 4.83% |
| XGBoost [86] | - | - | - | - | 10.08% |
| This Work | 4.97% | 6.34% | 6.16% | 5.44% | 4.23% |

5. Conclusion

In this study, we have taken the load data from a feeder and computed load forecasting for the next 24 hours, 72 hours, one-week, two-weeks, and one month to forecast the load demands for short-term and medium-term load demands. We optimized the AI algorithms such as MLP, LSTM, and CNN to improve the forecasting performance. We measured the forecasting performance based on the robust error metrics such as squared error, MAPE, MAE, MSE, and RMSE. The smallest error value between the observed and predicted values of these different ahead forecast value indicates that model best fit the data for the next 24 hours, 72 hours, one week and one-month load forecast. The smallest value of the R^2 statistical measure method indicates how closely the data to be fitted. The R-squared yields the smallest error in all these cases followed by MAPE, MAE, RMSE, and MSE. For STLF, the MLP and LSTM gives the better forecasting. However, for MTLF, the CNN and LSTM gives more better forecasting. It indicates that data demands are getting higher the deep learning models with a high number of neurons and optimized activation functions provide better predictions. The results indicate that the proposed model with optimizing the deep learning models yielded good predictions for short-term and medium-term forecasting. This indicated that the power systems with their complexity and growth and other different influential factors for power generation and consumption, power planning, etc. can be better predicted using this approach.

Limitations of study and future recommendations

There are various factors which effects the load growth such as peak hours in the day during which the demand of electricity increases or some environmental factors which effects on energy demand. Currently, we computed ahead load demands using different proposed models from January 2008 to December 2008 and July to December 2009 of one grid. We will considered each of these factors separately to see the effects of load growth.

Conflicts of interest

The authors declare no conflict of interest in this paper.

References

1. Y. Al-Rashid, L. D. Paarmann, Short-term electric load forecasting using neural network models,

- in: *Proc. 39th Midwest Symp. Circuits Syst.*, IEEE, (1996), 1436–1439.
2. T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, R. J. Hyndman, Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond, *Int. J. Forecast.*, **32** (2016), 896–913.
 3. T. Hong, P. Pinson, S. Fan, Global Energy Forecasting Competition 2012, *Int. J. Forecast.*, **30** (2014), 357–363.
 4. L. Hussain, M.S. Nadeem, S. A. A.Shah, Short term load forecasting system based on Support Vector Kernel methods, *Int. J. Comput. Sci. Inf. Technol.*, **6** (2014), 93–102.
 5. G. Gross, F. D. Galiana, Short-term load forecasting, *Proc. IEEE.*, **75** (1987), 1558–1573.
 6. P. J. Santos, A. G. Martins, A. J. Pires, Short-term load forecasting based on ANN applied to electrical distribution substations, in: *39th Int. Univ. Power Eng. Conf. UPEC 2004 - Conf. Proc.*, IEEE, **1**(2004), 427–432.
 7. E. Hossain, I. Khan, F. Un-Noor, S. S. Sikander, M. S. H. Sunny, Application of big data and machine learning in smart grid, and associated security concerns: A Review, *IEEE Access*, **7** (2019), 13960–13988.
 8. H. J. Sadaei, P. C. D.L. e Silva, F. G. Guimarães, M. H. Lee, Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series, *Energy*, **175** (2019), 365–377.
 9. H. M. Al-Hamadi, S. A. Soliman, Short-term electric load forecasting based on Kalman filtering algorithm with moving window weather and load model, *Electr. Power Syst. Res.*, **68** (2004), 47–59.
 10. K.-B. Song, Y.-S. Baek, D. H. Hong, G. Jang, Short-term load forecasting for the holidays using fuzzy linear regression method, *IEEE Trans. Power Syst.*, **20** (2005), 96–101.
 11. D. K. Ranaweera, N. F. Hubele, G. G. Karady, Fuzzy logic for short term load forecasting, *Int. J. Electr. Power Energy Syst.*, **18** (1996), 215–222.
 12. Y. He, Q. Xu, J. Wan, S. Yang, Short-term power load probability density forecasting based on quantile regression neural network and triangle kernel function, *Energy*, **114** (2016), 498–512.
 13. M. Q. Raza, A. Khosravi, A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings, *Renew. Sustain. Energy Rev.*, **50** (2015), 1352–1372.
 14. L. C. P. Velasco, C. R. Villezas, P. N. C. Palahang, J. A. A. Dagaang, Next day electric load forecasting using Artificial Neural Networks, in: *2015 Int. Conf. Humanoid, Nanotechnology, Inf. Technol. Control. Environ. Manag.*, IEEE, 2015, 1–6.
 15. L. Hernández, C. Baladrón, J. M. Aguiar, L. Calavia, B. Carro, A. Sánchez-Esguevillas, et al., Artificial Neural Network for short-term load forecasting in distribution systems, *Energies*, **7** (2014), 1576–1598.
 16. J. Buitrago, S. Asfour, Short-term forecasting of electric loads using nonlinear autoregressive Artificial Neural Networks with exogenous vector inputs, *Energies*, **10** (2017), 40.
 17. L. Suganthi, S. Iniyar, A. A. Samuel, Applications of fuzzy logic in renewable energy systems- A review, *Renew. Sustain. Energy Rev.*, **48** (2015), 585–607.
 18. K-H. Kim, J-K. Park, K-J. Hwang, S-H. Kim, Implementation of hybrid short-term load forecasting system using artificial neural networks and fuzzy expert systems, *IEEE Trans. Power Syst.*, **10** (1995), 1534–1539.
 19. D. X. Niu, H. F. Shi, D. D. Wu, Short-term load forecasting using bayesian neural networks learned by Hybrid Monte Carlo algorithm, *Appl. Soft Comput.*, **12** (2012), 1822–1827.
 20. F. Kaytez, M. C. Taplamacioglu, E. Cam, F. Hardalac, Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines, *Int. J. Electr. Power Energy Syst.*, **67** (2015), 431–438.

21. X. Li, X. Wu, Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition, in: *2015 IEEE Int. Conf. Acoust. Speech Signal Process*, IEEE, 2015, 4520–4524.
22. R. J. Williams, D. Zipser, A learning algorithm for continually running fully recurrent Neural Networks, *Neural Comput.*, **1** (1989), 270–280.
23. M. Ghofrani, M. Ghayekhloo, A. Arabali, A. Ghayekhloo, A hybrid short-term load forecasting with a new input selection framework, *Energy*, **81** (2015), 777–786.
24. G.-F. Fan, L.-L. Peng, W.-C. Hong, Short term load forecasting based on phase space reconstruction algorithm and bi-square kernel regression model, *Appl. Energy*, **224** (2018), 13–33.
25. R. Mamlook, O. Badran, E. Abdulhadi, A fuzzy inference model for short-term load forecasting, *Energy Policy*, **37** (2009), 1239–1248.
26. K. Metaxiotis, A. Kagiannas, D. Askounis, J. Psarras, Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher, *Energy Convers. Manag.*, **44** (2003), 1525–1534.
27. K. Fukushima, S. Miyake, A self-organizing neural network model for a mechanism of visual pattern recognition, in: *Comp. cooperation Neural Net.*, Springer Berlin Heidelb., 1982, 267–85.
28. Y. LeCun, P. Haffner, L. Bottou, Y. Bengio, Object recognition with gradient-based learning, in: *Shape, Cont. grop. comp. vision.*, Springer Berlin Heidelb., 1999, 319–345.
29. Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, Y. Guo, et al., Deep learning for visual understanding: A review, *Neurocomputing*, **187** (2016), 27–48.
30. H.-K. Yu, Weighted fuzzy time series models for TAIEX forecasting, *Phys. A Stat. Mech. Appl.*, **349** (2005), 609–624.
31. N.-Y. Wang, S.-M. Chen, Temperature prediction and TAIEX forecasting based on automatic clustering techniques and two-factors high-order fuzzy time series, *Expert Syst. Appl.*, **36** (2009), 2143–2154.
32. W.-J. Lee, J. Hong, A hybrid dynamic and fuzzy time series model for mid-term power load forecasting, *Int. J. Electr. Power Energy Syst.*, **64** (2015), 1057–1062.
33. R. Efendi, Z. Ismail, M. M. Deris, A new linguistic out-sample approach of fuzzy time series for daily forecasting of Malaysian electricity load demand, *Appl. Soft Comput.*, **28** (2015), 422–430.
34. R. Enayatifar, H. J. Sadaei, A. H. Abdullah, A. Gani, Imperialist competitive algorithm combined with refined high-order weighted fuzzy time series (RHWFTS–ICA) for short term load forecasting, *Energy Convers. Manag.*, **76** (2013), 1104–1116.
35. L. Hussain, W. Aziz, J. S. Alowibdi, N. Habib, M. Rafique, S. Saeed, et al., Symbolic time series analysis of electroencephalographic (EEG) epileptic seizure and brain dynamics with eye-open and eye-closed subjects during resting states, *J. Physiol. Anthropol.*, **36** (2017), 21.
36. L. Hussain, Detecting epileptic seizure with different feature extracting strategies using robust machine learning classification techniques by applying advance parameter optimization approach, *Cogn. Neurodyn.*, **12** (2018), 271–294.
37. L. Hussain, I. A. Awan, W. Aziz, S. Saeed, A. Ali, F. Zeeshan, et al., Detecting congestive heart failure by extracting multimodal features and employing machine learning techniques, *Biomed. Res. Int.*, **2020** (2020), 1–19.
38. Y. Asim, B. Raza, A. Kamran, M. Saima, A. K. Malik, S. Rathore, et al., A multi-modal, multi-atlas-based approach for Alzheimer detection via machine learning, *Int. J. Imag. Syst. Tech.*, **28** (2018), 113–123.
39. L. Hussain, S. Saeed, I. A. Awan, A. Idris, M. S. A. Nadeem, Q.-A. Chaudhry, et al., Detecting brain tumor using machines learning techniques based on different features extracting strategies,

- Curr. Med. Imag. Rev.*, **15** (2019), 595–606.
40. L. Hussain, W. Aziz, S. Saeed, I. A. Awan, A. A. Abbasi, N. Maroof, Arrhythmia detection by extracting hybrid features based on refined Fuzzy entropy (FuzEn) approach and employing machine learning techniques, *Waves Rand. Complex Media*, **30** (2020), 656–686.
 41. L. Hussain, W. Aziz, A. A. Alshdadi, M. S. A. Nadeem, I. R. Khan, Q.-A. Chaudhry, Analyzing the dynamics of lung cancer imaging data using refined fuzzy entropy methods by extracting different features, *IEEE Access*, **7** (2019), 64704–64721.
 42. A. A. Abbasi, L. Hussain, I. A. Awan, I. Abbasi, A. Majid, M. S. A. Nadeem, Q. -A. Chaudhary, Detecting prostate cancer using deep learning convolution neural network with transfer learning approach, *Cogn. Neurodyn.*, **5** (2020), 1–11.
 43. L. Hussain, A. Ahmed, S. Saeed, S. Rathore, I. A. Awan, S. A. Shah, et al., Prostate cancer detection using machine learning techniques by employing combination of features extracting strategies, *Cancer Biomark.*, **21** (2018), 393–413.
 44. C. Lang, F. Steinborn, O. Steffens, E. W. Lang, Applying a 1D-CNN network to electricity load forecasting, in: *Int. conf. time series forecas.*, Springer, Cham, 2020, 205–218.
 45. N. Javaid, NADEEM: A Novel Reliable Data Delivery Routing Protocol for Underwater WSNs, in: *Works. Int. Conf. Adv. Inf. Netw. App.*, Springer, Cham, 2019, 103–115.
 46. D. Hussain, T. Hussain, A. A. Khan, S. A. A. Naqvi, A. Jamil, A deep learning approach for hydrological time-series prediction: A case study of Gilgit river basin, *Earth Sci. Inform.*, **13** (2020), 915–927.
 47. A.M. Tudose, D.O. Sidea, I. I. Picioroaga, V. A. Boicea, C. Bulac, A CNN Based Model for Short-Term Load Forecasting: A Real Case Study on the Romanian Power System, in: *2020 55th Int. Univ. Power Eng. Conf.*, IEEE, 2020, 1–6.
 48. M. Solas, N. Cepeda, J. L. Viegas, Convolutional Neural Network for Short-term Wind Power Forecasting, in: *2019 IEEE PES Innov. Smart Grid Technol. Eur.*, IEEE, 2019, 1–5.
 49. A. E. Khantach, M. Hamlich, N. E. Belbounaguia, Short-term load forecasting using machine learning and periodicity decomposition, *AIMS Energy*, **7** (2019), 382–394.
 50. G. Dudek, Forecasting time series with multiple seasonal cycles using neural networks with local learning, in: *Int. Conf. Artif. Intell. Soft Comput.*, Springer, Berlin, Heidelberg., 2013, 52–63.
 51. M. A. Ghorbani, R. C. Deo, Z. M. Yaseen, M. H. Kashani, B. Mohammadi, Pan evaporation prediction using a hybrid multilayer perceptron-firefly algorithm (MLP-FFA) model: Case study in North Iran, *Theoret. App. Climat.*, **133** (2018), 1119–1131.
 52. M. E. Günay, Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicators and climatic conditions: Case of Turkey, *Energy Policy*, **90** (2016), 92–101.
 53. G. Dudek, Multilayer perceptron for short-term load forecasting : From global to local approach, *Neural Comput. Appl.*, **32** (2020), 3695–3707.
 54. X. Glorot, Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, in: *Proc. thirteenth int. conf. artif. intell. stat.*, 2010, 249–256.
 55. S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.*, **9** (1997), 1735–1780.
 56. H. Zheng, J. Yuan, L. Chen, Short-term load forecasting using EMD-LSTM neural networks with a Xgboost algorithm for feature importance evaluation, *Energies*, **10** (2017), 1168.
 57. Y. Zhao, Y. Shen, Y. Zhu, J. Yai, Forecasting wavelet transformed time series with attentive neural networks, in: *2018 IEEE Int. Conf. Data Min.*, IEEE, 2018, 1452–1457.
 58. K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, J. Schmidhuber, LSTM: A Search Space Odyssey, *IEEE Trans. Neural Networks Learn. Syst.*, **28** (2016), 2222–2232.

59. P. Nagabushanam, S. T. George, S. Radha, EEG signal classification using LSTM and improved neural network algorithms, *Soft Comput.*, **24** (2020), 9981–10003.
60. V. Y. Senyurek, M. H. Imtiaz, P. Belsare, S. Tiffany, E. Sazonov, A CNN-LSTM neural network for recognition of puffing in smoking episodes using wearable sensors, *Biomed. Eng. Lett.*, **10** (2020), 195–203.
61. A. Graves, J. Schmidhuber, Framewise phoneme classification with bidirectional LSTM and other neural network architectures, *Neural Networks*, **18** (2005), 602–610.
62. A. Graves, Long Short-Term Memory, in: *Sup. Seq. Label.Recurr. Neural Networks. Stud. Comput. Intell.*, Springer Berlin Heidelberg, 2012, 37–45.
63. X. Ma, Z. Tao, Z. Wang, H. Yu, Y. Wang, Long short-term memory neural network for traffic speed prediction using remote microwave sensor data, *Transp. Res. Part C Emerg. Technol.*, **54** (2015), 187–197.
64. E. Shelhamer, J. Long, T. Darrell, Fully convolutional networks for semantic segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.*, **39** (2017), 640–651.
65. Y. Taigman, M. Yang, M. Ranzato, L. Wolf, DeepFace: Closing the gap to human-level performance in face verification, in: *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2014, 1701–1708.
66. A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, L. Fei-Fei, Large-scale video classification with convolutional neural networks, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. IEEE*, 2014, 1725–1732.
67. T. Young, D. Hazarika, S. Poria, E. Cambria, Recent trends in deep learning based natural language processing, *IEEE Comput. Intell. Mag.*, **13** (2018), 55–75.
68. M. Cai, M. Pipattanasomporn, S. Rahman, Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques, *Appl. Energy*, **236** (2019), 1078–1088.
69. M. Canizo, I. Triguero, A. Conde, E. Onieva, Multi-head CNN–RNN for multi-time series anomaly detection: An industrial case study, *Neurocomputing*, **363** (2019), 246–260.
70. Z. Zeng, H. Xiao, X. Zhang, Self CNN-based time series stream forecasting, *Electron. Lett.*, **52** (2016), 1857–1858.
71. Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, *Proc. IEEE*, **86** (1998), 2278–2324.
72. L. Atlas, T. Homma, R. Marks, An artificial neural network for spatio-temporal bipolar patterns: Application to phoneme classification, in: *Neural Inf. Process*, 1988, 31–40.
73. J. Günther, P. M. Pilarski, G. Helfrich, H. Shen, K. Diepold, First steps towards an intelligent laser welding architecture using deep neural networks and reinforcement learning, *Procedia Technol.*, **15** (2014), 474–483.
74. N. Amjady, F. Keynia, Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm, *Energy*, **34** (2009), 46–57.
75. L. Hussain, S. Saeed, A. Idris, I. A. Awan, S. A. Shah, A. Majid, et al., Regression analysis for detecting epileptic seizure with different feature extracting strategies, *Biomed. Eng. Biomed. Tech.*, **64** (2019), 619–642.
76. S. Du, T. Li, X. Gong, Y. Yang, S.J. Horng, Traffic flow forecasting based on hybrid deep learning framework, in: *2017 12th Int. Conf. Intell. Syst. Knowl. Eng.*, IEEE, 2017, 1–6.
77. P.-H. Kuo, C.-J. Huang, A high precision artificial neural networks model for short-term energy load forecasting, *Energies*, **11** (2018), 213.
78. G. M. U. Din, A. K. Marnerides, Short term power load forecasting using Deep Neural Networks, in: *2017 Int. Conf. Comput. Netw. Commun.*, IEEE, (2017), 594–598.

79. S. Fan, L. Chen, Short-term load forecasting based on an adaptive hybrid method, *IEEE Trans. Power Syst.*, **21** (2006), 392–401.
80. T. Ouyang, Y. He, H. Li, Z. Sun, S. Baek, Modeling and forecasting short-term power load with copula model and deep belief network, *IEEE Trans. Emerg. Top. Comput. Intell.*, **3** (2019), 127–136.
81. X. Wang, F. Fang, X. Zhang, Y. Liu, L. Wei, Y. Shi, LSTM-based Short-term Load Forecasting for Building Electricity Consumption, in: *2019 IEEE 28th Int. Symp. Ind. Electron.*, IEEE, 2019, 1418–1423.
82. F. Rodrigues, C. Cardeira, J. M. F. Calado, The daily and hourly energy consumption and load forecasting using artificial neural network method: A case study using a set of 93 households in Portugal, *Energy Proced.*, **62** (2014), 220–229.
83. I. Shah, H. Iftikhar, S. Ali, Modeling and forecasting medium-term electricity consumption using component estimation technique, *Forecasting*, **2** (2020), 163–179.
84. S. H. Rafi, N. Al-Masood, Short Term Electric Load Forecasting Using High Precision Convolutional Neural Network, in: *2020 Int. Conf. Comput. Electr. Commun. Eng.*, IEEE, 2020, 1–7.
85. S. H. Rafi, N. Al-Masood, Highly Efficient Short Term Load Forecasting Scheme Using Long Short Term Memory Network, in: *2020 8th Int. Electr. Eng. Congr.*, IEEE, (2020), 1–4.
86. R. A. Abbasi, N. Javaid, M. N. J. Ghuman, Z. A. Khan, S. Ur Rehman, Amanullah, Short Term Load Forecasting Using XGBoost, in: *Work. Int. Conf. Adv. Info. Netw. App.*, (2019), 1120–1131.



AIMS Press

©2021 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)