A short-term hybrid forecasting model for time series electrical-load data using random forest and bidirectional long short-term memory

Zannatul Ferdoush[1], Booshra Nazifa Mahmud[2], Amitabha Chakrabarty[3], Jia Uddin[4]
[1,2,3]Department of Computer Science and Engineering, Brac University, Bangladesh
[4]Technology Studies Department, Endicott College, Woosong University, Republic of Korea

Article Info

ABSTRACT

In the presence of the deregulated electric industry, load forecasting is more demanded than ever to ensure the execution of applications such as energy generation, pricing decisions, resource procurement, and infrastructure development. This paper presents a hybrid machine learning model for short-term load forecasting (STLF) by applying random forest and bidirectional long short-term memory to acquire the benefits of both methods. In the experimental evaluation, we used a Bangladeshi electricity consumption dataset of 36 months. The paper provides a comparative study between the proposed hybrid model and state-of-art models using performance metrics, loss analysis, and prediction plotting. Empirical results demonstrate that the hybrid model shows better performance than the standard long short-term memory and the bidirectional long short-term memory models by exhibiting more accurate forecast results.

Corresponding Author:

Jia Uddin,
Technology Studies Department, Endicott College,
Woosong University,
Daejeon, Republic of Korea.
Email: jia.uddin@wsu.ac.kr

1.   INTRODUCTION

Short term load forecasting (STLF) is a crucial instrument for ensuring the supply of high-quality electricity to customers in a systematic economic manner. Forecasts of electricity serve as the basis of many operating decisions taken by the electric utilities such as resource acquisition, dispatch scheduling of generating capacity, plant maintenance schedule, reliability analysis, and efficient power plant design [1]. The necessity of reliable load forecasting is more significant than ever as most of the energy markets around the world are becoming deregulated [2]. In deregulated electricity markets, energy transactions and price of electricity are critically conditional on forecast electric loads. The operational cost may rise remarkably as an outcome of underestimating the energy demand [3].

Load forecasting is a multivariate time-series prediction problem and generally divided into three categories: short-term, medium, and long-term forecasts. The STLF serves as the base for the safe and economic operation of power systems by providing a valuable estimation of the load for the next half hour up to the next two weeks. In earlier, several forecasting models such as statistical analysis, time-series analysis, Box Jenkins [4], and regression analysis were used to construct noteworthy forecasting results. The regression-based approach is one of the earliest and widely used statistical techniques. In the statistical methods such as holt-winter exponential smoothing (ES) [5] and auto-regressive integrated moving average

(ARIMA) [6], we identify the dominant variables based on the correlation analysis with the load for forecasting. As statistical approaches struggled to achieve efficient results with the highly non-linear load data, researchers gradually moved towards the machine learning (ML), artificial intelligence (AI), and a hybrid of statistical and AI-based learning models [7] for forecasting. In recent years reserchers have intensively explored ML/AI-based algorithms like random forest (RF) [8], support vector regression (SVR) [9], boosting algorithms [10], fuzzy logic [11], and artificial neural networks (ANN) [12]. An introduction of adaptive pattern recognition and self-organizing techniques for STLF gained significant attention [13] and later developed into an adaptive neural network for short-term forecasting [14]. Moon *et al.* [15] categorized the load data using a decision tree and later fabricated a hybrid model using the RF and multilayer perceptron. Among all the forecasting methods, ANN received noteworthy acknowledgment as it brought forth advanced prediction systems through feeding patterns of information via the input unit.

Recurrent neural network (RNN), widely implemented for forecasting purposes, struggles to solve time-series issues due to vanishing and exploding gradient problems. Researchers started exploring the long short-term memory (LSTM) network to overcome this challenge and successfully created several load forecasting models applying the LSTM network. Marino *et al.* [16] employed an advanced load forecasting methodology using two different LSTM architectures producing promising results. Bouktif *et al.* [17] proposed an LSTM model using the feature selection and genetic algorithm. However, these deep learning methods can be improved further by extending the number of works with various forecasting setups [18, 19].

As of recent, both RF and LSTM have exhibited tremendous potential in the forecasting of time series data. However, RF faces a significant drawback during its execution when the increase in the sample does not improve its prediction accuracy. Also, the LSTM tends to overfit as well as execution time slows down when the number of parameters increases. This paper presents a hybrid model that takes advantage of RFs ability to segregate vital features for a forecast by isolating variables of the least significance. By omitting the trivial features lacking in assisting forecast, the hybrid model reduces the parameter quantity and enhances the bi-LSTM model to be simpler and faster to fit and predict.

2.   RESEARCH METHOD

This section explains the details of the proposed model that includes data analysis and the proposed hybrid model simulation.

2.1. Casual structure of proposed RF-bi-LSTM hybrid model

Figure 1 presents the overall architecture of the proposed hybrid model. In the first layer of the model, the RF is used to identify the prediction strength of the features in the electricity consumption dataset. To locate the dominant variables, we measure the total decline in node impurity using the mean decrease impurity (MDI) method. The MDI counts the events when the model uses a feature to branch out a node, weighted by the number of samples it splits across all decision trees of the ensemble.
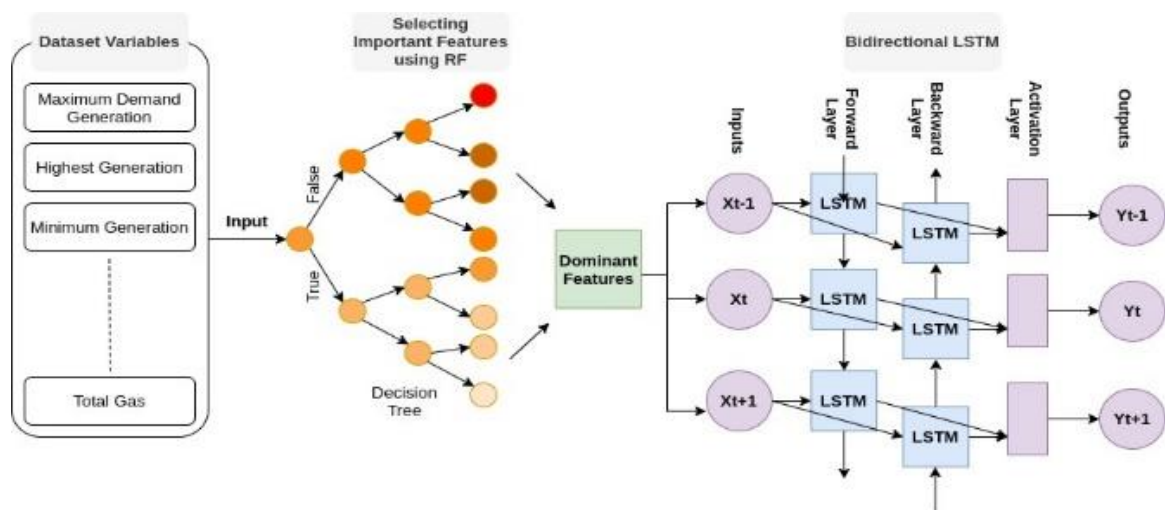

Figure 1. Proposed RF-bi-LSTM hybrid model

In the second layer, we utilized the bi-LSTM architecture. A memory block in the LSTM theoretically consists of four parts an input gate i, a forget gate f, an output gate o, and self-connected memory cells C.

$$F_t = (\sigma(W_f S_t - 1 + W_f X_t)) \tag{1}$$

$$i_t = (\sigma(W_i S_t - 1 + W_i X_t)) \tag{2}$$

$$o_t = \sigma(W_o S_t - 1 + W_o X_t) \tag{3}$$

$$C_{t\prime} = tanh(W_c S_t - 1 + W_c X_t) \tag{4}$$

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

In (1-5), the variables represent the following meanings: F_t = Forget gate, i_t = Input gate, o_t = Output gate, $C_{t\prime}$ = Intermediate cell state, and C_t = Cell state. These are the equations of gates and cell states of the LSTM.

$$h_t = o_t * tanh(C_t) \tag{6}$$

The LSTM has its unique trait of preserving information that previously passed through it by utilizing its hidden units [20, 21]. In (6), h_t represents the hidden state of the LSTM architecture. The Bi-LSTM model consists of two discrete LSTM networks where one access information in the forward direction and another access in the reverse direction [22]. This process helps to store knowledge from inputs of future and past at any point in time using two combined hidden states, and generate accurate output from the context of both the past and future data. Lastly, to generate forecast results, we feed the dominant features identified in the previous layer as the input to the bi-LSTM layer.

2.2. Dataset generation process

To evaluate the performance of the proposed model, we have used a dataset that is related to Bangladeshi electricity consumption data available at https://www.bpdb.gov.bd. For preparing the dataset, we utilized algorithm 1 that converts vital information of historical electrical load data from the PDF files to CSV format.

Algorithm 1: Data Extraction and Dataset Generation

```
1   Input: Start and end date of the PDF reports
2   Output: A CSV file with extracted data from all the PDF files between start and end
    date
3   for each date between the start date and end date do
4       download the PDF of this date and store into a specific folder

5   extracted_data ← []
6   for each PDF report from the downloaded reports do
7       extracted_row ← set of extratcted requisite data using regular expressions
8       append the extratcted_row into extracted_data

9   save extratcted_data as a CSV file
```

The automated tool to generate historical electricity consumption data consists of two subunits; the first part is a data extractor program that uses regular expressions (REs) to parse the PDF formatted files to extract specific information. The second part is called a CSV loader program that transfers the extracted data appended in a list into a CSV file. The data extracted for the period of June 1, 2015, to June 30, 2018, having a total of 1126 rows holds essential information from daily electricity consumption reports of mentioned 36 months.

2.3. Preliminary data analysis

To understand the different characteristics of the time-series, we analyzed the data to find the effects of the properties like seasonality, auto-correlation, and stationarity on the dataset. Figure 2 shows the decomposition factors of energy consumption data which reflects irregular trends and seasonality in our multivariate time-series data.
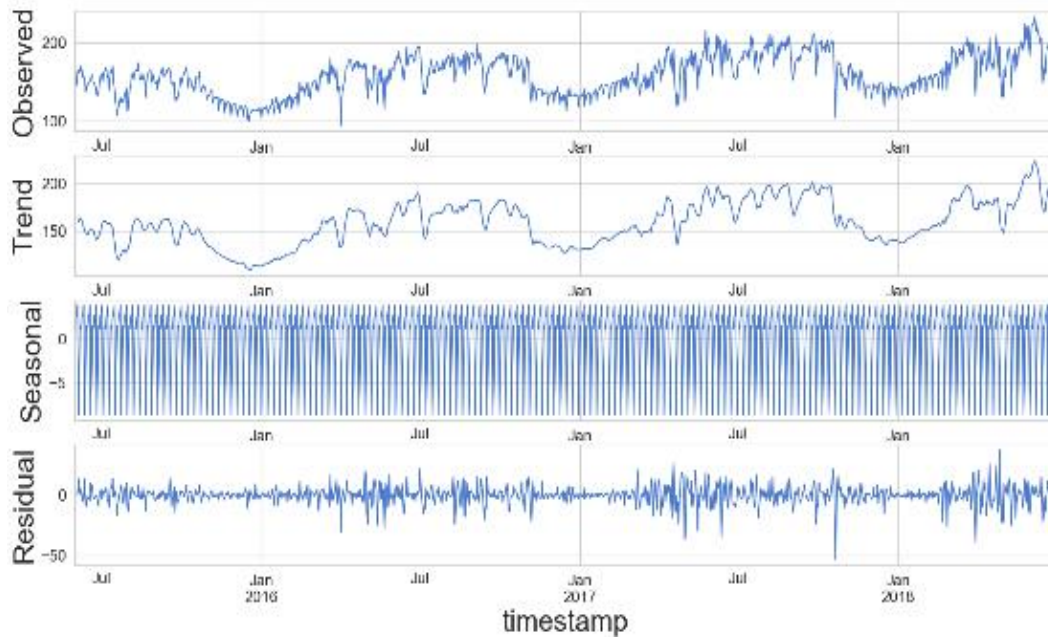
Figure 2. Time series decomposition of energy consumption data

We discerned that the electricity consumption has steadily increased over the years due to the stable economic growth as well as the rapid increase in the population [23]. The seasons in Bangladesh mainly fall into three categories; summer, rainy, and winter seasons. The power consumption rate naturally tends to be lower during the winter season that signifies the first and fourth quarters of the year. During the summer and the rainy seasons in the second and third quarters, the electricity consumption steeply progresses over the years. We observe that the rate drastically declines again when advancing through the winter season. These observations indicate the presence of repetitive short term seasonal cycles throughout our dataset.

Forecasts using the non-stationary time-series data may generate spurious conclusions by surmising false connections among variables. When a time series data following the time argument has invariant finite-dimensional distributions, it is called stationary data [24]. We employed tests like summary statistics and augmented dickey-fuller to detect stationarity in time-series data. In the summary statistics test, we split our data into two partitions and determined the mean and variance for both segments to compare. As depicted in Table 1, the mean and variance generated are not uniform over time, presenting our dataset to be non-stationary. Applying the Augmented Dickey-Fuller algorithm, we assessed the result using the p-value from the test. From Table 1, it is apparent that the series is non-stationary as p-value and ADF statistic both are greater than the threshold 5% significance level, and therefore we fail to reject the null hypothesis that unit root does exist. From the preliminary data analysis, we can conclude that the presence of a changing trend, seasonality, and non-stationarity of our data presents complex traits that can create unstable and deceptive results and direct to unreliable forecast results.

Table 1. Statistical values

No.	Type of Test	Generated Values
1	Summary Statistics	mean_1 = 150.51, mean_2 = 181.35, variance_1 = 458.79, variance_2 = 880.20
2	Augmented Dickey Fuller Test	ADF Statistic = -2.071, p-value = 0.257 Critical Values: 1%: -3.435, 5%: -2.864, 10%: -2.568

2.4. Proposed hybrid model simulation

The proposed RF-bi-LSTM model consists of two layers. To simulate the first layer of the hybrid model, we constructed an RF forecasting model that finds the variables with the highest importance for predicting forecasts. The initial input to the RF model is a complete set of features comprising time and electricity generation related factors mentioned in Table 2, and Table 3 describes the quantitative analysis of the power consumption dataset. We had to convert our input into a stationary dataset using the differentiation

method as from subsection 2.3. we learned our dataset to be non-stationary. Table 4 shows the parameter values used to train deep learning models. Through meticulous testing, we empirically decided the hyperparameter values to secure the most stable performance of the models. At first, we fit the model with default parameters to get a baseline idea of the performance, and later we confirmed the best values for the hyperparameters by increasing or decreasing them.

Table 2. The features of energy consumption dataset

No.	Attribute	Description
1	day_of_week	An integer value between 0 and 6
2	day_of_month	An integer value between 1 and 31
3	month	An integer value between 1 and 12
4	season	An integer value between 0 and 2
5	is_weekend	An integer value between 0 and 1
6	max_demand_gen	Maximum demand of power at generation end (in MW)
7	highest_gen	The highest amount of power generated at a specific hour in a day (in MW)
8	min_gen	The lowest amount of power generated at a specific hour in a day (in MW)
9	day_peak_gen	Total power generated at peak hour during daytime (in MW)
10	eve_peak_gen	Total power generated at peak hour during night time (in MW)
11	eve_peak_load_shedding	The amount of load-shedding at peak hour during night time (in MW)
12	max_temp	Maximum temperature of the particular day (℃)
13	total_gas	The amount of gas used to produce electricity in a day (in MMCFD)
14	total_energy	The total amount of energy consumed in a day (generated locally + imported from India) in a particular day (in MKWh)

Table 3. Quantitative form of the power consumption dataset

Attribute	timestamp	max_demand_gen (MW)	Highest_gen (MW)	min_gen (MW)	day_peak_gen (MW)	eve_peak_gen (MW)	eve_peak_load_shedding (MW)	max_temp (℃)	total_gas (MMCFD)	total_energy (MKWh)
Max	30/06/2018	10958	10958	8521.7	9178.2	10958	904	39	1152.3	233.9
Min	01/06/2015	5413	5413	1759	3634	5413	0	17.8	0	93.3
Mean		8045.6	8046.8	5426.7	6476.3	8017.6	22.7	31.3	958	159.2
Std. Dev		908.5	894.6	1088.5	1009.5	905.3	86.4	3.4	85.7	25.2

To evaluate the features with high importance, we estimated the decrease in variance impurity and considered the mean decline in precision over all trees of the ensemble. An outcome class additionally separates this significance measure, and the retrieved important features from the model are plotted in a sorted manner according to their values. This significance is a proportion which highlights by what amount expelling a variable decreases impurity, and vice versa. If a variable has next to no perceptive power, rearranging may prompt a slight increment inexactness because of random noise. In turn, this can offer ascent to negative significance scores, which can be regarded as equivalent to zero importance.

We decided on the conditions based on impurity, which in the case of regression trees is variance. If we examine the decision tree depicted in Figure 3 to understand the internal mechanism of the RF model, we observe that only a few variables assist in the forecast prediction. The tree in Figure 3 shows the underlying mechanism of the random forest to reveal how each feature adds to the end value.

We learn that the value of the prediction changes along the prediction path within the decision tree, together with the information which features caused the split. The presented tree also shows that a leaf node with higher value shifts to the right branch of the decision tree. As a result, we discover that several features such as maximum temperature, the total amount of gas are useless for forecasting.

In the second layer of the hybrid model, we simulated the bidirectional LSTM model by taking only the identified essential features from the first layer as input. These five features are the minimum generation (generation end), day-peak generation (generation end), highest generation (generation end), max. demand (generation end), and evening peak generation (generation end). We predict the total_energy (Generation + India Import) variable among the fourteen properties in Table 2 for power demand forecasting. Total_energy corresponds to the inclusive power consumption on a particular day.

Before training the model, we scaled the data using the RobustScaler to make it robust to outliers using the interquartile range. Table 4 shows the parameters used to initiate the LSTM model for forecasting. We concluded the hyperparameter values based on the trial and test approach to identify the parameter values that best suit our model. During the training, the data passes through the LSTM, Dropout, and dense layers [25], and finally fits to forecast electricity load based on unseen data.

Table 4. Parameters for machine learning techniques

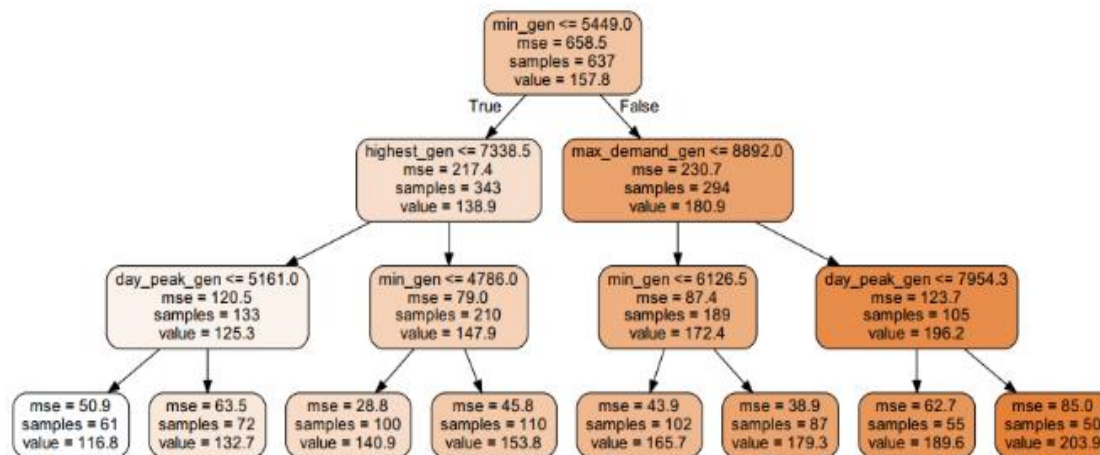| No.	Model	Parameters
1	Random Forest Regression	No of estimators = 200, min samples split = 2, max depth= None, min sample leaf = 2
2	LSTM, Bi-LSTM, RF-bi-LSTM	Activation = relu, weight optimization = adam, batch size = 512, number of epochs = 700, validation split = 0.1, learning rate = 0.01

Figure 3. A random tree upto three levels

3. RESULTS AND DISCUSSIONS

This section discusses the experimental results with a detailed performance analysis of the proposed RF-bi-LSTM hybrid model.

3.1. Performance comparison and result analysis

To observe the performance of the proposed model, we conducted a comparative analysis of our hybrid model with the standard LSTM and bi-LSTM models. We adapted cross-validation (CV) for optimum performance evaluation by splitting our data into test and train sets. In the experimental evaluation, we apply a total of 13 features to train the LSTM and the bi-LSTM models. Later we train the proposed RF-bi-LSTM hybrid model utilizing only the five most dominant features having positive values in terms of prediction strength. In our proposed model, we use the window of 10 days as an input to the bi-LSTM layer to predict the electricity consumption of the next day as a result. As for performance metrics, we utilize mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE).

Applying the performance metrics, we present the summary of the experimental results in Table 5, which concludes that the proposed RF-bi-LSTM model returns a more favorable output in terms of forecasting electrical load than the standard LSTM, and bi-LSTM models trained with all 13 features. By observing the values of the error metrics in Table 5, we can establish that the proposed method has a competitive advantage for power consumption prediction over other deep learning methods.

Table 5. Performance comparison of proposed method

Method	MSE	RMSE	MAE	MAPE
LSTM	0.4776	0.6911	0.5578	148.79
Bi-LSTM	0.2943	0.5425	0.4317	194.80
Proposed RF-bi-LSTM	**0.1673**	**0.4090**	**0.3070**	**193.49**

To illustrate the comparison better, we visualized our prediction data. In Figure 4, we discern a distinct difference between the predicted forecast results of the standard bi-LSTM and the proposed hybrid model. From Figure 4 (b), it is apparent that in maximum data points, our recommended model exhibits a lesser error. We can also conclude that we can predict the peak points in the energy consumption data more precisely using the hybrid model than the bi-LSTM model, see Figure 4.

Attributing to the values in Table 5, we can verify that our proposed hybrid model manifests better performance resulting in fewer error metric values than the LSTM and the bi-LSTM model. In the time-series decomposition graph in Figure 2, we can see that our data has features like changing trend, seasonality, and non-stationarity which make it quite challenging to forecast electricity consumption given it is a multivariate time-series data. Figure 4 (b) reflects that our proposed model has efficiently modelled these complex traits and performs superior to other deep learning models.
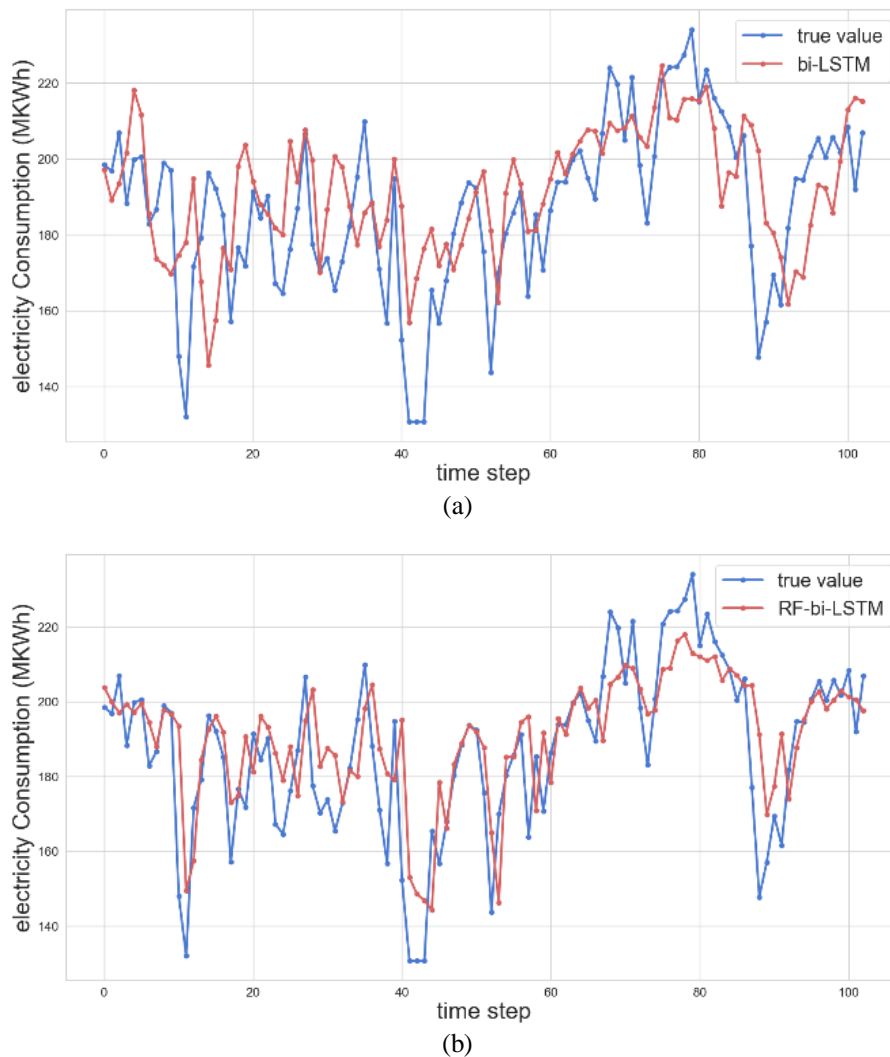

(a)

(b)

Figure 4. Graphs of predicted forecast results using a) the standard bi-LSTM model,
(b) the proposed RF-bi-LSTM mode

3.2. Learning loss analysis of proposed RF-bi-LSTM model

When we trained the LSTM and the bi-LSTM models with all features from the power consumption dataset, we detected that both of the models present an overfitting tendency as they grow more specialized with the training data. As a consequence of overfitting, the validation loss begins to rise after a specific point while the training loss sustains falling as the models gain experience. Furthermore, the models fail to accurately predict new observations that are not part of the original training data set, and generalization error rises.

On the contrary, our proposed model trained with significant features balances between an overfit and underfit, resulting in a good fit. In Figure 5, we can see the learning loss graph of our proposed RF-bi-LSTM model. From the loss graph, we can distinguish that our model delivers a good fit as the training and validation loss drop to the point of stability and has a modest generalization gap in between both curves. It exhibits stable learning characteristics and signifies our model's ability to produce a better forecast result

on unseen data. From the observation of learning loss between distinct models, it is apparent that our proposed model yields to be more efficient than the other state-of-art models.
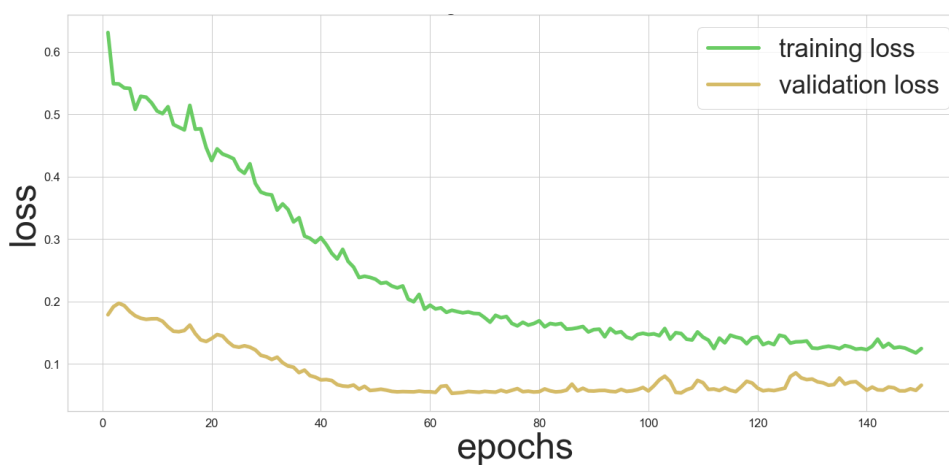

Figure 5. Loss and validation loss per epoch for proposed RF-bi-LSTM model

4. CONCLUSION

This paper presents a hybrid model for short-term load forecasting using the random forest and the bidirectional long short-term memory. To validate the model, we use a three years long electric power consumption dataset. For preparing the dataset, we collected daily PDF reports from the official website of BPDB and then converted it into the CSV format. Among the 13 features of the dataset, we have found the five most prominent features. In the experimental evaluation, using the five dominant features, the performance of the proposed model is compared with the state-of-art models by considering the different statistical parameters-MSE, RMSE, MAE, and MAPE. Experimental results demonstrate that the proposed hybrid model outperforms the compared models by exhibiting fewer error metrics.

REFERENCES

[1]   A. Ahmed and M. Khalid, "A review on the selected applications of forecasting models in renewable power systems," *Renewable and Sustainable Energy Reviews*, vol. 100, pp. 9-21, Feb. 2019.
[2]   K. Chen, K. Chen, Q. Wang, Z. He, J. Hu, and J. He, "Short-Term Load Forecasting with Deep Residual Networks," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3943-3952, Jul. 2019.
[3]   P. Linares, P. Pintos, and K. Wurzburg, "Assessing the potential and costs of reducing energy demand," *Energy Transitions*, vol. 1, no. 1, Jun. 2017.
[4]   T. GONEN, "Load Forecasting Using Box-Jenkins Methodology," *COMPEL – The international journal for computation and mathematics in electrical and electronic engineering*, vol. 3, no. 1, pp. 35-46, Jan. 1984.
[5]   J. W. Taylor and R. D. Snyder, "Forecasting intraday time series with multiple seasonal cycles using parsimonious seasonal exponential smoothing," *Omega*, vol. 40, no. 6, pp. 748-757, Dec. 2012.
[6]   D. Alberg and M. Last, "Short-term load forecasting in smart meters with sliding window-based ARIMA algorithms," *Vietnam Journal of Computer Science*, vol. 5, no. 3–4, pp. 241-249, Jun. 2018.
[7]   J. Zhang, Y.-M. Wei, D. Li, Z. Tan, and J. Zhou, "Short term electricity load forecasting using a hybrid model," *Energy*, vol. 158, pp. 774-781, Sep. 2018.
[8]   N. Huang, G. Lu, and D. Xu, "A Permutation Importance-Based Feature Selection Method for Short-Term Electricity Load Forecasting Using Random Forest," *Energies*, vol. 9, no. 10, p. 767, Sep. 2016.
[9]   Y. Chen et al., "Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings," *Applied Energy*, vol. 195, pp. 659-670, Jun. 2017.
[10] S. Ben Taieb and R. J. Hyndman, "A gradient boosting approach to the Kaggle load forecasting competition," *International Journal of Forecasting*, vol. 30, no. 2, pp. 382-394, Apr. 2014.
[11] D. K. Chaturvedi, A. P. Sinha, and O. P. Malik, "Short term load forecast using fuzzy logic and wavelet transform integrated generalized neural network," *International Journal of Electrical Power & Energy Systems,* vol. 67, pp. 230-237, May 2015.
[12] P.-H. Kuo and C.-J. Huang, "A High Precision Artificial Neural Networks Model for Short-Term Energy Load Forecasting," *Energies*, vol. 11, no. 1, p. 213, Jan. 2018.
[13] D. C. Park, M. A. El-Sharkawi, R. J. Marks, L. E. Atlas, and M. J. Damborg, "Electric load forecasting using an artificial neural network," *IEEE Transactions on Power Systems*, vol. 6, no. 2, pp. 442-449, May 1991.

[14] K. Y. Lee, Y. T. Cha, and J. H. Park, "Short-term load forecasting using an artificial neural network," *IEEE Transactions on Power Systems*, vol. 7, no. 1, pp. 124–132, 1992.

[15] J. Moon, Y. Kim, M. Son, and E. Hwang, "Hybrid Short-Term Load Forecasting Scheme Using Random Forest and Multilayer Perceptron," *Energies*, vol. 11, no. 12, p. 3283, Nov. 2018.

[16] D. L. Marino, K. Amarasinghe, and M. Manic, "Building energy load forecasting using Deep Neural Networks," *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*, 2016.

[17] S. Bouktif, A. Fiaz, A. Ouni, and M. Serhani, "Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches," *Energies*, vol. 11, no. 7, p. 1636, Jun. 2018.

[18] Z. Wu, X. Zhao, Y. Ma, and X. Zhao, "A hybrid model based on modified multi-objective cuckoo search algorithm for short-term load forecasting," *Applied Energy*, vol. 237, pp. 896-909, Mar. 2019.

[19] H. Zheng, J. Yuan, and L. Chen, "Short-Term Load Forecasting Using EMD-LSTM Neural Networks with a Xgboost Algorithm for Feature Importance Evaluation," *Energies*, vol. 10, no. 8, p. 1168, Aug. 2017.

[20] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, Nov. 1997.

[21] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A Search Space Odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222-2232, Oct. 2017.

[22] J. P. C. Chiu and E. Nichols, "Named Entity Recognition with Bidirectional LSTM-CNNs," *Transactions of the Association for Computational Linguistics*, vol. 4, pp. 357–370, Dec. 2016.

[23] M. Shahbaz, G. Salah Uddin, I. Ur Rehman, and K. Imran, "Industrialization, electricity consumption and CO2 emissions in Bangladesh," *Renewable and Sustainable Energy Reviews*, vol. 31, pp. 575–586, Mar. 2014.

[24] M. Jaworski, P. Duda, and L. Rutkowski, "New Splitting Criteria for Decision Trees in Stationary Data Streams," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 6, pp. 2516–2529, Jun. 2018.

[25] P. Hu, J. Tong, J. Wang, Y. Yang, and L. de Oliveira Turci, "A hybrid model based on CNN and Bi-LSTM for urban water demand prediction," *2019 IEEE Congress on Evolutionary Computation (CEC)*, 2019.

BIOGRAPHIES OF AUTHORS

Zannatul Ferdoush graduated in Computer Science and Engineering from Brac University in 2019. She is a Software Engineer at Bdjobs.com Limited. Her research interests include machine learning, data exploration and bioinformatics.

Booshra Nazifa Mahmud is a software developer at Monico Technologies Bangladesh (2020). She obtained Bachelor Degree in Computer Science and Engineering from Brac University (Bangladesh) in 2019. Her research interests are in fields of data science, machine learning, and web development.

Amitabha Chakrabarty received his Ph.D. degree from Faculty of Engineering, Dublin City University (DCU), Ireland, and M.Sc. degree in Telecommunication Engineering from Independent University, Bangladesh. He also received his B.Sc. degree in Computer Science and Engineering from Rajshahi University, Bangladesh. Dr. Chakrabarty is currently working as an Associate Professor of Computer Science and Engineering at Brac University. He is also the coordinator of M.Sc./M.Engg. in Computer Science and Engineering and Master of Computer Applications programs. His research areas are- IoT, Machine Learning, and Fog Computing.

Jia Uddin received a BSc degree in CCE from International Islamic University Chittagong, Bangladesh in 2005, and a MSc. degree in Telecommunications from the Blekinge Institute of Technology, Sweden, in 2010. He did Ph.D. in Computer Engineering from the University of Ulsan, Korea, in January 2015. Currently he is an Assistant Professor in Dept. of Technology Studies, Endicott College, Woosong University, Korea and an Associate Professor (On Leave), CSE at BracU, Bangladesh. His research interests include fault diagnosis, time series data analysis, and multimedia signal processing.