

PERFORMANCE COMPARISON OF DEEP LEARNING MODELS FOR SHORT TERM LOAD FORECASTING

Jaya Christa Sargunar Thomas*¹ and Krishna Veni Kaliappan Sivakumar²

^{1, 2}Department of Electrical and Electronics Engineering, Mepco Schlenk Engineering College, Sivakasi, Tamilnadu, India.

¹<https://orcid.org/0000-0001-9033-4407>, ²<https://orcid.org/0000-0003-2186-2534>

Email: *stchrista@mepcoeng.ac.in, krishnaveni.ks@mepcoeng.ac.in

ARTICLE INFO

Article History

Received: March 11, 2026

Reviewed: April 11, 2026

Accepted: July 11, 2026

Published: August 31, 2026

Keywords:

Short term load forecasting,
Convolutional neural network,
Long short term memory,
Bi-directional long short term-
memory.

ABSTRACT

In today's power industry, the ever-changing load environment makes accurate load forecasts even more important than ever. Even, day ahead load forecasting is crucial for the successful and efficient operation of the grid. With the absence of good reliable forecasts of daily load fluctuations, systems within the grid can become overloaded, resulting in a failure of the system, thus establishing the need for sound forecasting methods. As a solution to this challenge, different types of deep learning algorithms namely, Long Short-Term Memory, Bidirectional Long Short-Term Memory, Convolutional Neural Network based LSTM, Convolutional Neural Network based BiLSTM are applied for short-term load forecasting in this paper. The efficacy of the various deep learning techniques for short term load prediction is compared. The technique which produces the least value of root mean squared error between real and estimated load, is the most preferred one. The comparative analysis shows that Convolutional Neural Network based BiLSTM technique is a better option for short term load forecasting.



Copyright ©2026 by authors and Galileo Institute of Technology and Education of the Amazon (ITEGAM). This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

I. INTRODUCTION

A nation's economy mainly depends on the power sector. To improve the nation's economy, power sector must be improved. If power system is stable, then the social stability, national security, economic growth and the overall quality of life of people is stable for a nation. Consumers' expectation from the power system is steadily growing and the power management in the grid is also improving with time. Every Power dispatch center needs load forecasting data for the better load distribution. The power industries must improve their abilities to produce power efficiently to manage the nation's technological advancements and economic improvements [1]. The growing smart grid faces significant challenges in energy management because of the unpredictability of power supply from renewable energy generation and also the power demand. Accurately projecting future generation and load is essential for effective power management [1]. Any significant deviation in the forecasts could result in technical and economic challenges. Furthermore, in deregulated power system markets, the accuracy of forecasted demand is crucial as all participants, from energy producers to customers, rely directly on this information [2]. Load forecasting has drawn a lot of attention from academia and industry due to the use of sophisticated data acquisition systems (like smart meters and Advanced Metering Infrastructure meters, as well as the development of data analytic methods and artificial intelligence. Numerous sophisticated data analysis methods have been developed and are widely used in modern life. Data analytics are crucial to the daily planning and operation of power systems. Predictive models are being used to analyze the demand pattern, with specific electrical generators designated to meet this demand. Load forecasting has different types, depending on the forecasting time horizon. Medium and long-term load prediction cover periods of two weeks and three years, respectively, whereas short-term load forecasting covers intervals of several minutes to one week [3]. Several factors specifically daily, weekly and yearly cycles of demand as well as unpredictable electricity consumption pattern of consumers' appliances contribute to the difficulty of load forecasting. Load also displays a strong correlation with environmental variables like temperature, humidity, solar irradiance, wind speed, etc.

A direct method for forecasting net-load and an indirect method for forecasting gross load and renewable generation separately are discussed in [4]. The findings show that the forecast accuracy is higher in the direct approach. This research also aims to forecast the net load in a power system.

1.1 LITERATURE REVIEW

While Classical models and forecasting methods typically are implemented using ARMA, SARIMA, and ARMAX modelling techniques, in past few years Recurrent Neural Networks (RNN) and Artificial Neural Networks have also received some recognition for possible applications in forecasting [5]. As a result of these advancements in technology, developers have begun testing and using a variety of new models. These new models use Deep Learning techniques to identify key information within the load profile [6]. Although RNNs provide a number of benefits in modelling load time series data, they struggle to model long term dependencies accurately because of their architecture. To solve this issue, developers created Long Short-Term Memory Networks (LSTMs) [7]. Although the base architecture of an LSTM is similar to that of a standard RNN, it incorporates memory blocks to store information. The memory blocks contain both Subnets and Recurrent Connection units, as opposed to the Summation units that are used by the buried layers of RNNs. LSTMs have consistently provided superior performance for solving industrial problems involving time series data compared to RNNs. Through the incorporation of memory units, LSTMs have consistently provided superior results for modelling extremely long-term dependencies in multiple real-world sectors [8]. Long-term memory tasks such as predicting the secondary structure of proteins and context-free grammar were successfully handled by LSTM models [9]. Although LSTM can learn to identify long-term dependencies for sequential data, it is limited to working with historical data; Bi-directional LSTM (Bi-LSTM) is able to utilize both past and upcoming context to produce highly exact estimations. Therefore, when completing tasks where bidirectional information is necessary such as speech recognition or natural language processing, Bi-LSTM has a distinct advantage over LSTM. Whether LSTM or Bi-LSTM should be used in a time series forecasting application will largely depend on how the forecast will be used and on the data being analyzed [10].

By using Bi-LSTM, an analyst can represent complicated patterns and relationships in the data much more efficiently than with unidirectional LSTM operating only in one direction. Combining Empirical Mode Decomposition with the Bi-LSTM approach allows for highly accurate forecasting of both short-term and long-term dependencies of electrical load. The advantages of the new model are evident when considering that it produces more accurate short-term load forecasts than any of the previous models have produced; thus, it is very likely to be an effective method of forecasting energy consumption [11]. Recently, studies investigating the benefit of Convolutional Neural Networks (CNN) for short-term load prediction are conducted. In contrast to LSTM networks, temporal CNN could produce a forecasting model that was both effective and dependable. In this paper, it is suggested using an integrated CNN-LSTM model to predict the future load. Lastly, a unique parallel LSTM-CNN network that was trained for both CNN and LSTM was developed to resolve the short term load estimation problem in smart grids. Results proved that the LSTM-CNN demonstrated good performance [12]. CNN-Bi-LSTM is another model used for load forecasting. CNN handles the missing data, while Bi-LSTM is used for the long term load forecasting. By combining both the models, the results are better for load forecasting for discontinuous dataset [13].

1.2 OBJECTIVES OF THE RESEARCH

By looking into the literature, a detailed study comparing the various deep learning models for short term load estimation is missing. Hence, in this paper, an analysis is conducted on the four prominent deep learning models - LSTM, Bi-LSTM, CNN-LSTM, and CNN-Bi-LSTM, to determine the best technique for load estimation in accordance with the evaluation of results [14]. The primary goal is to identify the optimal approach for load estimation in the power system by assessing the deviation between the real load and the estimated load. In this paper, an aggregated dataset of electrical load profiles from New England, USA is utilized which is not used in the above mentioned papers. Program code for the deep learning models is created using Python.

II. DEEP LEARNING NETWORK ARCHITECTURES

II.1 LSTM ARCHITECTURE

Traditional RNNs have problems in capturing long-term dependencies because of their single hidden state that persists over time. LSTM networks address the issue with long-term storage found in traditional RNN models through a special component called the memory cell which has the ability to hold on information for much longer duration of time. Therefore, LSTM models are excellent at learning temporal relationships over many different time periods and are commonly used in areas such as voice-to-text transcription, translating between different languages and predicting future values based on the data available. Figure 1 shows LSTM model's architecture. In the overall structure of an LSTM model, there are three primary areas known as gates. The input gate is where the new piece of information enters; the forget gate is the place to remove data that is no longer find useful; and the output gate is the point where the information needed is taken out from the memory cell. By selectively controlling the retention and discarding of information at each step, LSTM networks effectively manage long-term dependencies and facilitate strong learning processes [15].

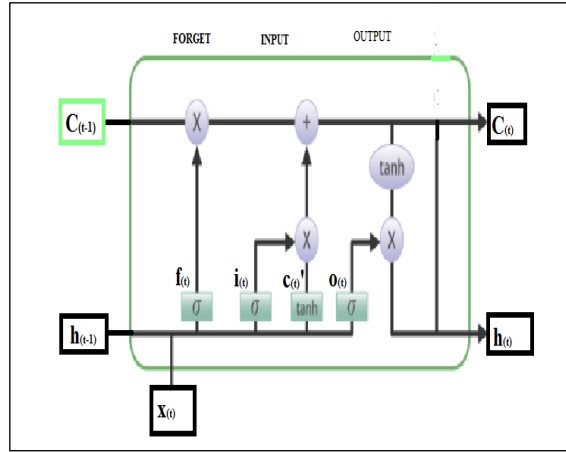


Figure 1: Architecture of LSTM.
Source: Authors, (2026).

II.2 BI-LSTM ARCHITECTURE

Figure 2 demonstrates the architecture of Bi-LSTM model which is a type of sequential model consisting of two LSTM layers: one for handling sequences in forward direction and another for treating sequences in reverse direction. The outputs of both LSTMs are converted to a probability vector and combined to create the final output [9]. Bi-LSTMs are often used when performing Natural Language Processing tasks because they allow the model to have a greater understanding of the complex interrelations between elements of a sequence by using the context to infer the meaning of certain elements.

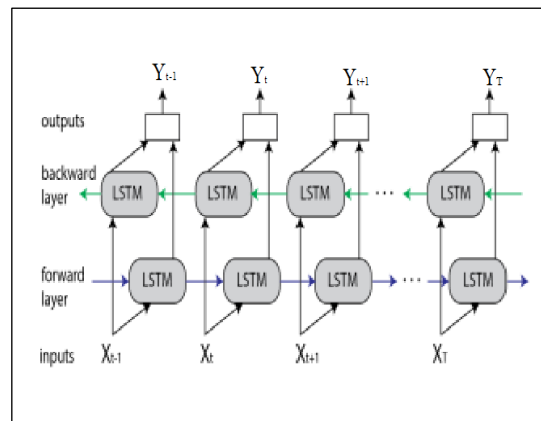


Figure 2: Architecture of Bi-LSTM.
Source: Authors, (2026).

II.3 CNN-LSTM ARCHITECTURE

Combining CNN with LSTM creates a fusion network for Deep Learning that can analyze different types of sequential information. It uses the benefits of both CNN and LSTM networks to learn both spatial and temporal characteristics in sequential data, creating an effective solution for tasks related to video classification, action recognition and spatiotemporal forecasting. In a typical CNN-LSTM architecture illustrated in Figure 3, the role of CNN is to perform the extraction of the spatial features of the data, while the role of the LSTM is to model the temporal dependencies and sequential nature of the data. CNN networks are particularly good at extracting information such as feature maps or key frames using their various convolutional and pooling layers. The CNN network produces a series of spatial feature vectors, which are then fed to the LSTM network to learn the temporal dependencies present inside the data. LSTMs are especially effective at modeling long-term dependence on sequential data since they maintain a memory cell to store and update the values of previous time steps. Therefore, if CNN is employed to pull out spatial peculiarities, and LSTM is used to simulate temporal dependencies of sequential data, then the combined model of CNN-LSTM can learn how the spatial feature vectors relate temporally. The combination of CNNs with LSTMs provides an open way for CNN-LSTM to process and understand complex and sequential information. Because an image or video may be viewed as a sequence of images overtime, this makes it very useful for a variety of applications such as video understanding, action recognition or forecasting time series data [11]. The CNN-LSTM architecture provides a method to analyze both spatial features and temporal features at the same time, making the architecture flexible and very effective at understanding the context of an input sequence as well as recognizing patterns and relationships between elements within that sequence.

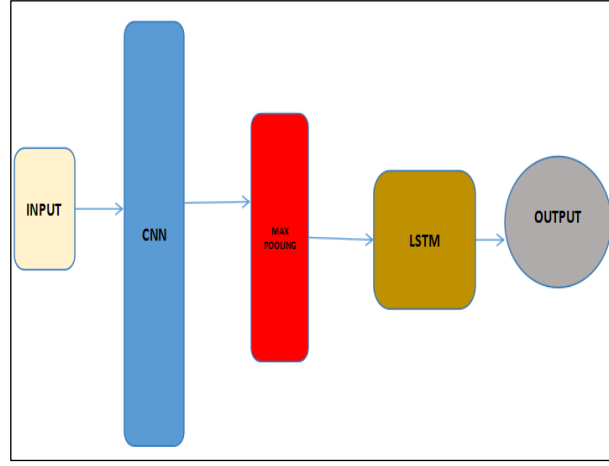


Figure 3: Architecture of CNN-LSTM.
Source: Authors, (2026).

II.4 CNN-BI-LSTM ARCHITECTURE

CNN-BiLSTM is a network architecture that integrates CNN with Bi-LSTM networks demonstrated in Figure 4. By utilizing the information from CNN's features before inputting them into Bi-LSTMs, they benefit from the existing knowledge of how to process inputs. Initially, the dataset was segmented into load profiles containing load values for each day (i.e., load profiles for a 24 hour period) and these profiles were then combined to create a single input for the One Dimensional (1D) CNN layer of the architecture. During the extraction of features from the load profile, the 1D CNN layer used 64 filters with a kernel size (i.e., width and height of the filter) of 3 by 1 and a default stride of 1 when generating a feature vector for each input load profile. After extracting the features from the input load profile data, those raw feature map data are still subjected to max-pooling which decreases the dimensionality of the features but also, likely, loses some of the important aspects of the data; therefore, this loss of important information may account for the decrease in accuracy experienced by the Bi-LSTM networks [12].

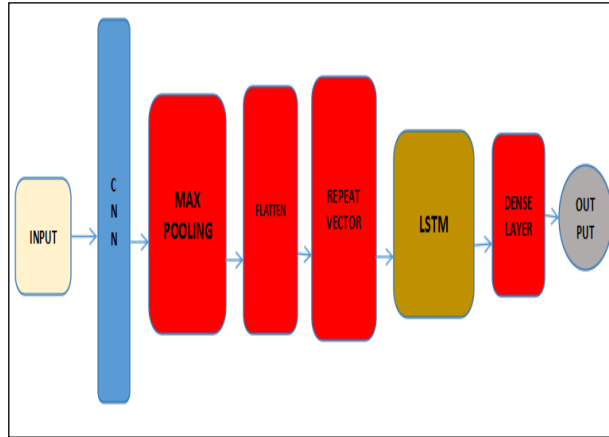


Figure 4: Architecture of CNN-Bi-LSTM.
Source: Authors, (2026).

The feature vector output from the extraction layer of the network will then feed into the Bi-LSTM networks to become the sequential load demand forecasts for the succeeding day's load profile. The fully-connected layer of the network will also include rectified linear unit activation functions for use when creating forecast load profiles, however, the CNN-LSTM versions of the network will only utilize standard, one-way directed LSTM networks for the sequential time series forecasts.

III. METHODOLOGY

III.1 DATASET PREPARATION

The data used in this research comes from aggregated electricity usage profiles in the New England Region in the United States [16]. This dataset used consists of hourly (60 minute) data for active power beginning from January 2019 to January 2024. As part of load forecasting, this study focuses solely on active power measurements. Additionally, as presented in Table 1, the data is classified according to time into training and test sets, which are created using the provided timestamps. In order to prepare the data for use with the models, min/max normalization will be applied to the dataset based on the equation (1).

$$X_{\text{norm}} = (X - X_{\text{min}}) / (X_{\text{max}} - X_{\text{min}}) \quad (1)$$

III.2 TRAINING AND TESTING

The dataset used in this study is made up of total load, which was first divided into training and test dataset. The training dataset is then utilized to build up various neural network architectures: LSTM, Bi-LSTM, CNN-LSTM, and CNN-Bi-LSTM; with the test dataset being specifically used to assess the model's efficacy. 70% of data are utilized for training the models and 30% for testing which is shown in Table 1.

Table 1: Data division into training and testing set.

Data Description	Starting date	Ending date
Training Set	06/01/2019	06/07/2022
Test Set	07/07/2022	09/01/2024

Source: Authors, (2026).

Before being passed onto the model during the actual training phase, the training dataset is normalised. The model has been tasked with predicting the demand for the next 24 hours in accordance with the previous history of data [17]. The time window will be moved forward by 1 hour, from the previous hour, during each iteration (1 hour), until all data entries in the dataset are used to train the model. In order to achieve optimal results, increased training set size and well-tuned parameters (Batch Size, Number of Cells, Number of Layers) are helpful. When the model has been trained, a walk-forward testing method can be used to assess the technique's efficacy. To evaluate the performance of the trained model, Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) are the metrics adopted. RMSE is given by equation (2).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}} \quad (2)$$

In this context, 'n' denotes total number of load data within the 24-hour period, where P_i denotes the real load at the 'ith' hour and O_i signifies the estimated load at the 'ith' hour [18]. MAE is the mean absolute deviations between estimated and actual loads. MAPE is the mean absolute percentage deviation between estimated and real loads.

IV. RESULTS AND DISCUSSION

During the testing phase, load is forecasted for the testing period day by day from 07/07/2022 to 09/01/2024, with load forecasted for one day at a time. Figure 5 depicts the plot of actual versus predicted load by all the models, over a 24-hour period for one such testing day which is 4th November 2023. The performance of each model is discussed below.

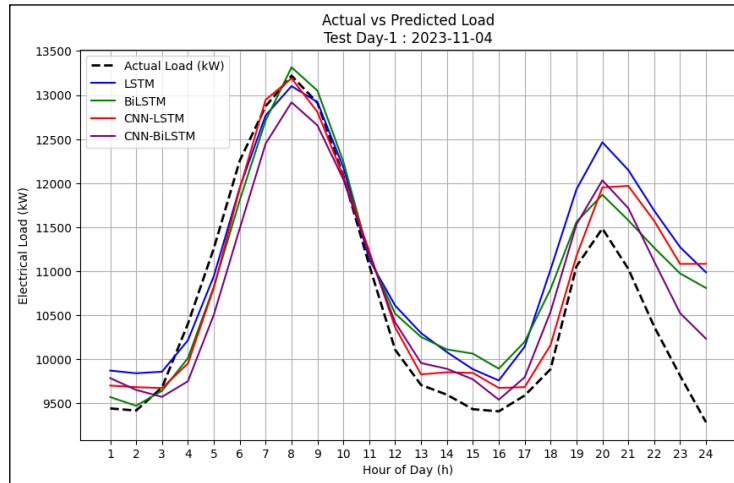


Figure 5: Real Vs forecasted load by all the models.

Source: Authors, (2026).

IV.1 LSTM BASED LOAD FORECASTING

Figure 6 depicts the plot of actual versus predicted load by LSTM model over a 24-hour period for the above mentioned testing day. Figure 7 displays an hourly RMSE plot that compares the real load with the forecasted load. Lower RMSE values indicate more precise estimations. The LSTM forecast method achieves an average daily RMSE score of 0.060051 kW considering all the test days.

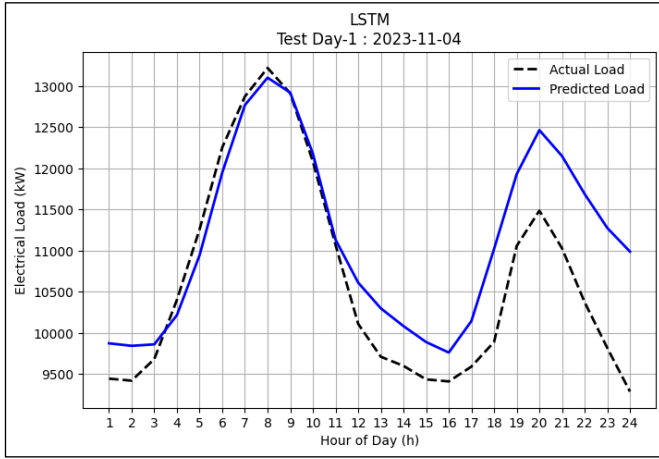


Figure 6: Real Vs forecasted load by LSTM.
Source: Authors, (2026).

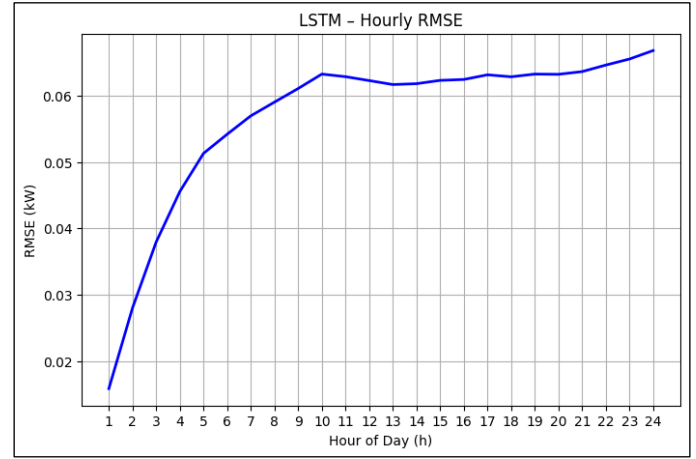


Figure 7: Hourly RMSE values for LSTM networks.
Source: Authors, (2026).

Figure 8 shows the plot between training and validation loss which shows that LSTM model is a good fit model and it converges in 47th epoch.

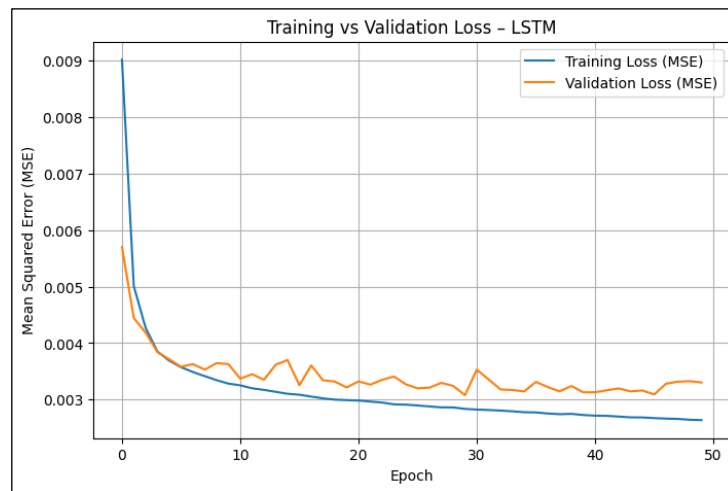


Figure 8: Training vs Validation loss for LSTM model.
Source: Authors, (2026).

IV.2 BI-LSTM BASED LOAD FORECASTING

Figure 9 illustrates the plot of actual versus predicted load over a 24-hour period by Bi-LSTM. Figure 10 showcases an hourly RMSE plot for the same test day. The BiLSTM forecast method achieves an average daily RMSE score of 0.060781 kW.

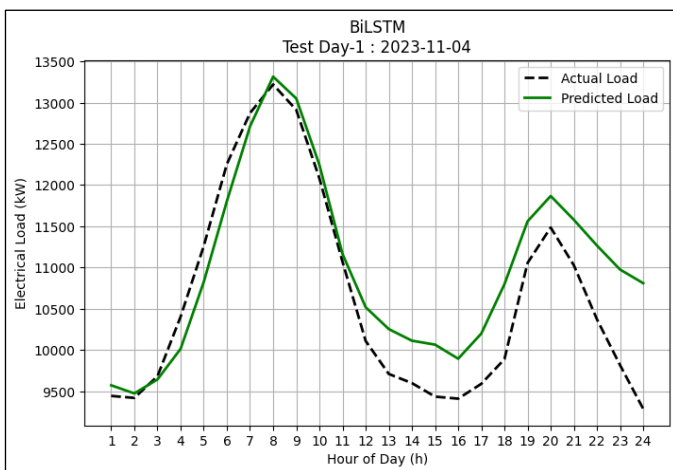


Figure 9: Real Vs forecasted load by Bi-LSTM.
Source: Authors, (2026).

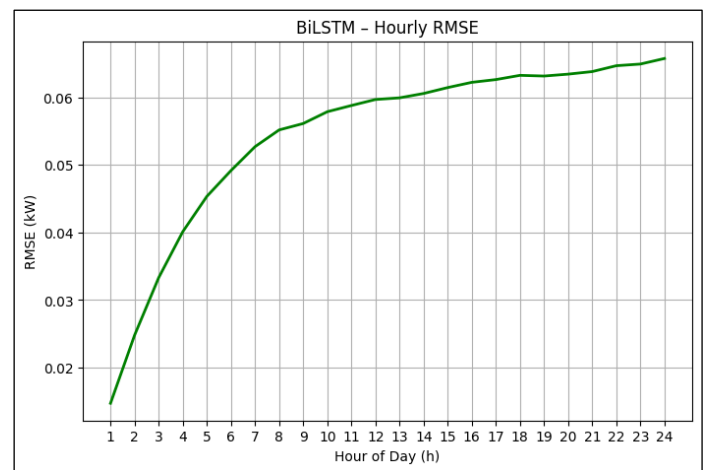


Figure 10: Hourly RMSE values for Bi-LSTM networks.
Source: Authors, (2026).

Figure 11 shows the plot between training and validation loss which shows that BiLSTM model is another good fit model and it converges in 49th epoch.

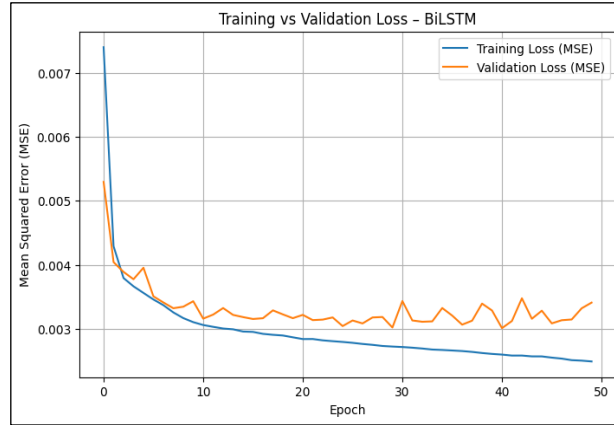


Figure 11: Training vs Validation loss for BiLSTM model.
Source: Authors, (2026).

IV.3 CNN-LSTM BASED LOAD FORECASTING

Figure 12 displays the plot of actual versus predicted load over a 24-hour period by CNN-LSTM model. Figure 13 presents an hourly RMSE plot comparing the actual and predicted load demand. The CNN-LSTM forecast method achieves an average daily RMSE score of 0.061192 kW.

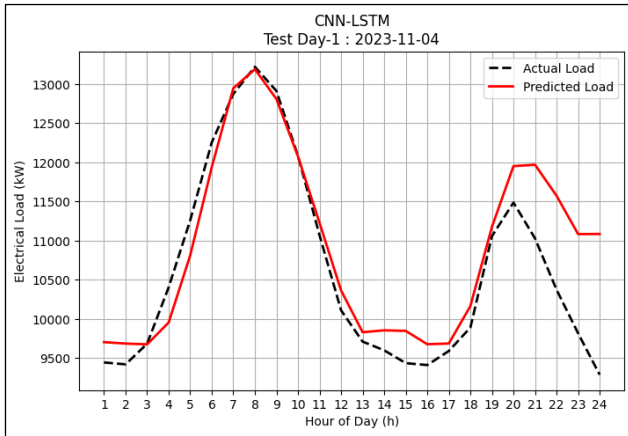


Figure 12: Real Vs forecasted load by CNN-LSTM.
Source: Authors, (2026).

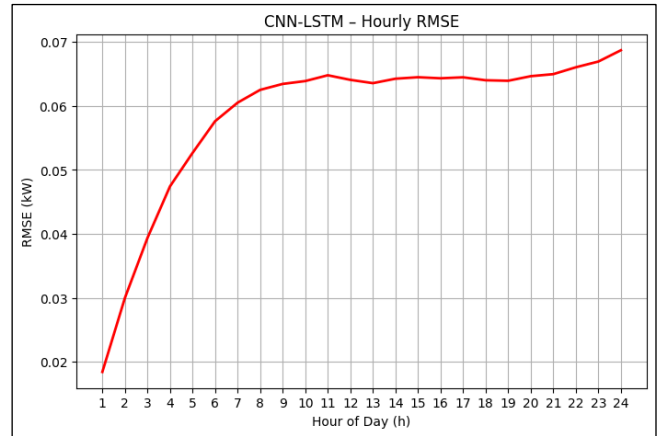


Figure 13: Hourly RMSE values for CNN-LSTM network.
Source: Authors, (2026).

Figure 14 shows the plot between training and validation loss which shows that CNN-LSTM model is also a good fit model and it converges in 49th epoch.

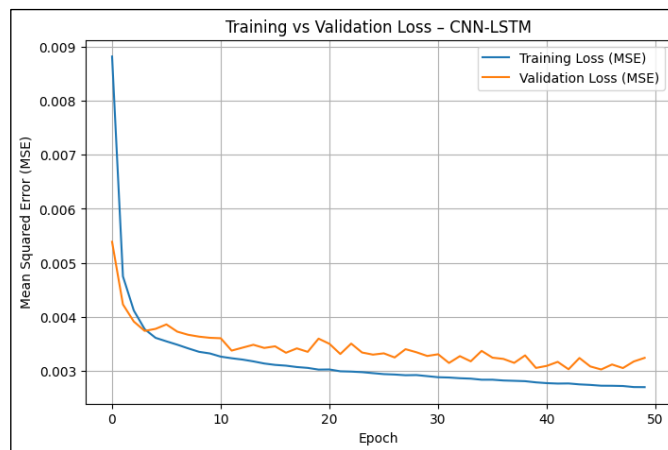


Figure 14: Training vs Validation loss for CNN-LSTM model.
Source: Authors, (2026).

IV.4 CNN-BiLSTM BASED LOAD FORECASTING

Figure 15 illustrates the plot of actual versus predicted load over a 24-hour period for CNN-BiLSTM model. Figure 16 depicts an hourly RMSE plot that showcases the contrast between the actual and predicted load demand. The CNN-BiLSTM algorithm brings about an average daily RMSE result of 0.058449 kW.

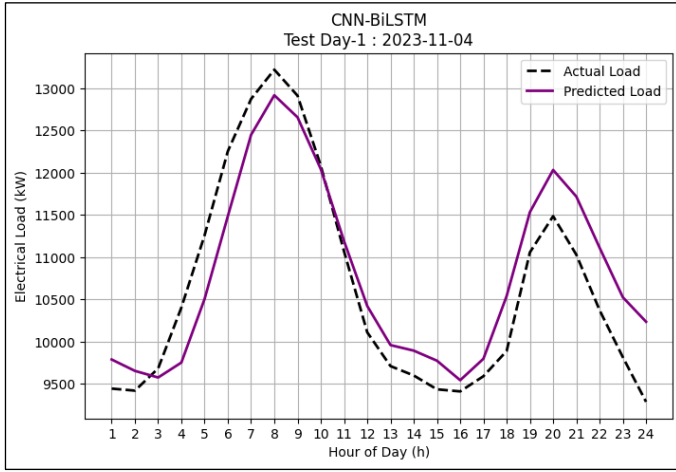


Figure 15: Real Vs forecasted load for CNN-BiLSTM.
Source: Authors, (2026).

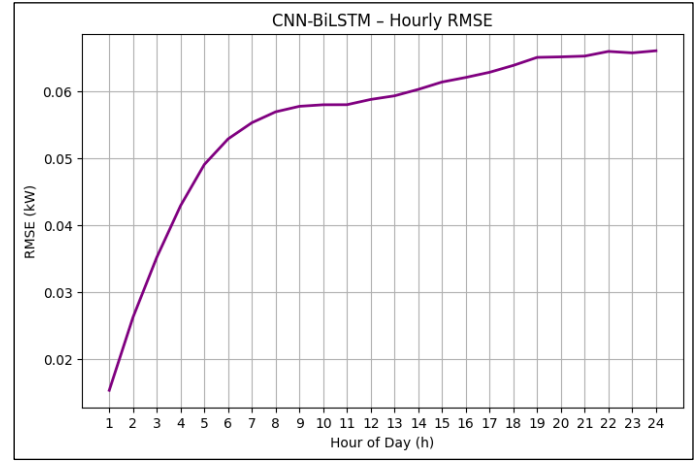


Figure 16: Hourly RMSE values for CNN-BiLSTM network.
Source: Authors, (2026).

Figure 17 shows the plot between training and validation loss which shows that CNN-BiLSTM model is a good fit model and it converges in 49th epoch.

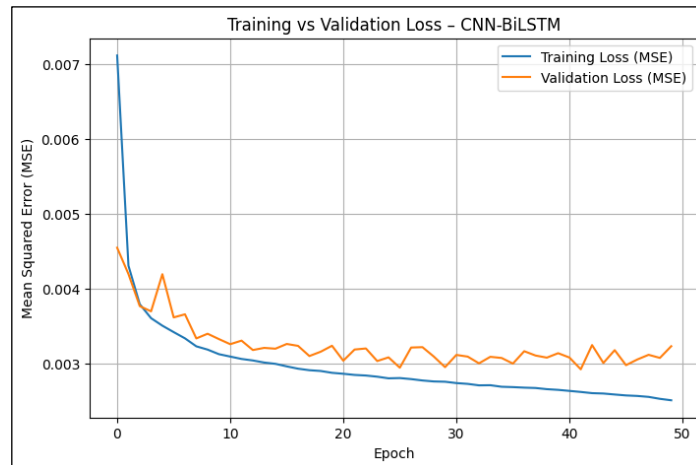


Figure 17: Training vs Validation loss for CNN-BiLSTM model.
Source: Authors, (2026).

IV.5 COMPARISON OF RESULTS

The average RMSE values, MAE values and MAPE % for the different models are presented in Table 2.

Table 2: Prediction error for various algorithms.

Algorithms	RMSE (kW)	MAE (kW)	MAPE (%)
LSTM	0.060051	0.043350	20.883416
BiLSTM	0.060781	0.043708	20.948451
CNN-LSTM	0.061192	0.044336	21.759486
CNN-BiLSTM	0.058449	0.042286	20.667672

Source: Authors, (2026).

According to the results shown in Table 2, the CNN-BiLSTM model exhibits the least RMSE, MAE values and MAPE percentage in comparison to the other techniques. The CNN-BiLSTM model's ability to extract features effectively and to process data bidirectionally, contributes to more valid estimations. For load forecasting in regions like New England, the CNN-BiLSTM model stands out as the best model based on its performance.

V. CONCLUSIONS

Due to the need for sound load forecasting methods for the power industry, this research paper deals with the application of four deep learning algorithms namely, LSTM, BiLSTM, CNN-LSTM, CNN-BiLSTM for short-term load forecasting. These models were trained to forecast the aggregated load of New England Region in the United States, one day in advance based on the previous history of load profile. Results from this study, indicate that the CNN-BiLSTM architecture produces the lowest RMSE error values of any of the models that were tested against it, including LSTM, BiLSTM and CNN-LSTM. The authors speculate that this is due to the combined strength of the hybrid model. Automatic Feature Extraction by CNN and Temporal Dependency Learning by BiLSTM attribute to an accurate modeling of electricity demand patterns.

This research is very much useful to the power industry to advance power system management strategies in the framework of increasing renewable energy integration and smart home technologies. Future evaluations will progressively compare the performance of the models using more dynamic and less predictable data-sets like weather information and wide ranging load profiles.

VI. AUTHORS' CONTRIBUTION

Conceptualization: Jaya Christa Sargunar Thomas and Krishna Veni Kaliappan Sivakumar.

Methodology: Krishna Veni Kaliappan Sivakumar.

Investigation: Jaya Christa Sargunar Thomas and Krishna Veni Kaliappan Sivakumar.

Discussion of results: Jaya Christa Sargunar Thomas.

Writing – Original Draft: Jaya Christa Sargunar Thomas.

Writing – Review and Editing: Jaya Christa Sargunar Thomas and Krishna Veni Kaliappan Sivakumar.

Resources: Krishna Veni Kaliappan Sivakumar.

Supervision: Jaya Christa Sargunar Thomas.

Approval of the final text: Jaya Christa Sargunar Thomas.

VII. ACKNOWLEDGEMENTS

The authors thank the Mepco Schlenk Engineering College Management Committee for providing support to conduct this research.

VIII. REFERENCES

- [1] L. Wang, S. Mao, and B. Wilamowski, "Short-term load forecasting with LSTM-based ensemble learning," in Proc. 2019 Int. Conf. on Internet of Things (iThings), IEEE Green Computing and Communications (GreenCom), IEEE Cyber, Physical and Social Computing (CPSCom), and IEEE Smart Data (SmartData), 2019, pp. 793–800.
- [2] M. A. Alotaibi, "Machine learning approach for short-term load forecasting using deep neural network," *Energies*, vol. 15, no. 17, pp. 1–23, 2022, doi: 10.3390/en15176261.
- [3] M. Madhukumar, A. Sebastian, X. Liang, M. Jamil, and M. N. S. K. Shabbir, "Regression model-based short-term load forecasting for university campus load," *IEEE Access*, vol. 10, pp. 8891–8905, 2022, doi: 10.1109/ACCESS.2022.3144206.
- [4] O. Olasupo, T. Ayodele, and A. Ogunjuyigbe, "An Evaluation of Direct and Indirect Strategies for Hierarchical Net-Load Forecasting," *JETIA*, vol. 12, no. 57, pp. 375–388, 2026.
- [5] S. Muzaffar and A. Afshari, "Short-term load forecasts using LSTM networks," *Energy Procedia*, vol. 158, pp. 2922–2927, 2019.
- [6] C. Cai, Y. Tao, T. Zhu, and Z. Deng, "Short-term load forecasting based on deep learning bidirectional LSTM neural network," *Applied Sciences*, vol. 11, no. 17, pp. 1–16, 2021.
- [7] D. G. da Silva and A. A. de M. Meneses, "Comparing long short-term memory (LSTM) and bidirectional LSTM deep neural networks for power consumption prediction," *Energy Reports*, vol. 10, pp. 3315–3334, 2023, doi: 10.1016/j.egy.2023.09.175.
- [8] D. P. Mohan and M. S. Subathra, "A comprehensive review of various machine learning techniques used in load forecasting," *Recent Advances in Electrical & Electronic Engineering*, vol. 16, no. 3, pp. 197–210, 2023, doi: 10.2174/2352096515666220930144336.
- [9] N. Mounir, H. Ouadi, and I. Jhrilifa, "Short-term electric load forecasting using an EMD-BI-LSTM approach for smart grid energy management system," *Energy and Buildings*, vol. 288, pp. 1–10, 2023.
- [10] B. Ibrahim, L. Rabelo, E. Gutierrez-Franco, and N. Clavijo-Buritica, "Machine learning for short-term load forecasting in smart grids," *Energies*, vol. 15, no. 21, pp. 1–19, 2022, doi: 10.3390/en15218079.
- [11] S. Dai, "Short-term load forecasting with deep learning techniques," *Journal of Physics: Conference Series*, vol. 2547, pp. 1–8, 2023, doi: 10.1088/1742-6596/2547/1/012025.
- [12] Y. Hong, Y. Zhou, Q. Li, W. Xu, and X. Zheng, "A deep learning method for short-term residential load forecasting in smart grid," *IEEE Access*, vol. 8, pp. 55785–55797, 2020.
- [13] B. Zhong, "Deep learning integration optimization of electric energy load forecasting and market price based on the ANN–LSTM–transformer method," *Frontiers in Energy Research*, vol. 11, pp. 1–17, 2023, doi: 10.3389/fenrg.2023.1292204.
- [14] D. L. Marino, K. Amarasinghe, and M. Manic, "Building energy load forecasting using deep neural networks," in Proc. 42nd Annu. Conf. IEEE Ind. Electron. Soc. (IECON), 2016, pp. 7046–7051, doi: 10.1109/IECON.2016.7793413.
- [15] A. Narayan and K. W. Hipel, "Long short term memory networks for short-term electric load forecasting," in Proc. IEEE Int. Conf. Systems, Man, and Cybernetics, 2017, pp. 2573–2578, doi: 10.1109/SMC.2017.8123012.
- [16] ISO New England Web Page of Energy, Load, and Demand Reports <https://www.iso-ne.com/isoexpress/web/reports/load-and-demand> (accessed January 6, 2026).
- [17] M. Jacob, C. Neves, and D. Vukadinovic Greetham, "Short term load forecasting," in *Forecasting and Assessing Risk of Individual Electricity Peaks*, Mathematics of Planet Earth, Springer, 2020, pp. 15–37, doi: 10.1007/978-3-030-28669-9_2.
- [18] A. Ahmad, N. Javaid, A. Mateen, M. Awais, and Z. A. Khan, "Short-term load forecasting in smart grids: An intelligent modular approach," *Energies*, vol. 12, no. 1, pp. 1–21, 2019, doi: 10.3390/en12010164.