SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V11I10P107

Volume 11 Issue 10, 61-70, October 2024

Original Article

Enhanced Load Forecasting Using CNN-BiLSTM Models in University Buildings with Solar PV

Muhammad Zulhamizan Ahmad[1]*, Nofri Yenita Dahlan[1], Zuhaila Mat Yasin[1]

[1]*Solar Research Institute (SRI), Universiti Teknologi Mara (UiTM), Shah Alam, Selangor, Malaysia.*

[1]*Corresponding Author : mzulhamizan98@gmail.com*

Received: 04 August 2024 Revised: 05 September 2024 Accepted: 06 October 2024 Published: 30 October 2024

Abstract *- Energy management is an imperative practice involving the detailed monitoring, regulation, and optimization of energy consumption within various domains aimed at conserving resources and controlling energy costs. The escalating demand for electricity and integrating renewable energy sources has brought forth an array of complexities that challenge energy management efforts. As a result, the need to enhance the precision of load forecasting has surged in importance, attracting significant attention from researchers and organizations alike. Traditional time series models have their own limitations. These models rely on the assumption of linear relationships and stationary time series data, thereby potentially falling short of capturing the intricate, non-linear variations often present in energy consumption patterns. This limitation necessitates the exploration of more advanced and adaptable forecasting techniques. Campus buildings present some challenges for load forecasting. These challenges arise from the dynamic and ever-changing load patterns within educational institutions, which can fluctuate significantly based on various factors such as lecture schedules, semester breaks, and special occasions. This study introduces an innovative hybrid model called CNN-BiLSTM, which integrates Convolutional Neural Networks (CNN) with Bidirectional Long Short-Term Memory (BiLSTM) models to address the complexities of load patterns to produce accurate forecasts. The proposed model is thoroughly benchmarked against traditional Artificial Neural Networks (ANN) and BiLSTM models. Load data from the UiTM Permatang Pauh campus building, which encompasses 343 days of data collected at 30-minute intervals, a total of 16,464 data points for analysis. Leveraging this load data, comprehensive feature engineering was conducted, leading to the generation of categorical data such as hour, calendar attributes, and semester status. The CNN-BiLSTM model outperforms its counterparts, achieving a remarkable Mean Absolute Percentage Error (MAPE) of 6.9%. Therefore, as demonstrated through rigorous benchmarking, the model's superior performance highlights its potential significance for improving energy management in educational institutions and other domains with similar load complexity.*

Keywords *- Load forecasting, Artificial Neural Network, Bidirectional, Long Short-Term Memory. Convolutional Neural Network, CNN-LSTM, Campus building.*

1. Introduction and Literature Review

Energy management refers to the systematic and strategic monitoring, control, and optimization of an entity's energy consumption to conserve resources and minimise energy costs. The growing electricity demand, combined with the integration of renewable energy, is making energy management more difficult since renewable energy leads to greater fluctuations in power levels within the system [1]. Organizations must prioritize their power systems' reliable, secure, and efficient operation. Therefore, improving accuracy in load forecasting has gained interest among researchers. Accurate load consumption forecasting is important in energy management as it gives insights to the energy manager on future or estimated load usage, which helps to ensure power system stability, safety, and optimized operational costs [2]. Load dispatch planning relies on an accurate load forecasting model to improve power system performance. Ensuring the power supply's reliability and the grid's stability hinges on the efficient dispatch of electricity. Nevertheless, the task of managing dispatch becomes intricate due to the inherent fluctuations and intermittent nature of electric power [3]. These challenges demand high-accuracy load forecasting to improve power distribution and efficient energy usage [4].

Buildings account for approximately 40% of global energy consumption [5]. It is projected that by 2030, this proportion will escalate to 50%. Within Malaysia, buildings account for a significant 48% of the country's generated electricity consumption [5]. The substantial energy consumption involves both residential and commercial buildings, a trend accentuated by the growing population. Consequently, the significance of effective energy management in handling building energy becomes paramount, extending even to campus facilities. According to the

Malaysian Ministry of Education, there are more than 100 educational institutions in Malaysia, including both public and private universities [6]. Given Malaysia's considerable number of campus buildings, there is a pressing need to enhance current energy management practices, with a particular emphasis on refining load forecasting accuracy to effectively anticipate energy consumption within these facilities [7]. Modelling precise load forecasting for campus buildings poses challenges due to varying load patterns arising from lecture sessions, semester breaks, and, notably, elevated weekend loads attributed to student residences [8]. Given the inherent complexity and fluctuations present in load consumption patterns, traditional forecasting models often struggle to deliver optimal performance [9]. Given these intricate challenges, the core of this study lies in enhancing the precision of load forecasting models specifically designed for campus buildings with solar PV.

Load forecasting is a technique energy providers use to predict the amount of power or energy required to ensure that demand and supply remain balanced [10]. The precision of these predictions holds vital importance for utility companies' operational and managerial aspects. Most forecasting techniques rely on numerical methods or Artificial Intelligence algorithms like regression, neural networks, and fuzzy logic [11]. Long-term, medium-term, and short-term are the time frames encompassed by load forecasting [12]. In the short term, which spans from one hour to a week, forecasts are instrumental in managing daily operational systems and unit commitment. Medium-term load forecasting, covering one week to a year, aids in tasks like fuel supply scheduling and unit management. Long-term load forecasting, on the other hand, extends beyond a year and involves predicting loads over extended periods. As short-term load forecasting provides efficiency in energy management, this study focuses on day-ahead load consumption forecasting, which falls into short-term categories.

The area of load forecasting primarily revolves around the utilization of time series models, as highlighted in literatures [8-11]. This forecasting approach encompasses both conventional techniques and modern machine learning methods, offering a diverse array of tools to predict future load patterns based on historical data. Seasonal Autoregressive Integrated Moving Average (SARIMA), Autoregressive Integrated Moving Average (ARIMA), Autoregressive Moving Average (ARMA), Moving Average (MA), and Autoregression (AR) are among the commonly used models in traditional time series forecasting methods. In-depth investigations into these methodologies are conducted in [12-14], specifically examining their utility in shaping load forecasting. One prominent model within this category, the Auto Regressive Integrated Moving Average (ARIMA), is particularly remarkable. This predictive approach examines the underlying laws governing time series correlations, making it a suitable tool for analysing load patterns and making future predictions. Nevertheless, acknowledging the limitations of the ARIMA model is crucial. This approach assumes linear relationships and stationary time series data, which might not adequately capture complex and nonlinear load variations [19].

Additionally, ARIMA may struggle with handling seasonal and irregular data patterns, potentially leading to less accurate forecasts in situations where these factors play a significant role. To overcome these challenges, the involvement of machine learning methods, particularly Artificial Neural Networks (ANN), emerges as a promising solution. These techniques have the capacity to learn intricate nonlinear relationships within data, enabling them to handle both complex load variations and irregular patterns [20]. By incorporating ANN into load forecasting practices, the shortcomings of traditional models like ARIMA can be effectively addressed, resulting in more accurate predictions of load consumption.

The author in [20] conducted a comprehensive analysis of household electricity consumption forecasting, comparing the performance of the ARIMA model and Artificial Neural Networks (ANN). The study sought to assess how effectively an Artificial Neural Network (ANN) captures the non-linear patterns in household electricity consumption, which are frequently difficult for traditional time series models such as ARIMA to handle. The results indicated that the ANN model significantly outperformed ARIMA, reducing the forecasting error by 30%. ANN excels in modelling the complex and non-linear dynamics of household electricity consumption, as evidenced by the significant reduction in error, making it a more reliable tool for accurate load forecasting.

A comprehensive examination involving Quantile Regression, Decision Tree, and Artificial Neural Network (ANN) methods is carried out in the literature [21]. The investigation revolves around household electricity consumption data from Cameroon, utilizing it for simulation purposes. The outcomes highlight that, when compared to Quantile Regression and Decision Tree, ANN yields notably improved scores in metrics, including Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). Additionally, the Artificial Neural Network (ANN) demonstrates a higher coefficient of determination (R^2) value. Similarly, in [22], the authors explore a comparison between ANN and traditional statistical methods, specifically linear regression, for the task of load forecasting in building contexts during both working and non-working days. The results indicate that the ANN model provides more accurate predictions of load consumption on working days.

Meanwhile, there is no significant accuracy difference between ANN and Linear Regression in the scenario of non-working days. The study in [23] conducted a short-term (1 hour ahead) load forecasting for educational buildings using various machine learning models, including Gradient Boosting Machine, XGBoost, and Random Forest. The results indicate that the performance of these techniques was not particularly significant. While ANN and other traditional machine learning techniques appear sufficient for load forecasting modelling, it is essential to acknowledge that various types of neural networks each have their distinct advantages. For instance, Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN) each possess specialized capabilities.

A Recurrent Neural Network RNN is an artificial neural network built to handle data sequences by creating connections that enable information to be transmitted from one step in the sequence to the subsequent steps [24]. Unlike normal ANNs, RNNs have loops within their architecture that enable them to keep information of memory in previous inputs. This built-in memory mechanism makes RNNs especially effective for tasks that involve sequential or time-dependent data, where the order of inputs and the context of previous inputs are crucial for accurate predictions or classifications.

The Long Short-Term Memory (LSTM) model is a recent development in RNN technology, created to tackle the gradient vanishing issue prevalent in conventional RNNs [25]. A comparative analysis undertaken in [18, 19] compares LSTM with Autoregressive Integrated Moving Average (ARIMA) in the domain of load forecasting, focusing solely on meteorological data as input variables. The findings indicate that LSTM exhibits superior performance to ARIMA in terms of Root Mean Square Error (RMSE), Mean Square Error (MSE), and Mean Absolute Percentage Error (MAPE) measurements. The Convolutional Neural Network (CNN) is commonly employed in image classification due to its ability to extract meaningful features from input data.

Nonetheless, significant investigations have ventured into the utilization of CNN for load forecasting. Specifically, studies detailed in [4, 26] have crafted CNN models to predict load consumption in small-scale integrated energy systems and residential areas situated in Sceaux, France. Impressively, the proposed models in both studies yielded positive outcomes, highlighting CNN's capacity to effectively extract valuable features from input data, thereby enhancing its applicability in load forecasting.

In [27], an investigation compares univariate time series load forecasting models that utilize CNN and LSTM techniques. Each setup includes models arranged in one, two, and three layers. The study encompasses both 1-day and 2-day-ahead load forecasting scenarios. The results reveal that in the case of 1-day ahead forecasting, the 1-layer CNN outperforms the 1-layer LSTM, displaying a lower RMSE by 0.4 compared to the LSTM. Conversely, for both 1-day and 2-day-ahead load forecasting, the 3-layer LSTM demonstrates an RMSE approximately 2.0 lower than the 3-layer CNN. It's worth noting that CNN possesses limitations when compared to other neural network types.

In contrast to RNN, which excels in handling sequential and time-dependent data due to its inherent memory capabilities, CNN lacks this recurrent memory mechanism [28]. This can potentially limit CNN's effectiveness in capturing temporal dependencies present in load forecasting data, where time sequences play a crucial role. A fusion of both CNN and RNN architectures could be proposed to mitigate this. By integrating the strengths of CNN's feature extraction with RNN's temporal understanding, a hybrid model could be created that harnesses the power of both approaches. This could lead to improved load forecasting accuracy by effectively capturing both spatial features and temporal patterns inherent in energy consumption data [29].

The research documented in the literature [29] introduces the concept of univariate time series load forecasting through a hybrid approach, combining CNN and LSTM methodologies specifically tailored for the Bangladesh Power System. Also, authors detailed in the literature [30] dive into Univariate time series load forecasting, employing a fusion of CNN and LSTM techniques. The dataset utilized for their study encompasses load consumption data from New South Wales, Australia, spanning the years 2006 to 2010, without incorporating solar integration. The primary objective is to predict 24-hour load consumption ahead.

Notably, the hybrid CNN-LSTM approach in previous studies [23, 24] showcases enhanced predictive accuracy compared to standalone LSTM models. Existing research on CNN-LSTM models for load forecasting primarily centres around residential and commercial buildings, often disregarding the inclusion of solar Photovoltaic (PV) systems. However, existing research on CNN-LSTM models for load forecasting has predominantly focused on residential and commercial buildings, often excluding solar Photovoltaic (PV) systems. A significant gap exists in investigating educational buildings, which face unique challenges due to varying load patterns during semester breaks, ongoing semesters, and special events. This underrepresentation in the educational sector highlights the need for more comprehensive research in load forecasting.

Moreover, there is a noticeable scarcity of studies utilizing advanced deep learning techniques, such as CNN-BiLSTM models, for day-ahead load forecasting in educational buildings. This gap in the literature underscores the novelty and importance of exploring CNN-BiLSTM models in this context, particularly with the inclusion of solar PV integration, to better address the complex load patterns characteristic of educational institutions.

This paper's key contributions can be encapsulated as follows:

- This study introduces an innovative approach to day-ahead load consumption forecasting for university buildings by employing a hybrid model that combines CNN and BiLSTM architectures. The aim is to effectively capture the complex load patterns that occur during semester running and semester breaks, addressing the unique challenges these distinct occasions pose.
- To enhance the accuracy of the proposed technique, an optimization process for hyperparameters has been integrated using the Bayesian optimization algorithm. This approach facilitates the identification of the best hyperparameter values, leading to enhanced performance and efficiency of the CNN-BiLSTM model.
- The study thoroughly compares ANN, BiLSTM, and CNN-BiLSTM models to observe their varying performances. This detailed analysis provides insights into the differences among these models.

In summary, the paper's novelty lies in its innovative hybrid CNN-BiLSTM approach, the integration of hyperparameter optimization using Bayesian optimization, and the detailed comparative analysis of different neural network models. These factors work together to enhance the accuracy of day-ahead load forecasting for university buildings, specifically dealing with the challenges of changing load patterns during different academic periods.

2. Methodology

Neural Networks (NNs) are primarily grounded in simplified mathematical representations of how we believe the human brain functions. Figure 1 illustrates that a neural network architecture typically comprises three or more layers: an input layer, an output layer, and one or more hidden layers.

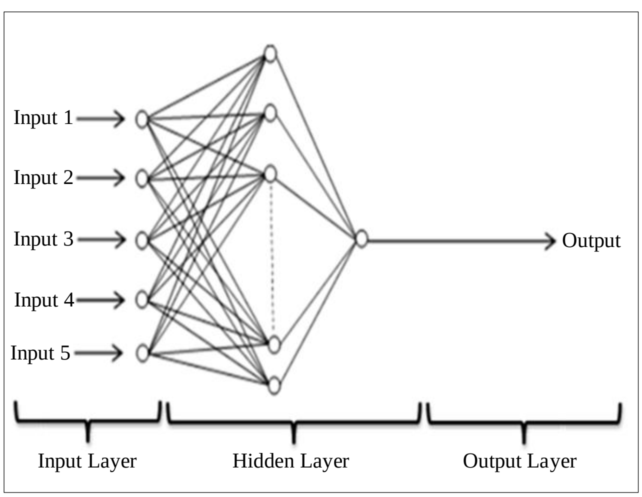

Fig. 1 The basic structure of the Neural Network (NN) model [31]

This section examines the models evaluated in the study, including BiLSTM, Convolutional Neural Network CNN-BiLSTM, and Artificial Neural Network (ANN). The fundamental concepts behind each model are outlined. Furthermore, the framework of the proposed model is comprehensively explained, offering a clear understanding of its design.

2.1. Artificial Neural Network (ANN)

An Artificial Neural Network (ANN) is a computational model modeled after the structure and operation of the human brain [32]. It consists of nodes, commonly known as neurons that are interconnected and organized into several layers. Each neuron processes inputs, applies weights to them, and generates an output. ANNs are designed to identify patterns and relationships within the data through training by adjusting the weights according to the input data and the desired output, enabling them to generate predictions, classifications, or other data-driven tasks [17]. At the core of an ANN lies the artificial neuron, arranged in layers and linked to neurons in different layers through connections referred to as synaptic weights. During the training process, a key goal is to determine and adjust these weights, facilitating the network's ability to learn and adapt. A neuron's activation is calculated by summing its weighted inputs, a process represented mathematically in Equation (1):

$$O = f\left(\sum(W_{ij}X_j)\right) \quad (1)$$

Where, O represents the output of the neuron, X_j is the input to that neuron, W_{ij} is the weight of the connection of the input to the neuron with f as a transfer function. The sigmoidal function is typically used in the transfer function of neural networks.

2.2. Bidirectional Long Short-Term Memory (BiLSTM)

Long Short-Term Memory (LSTM), a specialized form of Recurrent Neural Network (RNN), is engineered to address the problem of vanishing gradients. The LSTM architecture comprises essential components, including a cell, an input gate, an output gate, and a forget gate [33]. Figure 2 provides a detailed visualization of the LSTM cell diagram, highlighting the structure of the Long Short-Term Memory (LSTM) neural network. This diagram outlines the various components and their interactions within the LSTM, incorporating the input, forget, and output gates, along with the cell and hidden states. The forget gate is calculated mathematically as denoted in Equation (2), while Equation (3) denotes the input gate formula.

$$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f) \quad (2)$$

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i) \quad (3)$$

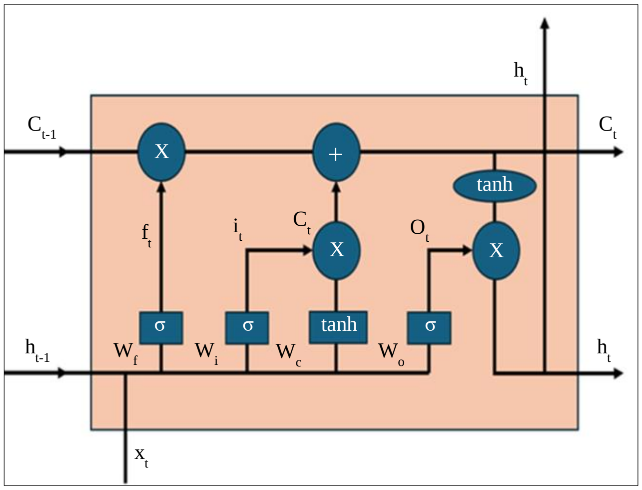

Fig. 2 LSTM cell diagram

$$\breve{C_t} = \tanh(W_c * [h_{t-1} - x_t] + b_c) \quad (4)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C_t} \quad (5)$$

The forget gate f_t , is essential to the LSTM's architecture by employing a sigmoid function to assess both the previously hidden state, denoted as h_{t-1}, and the current input at the time step, represented as x_t. Its primary function is to selectively eliminate information deemed less important or irrelevant. Within the input gates layer i_t, a sigmoid (σ) operation serves the purpose of determining which values should undergo an update, as outlined in Equation (3). Subsequently, a hyperbolic tangent (tanh) layer produces a fresh vector of candidate values, denoted as $\widetilde{C_t}$, which may potentially be incorporated into the state in accordance with Equation (4). The existing cell state, C_{t-1}, is then replaced with the newly computed cell state, C_t, following the principles set forth in Equation (5).

$$o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = o_t * tanh(C_t) \quad (7)$$

The output gate (o_t) functions by generating an output determined by the cell state. A sigmoid layer is employed to decide which segment of the cell state will be emitted, as indicated in Equation (6). To determine which segment of the cell state will be emitted, the cell state undergoes a transformation through the hyperbolic tangent (tanh) function, which scales values to range between -1 and 1 [34].

The result is then multiplied by the output from the sigmoid gate, as detailed in Equation (7), h_t is the new updated hidden state. Where, $W_f, W_c, W_{i,}$ and W_o are the rectangular weight arrays for forget gate (f), cell state (c), input gate (i) and output gate (o). While b_f, b_i, b_c and b_o are the bias vectors, and sigma is the logistic sigmoid for each respective gate and cell state.

A BiLSTM is a neural network that utilizes two LSTM layers, where one layer processes the sequence from beginning to end, and the other layer processes it from end to beginning. The core idea behind a BiLSTM is to allow the network to access information from both past and future parts of the input sequence.

The forward LSTM layer processes data in the usual chronological sequence, while the backward LSTM layer handles data in reverse chronological order. [35]. These two LSTM layers are connected and interact with each other, allowing the BiLSTM model to build a more comprehensive understanding of the input data compared to a standard unidirectional LSTM [36]. The network receives the input data in both the normal time sequence and the reverse time sequence, providing a richer representation of the sequential information.

This bidirectional flow of information through the two LSTM layers allows the BiLSTM model to better model and learn from sequential data, as it can leverage both past and future context at each time step [36]. The final output of the BiLSTM is a combination of the outputs from the forward and backward LSTM layers. In summary, the BiLSTM architecture harnesses the strengths of two complementary LSTM layers operating in opposite directions to capture a more complete understanding of the input sequence.

2.3. Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) stands out as a unique class of neural network designed with the capability to extract important elements from input data. It operates as a feedforward neural network featuring a multi-layer structure incorporating convolutional computations. Traditionally, CNNs have been employed for image classification tasks [29]. However, their adaptability extends to time-series forecasting applications.

The framework of CNN includes several essential components: convolutional layers, pooling layers, flattening layers, and fully connected layers. Convolutional layers take the input data, and a filter is applied to generate a feature map. An activation function is subsequently applied to the result. By leveraging information from the convolutional layer, the pooling layer decreases the size of the feature map. Finally, the flattening layer converts the reduced feature map into a one-dimensional array, preparing it for the subsequent fully connected layer. In this layer, weights are applied to process the data efficiently.

2.4. Data Collection

This section provides a comprehensive explanation of the data collection process. The dataset in question was obtained from UiTM Permatang Pauh, specifically selected as a case study due to the presence of its substantial 2.8 MW solar facility.

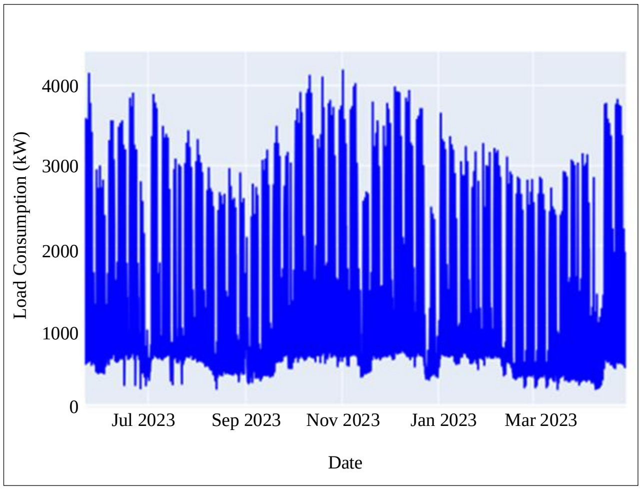

Fig. 3 Load consumption data of UiTM Permatang Pauh campus

Figure 3 visually presents the load data for UiTM Permatang Pauh, encompassing a total of 343 days ranging from May 22, 2023, until April 30, 2024. The data is at 30-minute intervals, for a total of 16464 data points. Leveraging this load data, comprehensive feature engineering was conducted, leading to the generation of categorical data such as hour, calendar attributes, and semester status. These additional data can further the analysis of developing an accurate model.

2.5. Overall Modelling Framework

This section provides an in-depth explanation of the overall modelling framework, which includes our proposed load forecasting model designed specifically for campus buildings. The proposed model combines the strengths of CNN and BiLSTM, with CNN serving as the foundational layer, followed by an additional seven layers of BiLSTM. This unique layering strategy is summarized to give a concise overview of the model's architecture.

Figure 4 hows the overall modelling framework. For model training, 70% of the collected data is utilized, featuring various input variables such as the day, hour, lagged load consumption from the previous week and day, the presence of a public holiday, the current semester status (whether it's lecture or office), and any ongoing semester breaks. These parameters are leveraged to predict the target variable, the day-ahead load consumption.

The ANN model consists of 8 layers, which have 100 neurons for the first layer, 100 for the second layer, 80 for the third layer, 120 for the fourth layer, 100 for the fifth layer, 30 for the sixth layer, 90 for the seventh layer and 1 for the eighth layer. At the same time, the BiLSTM model consists of 7 layers, which have 180 neurons for the first layer, 80 for the second and third layers, 50 for the fourth layer, 10 for the fifth layer, 15 for the sixth layer and 1 for the seventh layer. The proposed CNN-BiLSTM model consists of 128 filters of 2D CNN with 4 kernel sizes and 7 BiLSTM layers, like previous BiLSTM used. These configurations were obtained by using the Bayesian Optimization Algorithm.

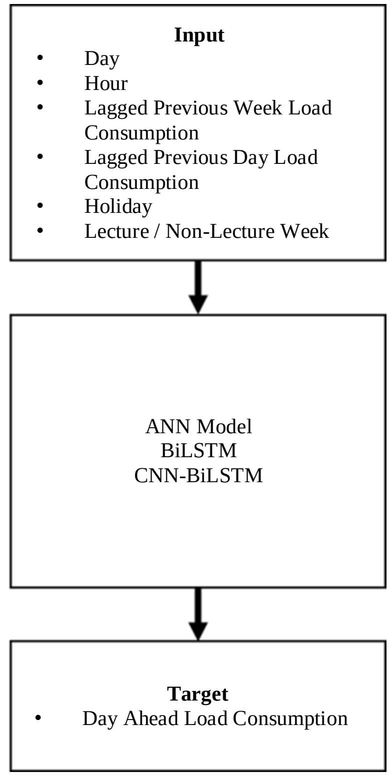

Fig. 4 Overall modelling framework

The ANN, LSTM, and CNN-BiLSTM models were all trained using identical input data. Subsequently, these trained models were assessed using a testing dataset, comprising 30% of the total dataset, to forecast day-ahead load consumption. The results were evaluated based on several metrics, including Mean Absolute Percentage Error, Root Mean Square Error (RMSE) and Mean Square Error (MSE). The formula for evaluation metrics is shown in Equation (8)-(11), where A_i represents the actual value and P_i is the predicted value.

$$MAPE = \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{A_i - P_i}{P_i}\right| \quad (8)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(A_i - P_i)^2} \quad (9)$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(A_i - P_i)^2 \quad (10)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{A_i - P_i}{P_i}\right| \quad (11)$$

Figure 5 illustrates the CNN-LSTM framework proposed in this study. The CNN component serves as a feature extractor tasked with capturing load patterns based on the provided labels. The BiLSTM layer then utilizes the features extracted by the CNN. It conducts the sequence learning to capture temporal dependencies, ultimately leading to the generation of more accurate output with minimized errors.

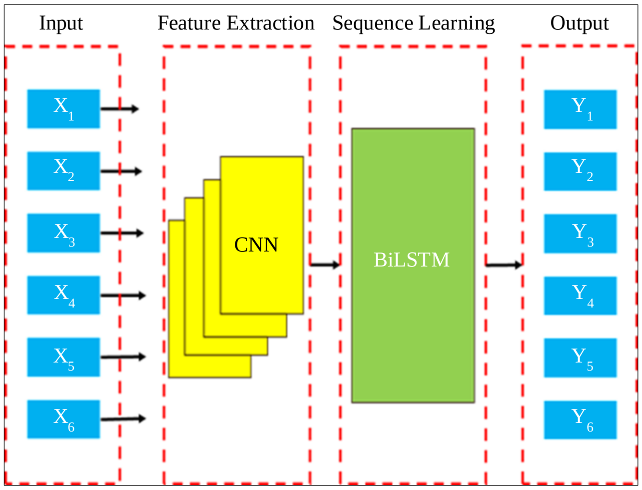

Fig. 5 Proposed model framework

3. Results and Discussion

This section presents the results of the data analysis and model testing phase, where the model architecture was obtained by implementing hyperparameter optimization. This section also illustrates the evaluation scores, as well as the forecast and actual plot. An analysis result of load consumption pattern during weekdays and weekends on lecture week and non-lecture week is shown in Figure 6 and Figure 7. The blue line in the plot shown is for load during lecture week, and the red line is during non-lecture week. The load consumption during lecture weeks is higher than in non-lecture weeks on weekdays and weekends. Detailed statistical analysis is shown in Tables 1and 2. These tables present load consumption statistics for weekdays and weekends during lecture and non-lecture weeks.

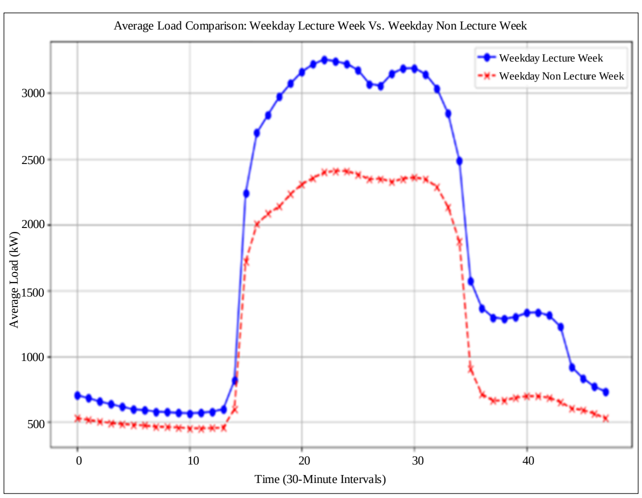

Fig. 6 Average load comparison for weekday

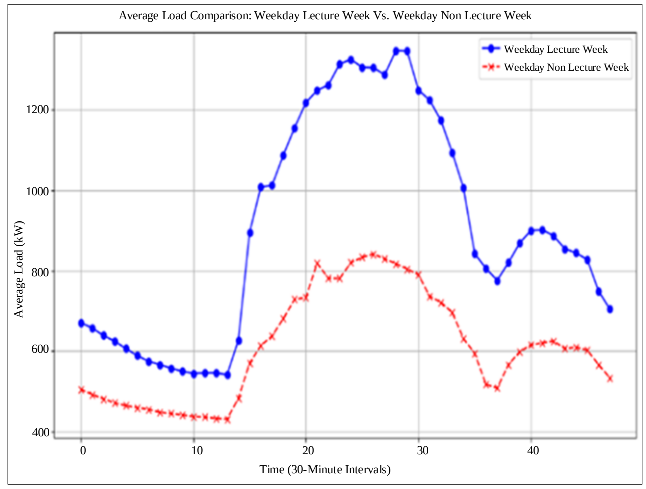

Fig. 7 Average load comparison for the weekend

On weekdays, the average load consumption during lecture weeks is 1766.95 kW, significantly higher than the 1263.95 kW recorded during non-lecture weeks. This suggests that energy demand is higher when academic activities are ongoing. Additionally, the minimum and maximum loads during lecture weeks are 565.55 kW and 3255.73 kW, respectively, compared to 450.96 kW and 2408.12 kW during non-lecture weeks, indicating greater fluctuations in load consumption when lectures are in session.

A similar pattern is observed on weekends, with the average load during lecture weeks at 905.89 kW, higher than the 611.04 kW during non-lecture weeks. The minimum load during lecture weekends is 541.64 kW, while non-lecture weekends see a lower minimum of 431.60 kW. The maximum load also follows this trend, reaching 1346.65 kW during lecture weeks and 841.47 kW during non-lecture weeks.

Table 1. Statistical description for weekday load profile

Statistic	Weekday Lecture Week	Weekday Lecture Week
Mean	1766.95 kW	1263.95 kW
Minimum	565.55 kW	450.96 kW
Maximum	3255.73 kW	2408.12 kW

Table 2. Statistical description for weekend load profile

Statistic	Weekend Lecture Week	Weekend Lecture Week
Mean	905.89 kW	611.04 kW
Minimum	541.64 kW	431.60 kW
Maximum	1346.65 kW	841.47 kW

Table 3. Optimized model's architecture

Model	Neurons per Layer	Learning Rate
ANN	100,100,80,120,100,30,90,1	0.001
BiLSTM	180,80,80,50,10,15,1	0.001
CNN-BiLSTM	128,4,180,80,80,50,10,15,1	0.001

Overall, the data shows that both weekdays and weekends during lecture weeks consistently have higher load consumption and larger fluctuations compared to non-lecture weeks, highlighting the significant impact of academic activities on energy usage patterns. Table 3 illustrates the model architecture for ANN, BiLSTM and CNN-BiLSTM. All the hyperparameters are obtained by using the Bayesian Optimization algorithm. All three models have a similar batch size, 32, but different neurons and the number of layers. The number of neurons for each model is illustrated in sequence, starting from the first layer and continuing to the last layer.

Table 4. Performance evaluation of models

Model	RMSE (kW)	MSE (kW)	MAE (kW)	MAPE (%)
ANN	240.94	62973.20	187.59	13.03
BiLSTM	198.12	42772.47	144.57	8.77
CNN-BiLSTM	165.87	34067.52	115.23	6.99

Table 4 illustrates the model performances based on RMSE, MSE, and MAE. The CNN-BiLSTM model delivers improved results over the ANN and BiLSTM, with a MAPE score of 6.99%. The higher error observed with the ANN in this finding can be attributed to its simpler algorithm compared to the more complex models, which limits its ability to capture intricate patterns in the data.

Figures 8, 9, and 10 present the actual versus forecasted plots for the ANN, BiLSTM, and CNN-BiLSTM models. The figure accompanying this analysis illustrates a 7-day plot comparing the actual versus forecasted load. The time steps between 0-250 correspond to weekdays, while the range from 250-360 represents the weekend. During the weekdays, the proposed model maintains close alignment with the actual load values, demonstrating its robustness in handling fluctuations in working-day demand.

As we move into the weekend, slight variations are observed, yet the model still performs accurately, confirming its ability to generalize across different days of the week. This visual representation supports the conclusion that the CNN-BiLSTM model is highly effective for load forecasting tasks, especially in environments with varying load patterns.

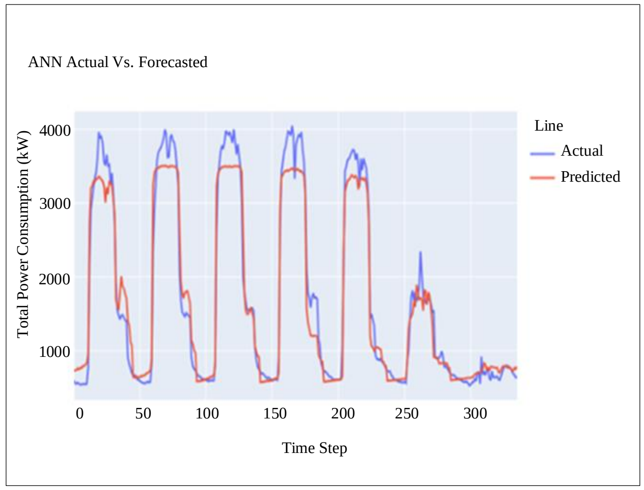

Fig. 8 ANN actual vs forecasted plot

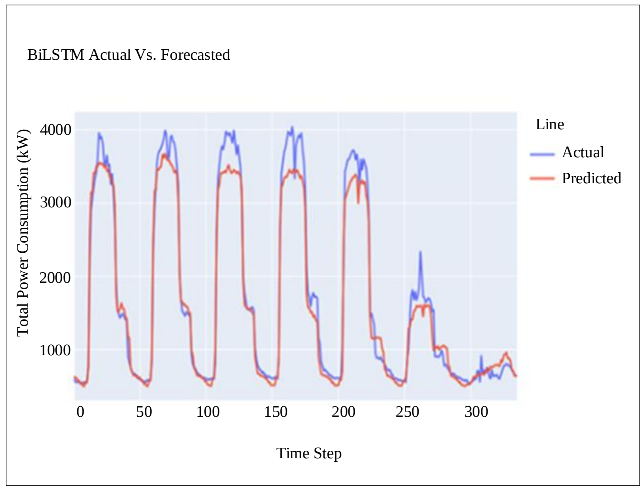

Fig. 9 BiLSTM actual vs forecasted plot

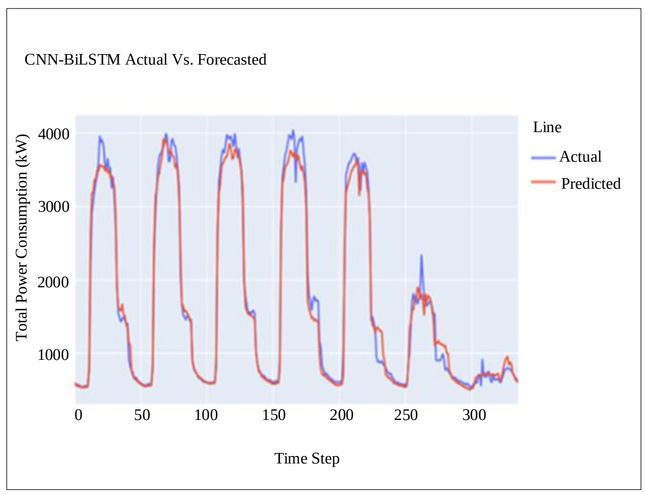

Fig. 10 CNN-BiLSTM actual vs forecasted plot

In comparison, the plots for the BiLSTM and ANN models reveal noticeable under-forecasting, especially on weekdays. Both models predict lower load values than the actual data, particularly during peak load periods. The ANN model shows the most significant discrepancy, attributed to its simpler algorithm, lacking the depth and complexity of the CNN-BiLSTM model. Meanwhile, the BiLSTM performs better than ANN but still struggles to match the predictive accuracy of CNN-BiLSTM. This comparison highlights that the CNN-BiLSTM model's superior architecture effectively addresses the limitations encountered in models like ANN and BiLSTM, particularly in environments with varied load patterns.

4. Conclusion

In conclusion, this paper explored the application of various neural network models for load consumption forecasting. Specifically, we examined the performance of ANN, BiLSTM, and CNN-BiLSTM. Notably, the proposed CNN-BiLSTM model emerged as the frontrunner in this analysis, outperforming both ANN and BiLSTM by achieving a remarkable MAPE value of 6.99%. This observation underscores the critical importance of accuracy in load forecasting models, as it directly enhances energy management systems. As we address the ever-growing energy demands, the utilization of advanced neural network architectures like CNN-BiLSTM proves to be a promising avenue for achieving more precise load forecasts and optimizing energy resource allocation. This study has the potential to significantly contribute to energy management strategies, particularly in academic building environments.

Acknowledgments

Gratitude is extended to the facility of UiTM Shah Alam for providing their property as the study site. Appreciation is also given to colleagues and collaborators for their valuable support and assistance throughout the research process.

References

[1] Minan Tang et al., "Short-Term Load Forecasting of Electric Vehicle Charging Stations Accounting for Multifactor IDBO Hybrid Models," *Energies*, vol. 17, no. 12, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Zulfiqar Ahmad Khan et al., "Efficient Short-Term Electricity Load Forecasting for Effective Energy Management," *Sustainable Energy Technologies and Assessments*, vol. 53, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Altaf Hussain et al., "A Hybrid Deep Learning-Based Network for Photovoltaic Power Forecasting," *Complexity*, vol. 2022, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] Mosbah Aouad et al., "A CNN-Sequence-to-Sequence Network with Attention for Residential Short-Term Load Forecasting," *Electric Power Systems Research*, vol. 211, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] J.S. Hassan et al., "Building Energy Consumption in Malaysia: An Overview," *Jurnal Teknologi (Sciences & Engineering)*, vol. 70, no. 7, pp. 33-38, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[6] K.P.M Kementerian Pelajaran Malaysia, Malaysia Educational Statistic 2020, 2020.

[7] Yisheng Cao et al., "PSO-Stacking Improved Ensemble Model for Campus Building Energy Consumption Forecasting Based on Priority Feature Selection," *Journal of Building Engineering*, vol. 72, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Rijalul Fahmi Mustapa et al., "Non-Linear and Linear Baseline Energy Modelling Comparative Studies in Educational Buildings," *International Journal of Engineering & Technology*, vol. 7, no. 4.22, pp. 61-66, 2018. [Publisher Link]

[9] Neelam Mughees et al., "Deep Sequence to Sequence Bi-LSTM Neural Networks for Day-Ahead Peak Load Forecasting," *Expert Systems with Applications*, vol. 175, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] Gokcen Ozdemir et al., "The Role of eXplainable Artificial Intelligence (XAI) in Smart Grids," *Big Data Application in Power Systems*, pp. 349-370, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Serge Guefano et al., "Forecast of Electricity Consumption in the Cameroonian Residential Sector by Grey and Vector Autoregressive Models," *Energy*, vol. 214, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[12] Bobo Wang et al., "Ss5:A Neural Network-Based Energy Consumption Prediction Model for Feature Selection and Paremeter Optimization of Winders," *2020 IEEE International Conference on Networking, Sensing and Control (ICNSC)*, Nanjing, China, pp. 1-5, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[13] Tahreem Anwar et al., "Introduction to Load Forecasting," *International Journal of Pure and Applied Mathematics*, vol. 119, no. 15c, pp. 1527-1538, 2018. [Google Scholar] [Publisher Link]

[14] Yousuf M.S. Khara, Asad Mahmood, and Kafait Ullah, "Load Forecasting of an Educational Institution Using Machine Learning: The Case of Nust, Islamabad," *Pakistan Journal of Science*, vol. 71, no. 4, pp. 252-257, 2019. [Google Scholar]

[15] Y. Wang, "Short-Term Power Load Forecasting Based on Machine Learning," 2020. [Google Scholar]

[16] Wang Hao, Hayinaer Tueraili, and Song Bin, "Refined Load Forecasting of Distribution Network Based on ARIMA Model with Multi-Factor Combination Analysis," *2021 IEEE 4th International Electrical and Energy Conference (CIEEC)*, Wuhan, China, pp. 1-4, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] A.A. Komyakov, V.V. Erbes, and V.I. Ivanchenko, "Application of Artificial Neural Networks for Electric Load Forecasting on Railway Transport," *2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC)*, Rome, Italy, pp. 43-46, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[18] Lingling Tang, Yulin Yi, and Yuexing Peng, "An Ensemble Deep Learning Model for Short-Term Load Forecasting Based on ARIMA and LSTM," *2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, Beijing, China, pp. 1-6, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[19] Innocent Mpawenimana et al., "A Comparative Study of LSTM and ARIMA for Energy Load Prediction with Enhanced Data Preprocessing," *2020 IEEE Sensors Applications Symposium (SAS)*, Lumpur, Malaysia, pp. 1-6, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[20] Chafak Tarmanini et al., "Short Term Load Forecasting Based on ARIMA and ANN Approaches," *Energy Reports*, vol. 9, pp. 550-557, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Jean Calvin Nsangou et al., "Explaining Household Electricity Consumption Using Quantile Regression, Decision Tree and Artificial Neural Network," *Energy*, vol. 250, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[22] Moon Keun Kim, Yang-Seon Kim, and Jelena Srebric, "Predictions of Electricity Consumption in a Campus Building Using Occupant Rates and Weather Elements with Sensitivity Analysis: Artificial Neural Network vs. Linear Regression," *Sustainable Cities and Society*, vol. 62, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[23] Jihoon Moon, Seungmin Rho, and Sung Wook Baik, "Toward Explainable Electrical Load Forecasting of Buildings: A Comparative Study of Tree-Based Ensemble Methods with Shapley Values," *Sustainable Energy Technologies and Assessments*, vol. 54, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[24] Lakhmi C. Jain, and Larry R. Medsker, *Recurrent Neural Networks (Design and Applications)*, CRC Press, Inc., 1999. [Google Scholar] [Publisher Link]

[25] Sepp Hochreiter, and Jürgen Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997. [CrossRef] [Google Scholar] [Publisher Link]

[26] Ge Zhang, Xiaoqing Bai, and Yuxuan Wang, "Short-Time Multi-Energy Load Forecasting Method Based on CNN-Seq2Seq Model with Attention Mechanism," *Machine Learning with Applications*, vol. 5, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[27] First Ali Agga et al., "Short-Term Load Forecasting Based on CNN and LSTM Deep Neural Networks," *IFAC-PapersOnLine*, vol. 55, no. 12, pp. 777-781, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[28] Liu Wenya, "Cooling, Heating and Electric Load Forecasting for Integrated Energy Systems Based on CNN-LSTM," *2021 6th International Conference on Power and Renewable Energy (ICPRE)*, Shanghai, China, pp. 808-812, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[29] Shafiul Hasan Rafi et al., "A Short-Term Load Forecasting Method Using Integrated CNN and LSTM Network," *IEEE Access*, vol. 9, pp. 32436-32448, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[30] Xifeng Guo et al., "A Short-Term Load Forecasting Model of Multi-Scale CNN-LSTM Hybrid Neural Network Considering the Real-Time Electricity Price," *Energy Reports*, vol. 6, no. 9, pp. 1046-1053, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[31] Phornnara Nop, and Zhijun Qin, "Transmission System Load Forecasting Using Neural Network Approach, Case Study: Electricité du Cambodge (EDC)," *2020 IEEE Student Conference on Electric Machines and Systems (SCEMS)*, Jinan, China, pp. 329-333, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[32] Lemuel Clark P. Velasco, Karl Anthony S. Arnejo, and Justine Shane S. Macarat, "Performance Analysis of Artificial Neural Network Models for Hourahead Electric Load Forecasting," *Procedia Computer Science*, vol. 197, pp. 16-24, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[33] Tasarruf Bashir et al., "Short Term Electricity Load Forecasting Using Hybrid Prophet-LSTM Model Optimized by BPNN," Energy Reports, vol. 8, pp. 1678-1686, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[34] Nirmala C.R. et al., "Forecasting Price of the Crude Oil Using LSTM Based on RNN," *International Journal for Research in Applied Science & Engineering Technology*, vol. 10, no. 7, pp. 196-206, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[35] Maged Mamdouh, Mostafa Ezzat, and Hesham Hefny, "Improving Flight Delays Prediction by Developing Attention-Based Bidirectional LSTM Network," *Expert Systems with Applications*, vol. 238, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[36] Umar Javed et al., "A Novel Short Receptive Field Based Dilated Causal Convolutional Network Integrated with Bidirectional LSTM for Short-Term Load Forecasting," *Expert Systems with Applications*, vol. 205, 2022. [CrossRef] [Google Scholar] [Publisher Link]