COST AND DEMAND PREDICTION OF ELECTRICITY USING NEURAL NETWORKS AND BIG DATA IN SMART GRIDS

V.Senthil Nayagam

Assistant Professor ,EEE Department ,Sathyabama Institute of Science and Technology,Tamil Nadu,India.

E-mail:sensribala123@gmail.com

Abstract - Smarter and sustainable electrical power systems are being developed in recent times. These tendencies have led to various beneficial effects, including active customer engagement in the electrical market. Nevertheless, the resultant flexibility in demand leads to rapid demand changes and raises the challenge of maintaining the power balance and smart grids' (SG) reliability. The pricing of electricity and forecasts of demand plays a crucial influence on SG's dependability and sustainability. Big data forecasting has become a new hot issue as huge amounts of information are created and saved in the SG context. With improved pricing information and power demand, energy users may efficiently regulate their load. In this article, an electricity price and demand prediction is employed with Big Data using multiple long, short-term memory variables (MV-LSTM). Researchers are continually engaged in proposing novel prediction models. These models have a single input variable and several other variables. Therefore, the suggested model employs several factors as input and estimates future energy consumption and cost values. The parameter setting is needed since the predictive model's performance depends on these parameter values. Inaccurate forecasts can lead to the selection of unsuitable values. Including an optimization, mechanism enhances predictability with minimal effort by the user. Data gets preprocessed and improved by the z-score procedure for efficient predictions of missing values and outliers. In addition, before prediction, the data is standardized. The accuracy of the forecast of the suggested model is measured by the root mean square error (RMSE).
Keywords: Big Data, Smart Grids, Cost and demand prediction, LSTM, RMSE.

1 INTRODUCTION

The existing electricity system solely depends on natural fuel, and energy is generated at a site far off from customers [1]. The conventional energy infrastructure has been reported to have energy losses due to the enormous distance between energy users and production. The main reason for blackouts [2] is thus the transmission and distribution lines. To address these issues, power facilities that are close to the demand zones are therefore required. This refers obviously to greener sources of energy such as microgrids [3]. The serious carbon emission problem will immediately be handled if the microgrid generates a maximal share of electricity consumption. The energy sector creates about 42% carbon footprint, and transport contributes 19% of the total greenhouse gasses worldwide [4]. This is why the conventional power source is not possible to meet the future demands of electricity users.

The existing power plants also produce large greenhouse gas emissions as they are heavily dependent on fossil fuels. With the development of MGs, the lowest use of fossil fuels and reduced carbon emissions are possible[5]. Moreover, in two methods, such as island and grid connections, the microgrid links in the smart house. The exterior grid cannot sell/buy energy together in the islanded mode[6]. However, the latter option permits the microgrid to sell and buy power from/to the exterior grid. The microgrid produces electricity from renewable sources of energy (RES), such as windmills, solar cells, tidal systems, etc., that greatly contribute to reducing carbon emissions[7]-[9].

In addition, because of their intermittent nature, RESs are unreliable sources of power generation. Energy storage systems (ESS) are therefore regarded vital for optimal use of sustainable power. Demand response (DR) plays an important role[10]. DR promotes end-users to control their power requirements or to offer reward and mitigation measures for power. By scheduling household appliances, a consumer may minimize costs and boost contentment. Residential electricity programs, DR schemes have been explored extensively[11].

In several residences with identical living patterns, the problem of DR has been examined in [12]. However, numerous households cannot possibly have the same equipment with the same operating time (LOT) and power ratings (kW). The timetable of consumption may be determined ideally by an effective scheduling technique. Therefore, the sequencing pattern shows the same characteristics as the generation of microgrid electricity[13]. Consumers can consequently conserve energy expenditures and import sufficient power from their exterior grid. In addition, accurate and steady microgrid operations may ensure efficient scheduling[14].

The remaining document is organized as follows. Section 2 explores related works on the 5G heterogeneous system. An electricity price and demand prediction are employed in section 3 with Big Data using multiple variables, long-term, short-term memory (MV-LSTM). Section 4 consists of the analysis and findings obtained from the proposed model. The conclusion and possible studies have been outlined in Section 5.

2 RELATED WORKS

Various methods and techniques have been developed for the Seasonal and Tendency breakdown using Loess Forecasting (STLF). Three primary groups called statistics, machine education, and hybrid approaches may be categorized into the STLF. The statistical techniques employ statistical time series modeling methodologies and predict the load using well-known time-series data, direction and seasonal characteristics. These quick approaches confront certain constraints with unilinear, complicated and turbulent time series. Initially, in STLF, statistical methods such as Multiple regression [15], state-space [16] and Exponential smoothing [17] have been used.

The second scenario is machine learning techniques for modeling historical data for training, which allows the system to be implemented for any uncertain future data after the framework has merged. Specific assumptions regarding time series requirements are not required for these methods[18]. They are also sensitive to uniformity or normalcy, which enables them to represent complicated time series. Artificial Neural Networks (ANN), Support Vector Machines (SVM) and Support Vector Regression (SVR) are the most recognizable approaches in this category[19].

There are other ANN-based forecasting methods in the literature that come from certain amendments to the core notion. ANN is, for instance, commonly trained by backpropagation (BP). However, this approach has several limitations, such as slow convergence and lock-in the minimal locality[20]. Stochastic weight allocation and logical connection between the hidden layer and output layer have increased the efficiency of the ANN training. This enhancement may be shown in the random weight neural network with a single hidden layer [21]. Finally, hybrid or blended approaches are included in the last category. For example, the neural network techniques have been combined with the other algorithms [22].

The research in [23] seeks to employ large data for regulating, processing and protecting functions in the electrical power system. As the volume of information is large and quickly growing, conventional database techniques cannot process this data. To manage this sort of data, special technologies are necessary. Thus, this research discusses three crucial elements of the use of large data: functional extraction, integration, and application[24]. In addition, this study describes the use of large data in asset management, defect detection, operational planning, and dispersed generation electrical systems. Information management processes have been presented, and future research guidelines for big data involvement in SG have been outlined [25].

Authors in [26] have discussed the use of IoT in SG. The incorporation of RES in the system has also been proposed. SG-enabled IoT produces enormous data every day that may be archived and used in the long term. As these figures are gigantic in amount, obtained from disparate sources and increased with time, they are called big data. This large amount of data is important for the operation, planning and efficiency of future SGs. This study also discusses the technologies utilized to store and analyze massive data. In addition, a multiple-level IoT framework has been presented for SG.

In [27], a framework has been suggested for anomaly detection in real-time. With the use of intelligent meter collected data, this framework can identify problems easily. The mistake count has been measured on each consumer's side and supplied to the system that connects the customer. The authors noted in [28] that large data had been generated by the SG's deployment of monitoring equipment. To maximize the effectiveness of the SG system, the usage of this large data is a crucial part.

The researchers focus on the questions of the data gathering in this study. In [29], a predictive model has been proposed for the electricity cost and load. The model employs an iterative neural network to predict future power charges and pricing. This is the hybrid prediction model. The results showed that the error predicted between projected model values and total power load and pricing values is minimal. In [30], a load and pricing prediction mechanism have been proposed. They asserted that the electricity cost fluctuates automatically, depending on the mismatch between the current and the demand for load at a specific period. The prediction of these variables is of major relevance for managing electricity consumption following the available power.

3 ELECTRICITY PRICE AND DEMAND PREDICTION IS EMPLOYED WITH BIG DATA USING MV-LSTM METHOD

The neural networks function well with big and complicated data sources. Functional engineering does not necessarily make them versatile. Furthermore, MV optimization algorithms have been used in hyperparameters, epoch, batch size, and window size, and this prediction model is called MV-LSTM. The use of MV-LSTM rather than LSTM allows the system to accommodate several input characteristics and modified hyperparameter values according to the set of data, which enhances prediction efficiency.

The Big Data Framework has been presented in Figure 1. The creation of data is the initial stage. It collects information collected from many resources, for example, microgrids, intelligent buildings, smart meters, electric cars, industries etc. These data are unprocessed and produced in huge quantities. The second massive step is the future storage of this data. These data might be kept in storage warehouses or cloud computing. Data preprocessing is an essential step when data are obtained for usage. This process will clean up enormous data, identify significant characteristics, filter out extraneous and unclear data, and transform the data to the database to use different analysis techniques. The next stage is to obtain the relevant information through detection and data mining techniques.

It is utilized to make significant judgments in several sectors following information extraction. In the procedure outlined, MV-LSTM has been used to analyze the data, predict power demand and cost. On the consumer end and the utility side, the knowledge of the predicted parameters might be utilized. Users may use this data in power control units to regulate their power effectively, cut their power costs, boost convenience and lower the PAR. In contrast, the company may utilize this information to regulate availability and price and execute other key energy production choices.

3.1 Proposed MV-LSTM model

Figure 2 shows the electricity price and demand prediction model employed with Big Data using the MV-LSTM framework. The phases of this model are discussed in detail.

3.1.1 Input Electricity Cost And Demand Data

The data needed for the prediction are collected at this phase. Two sets of data, i.e., power demand and cost information, are employed in this model[30]. The data set includes information gathered per hour. To reproduce the findings, publicly available information has been used. A range of factors for both cost and demand predictions are utilized in the study as input variables. The increasing input parameters make the model complicated and need more time and technology to predict effectively. Studies are also available with only one predicting input variable. These models take past values of input parameters and forecast their possible trends. Thus, the most dependent parameters that meet the basic requirements of the strategy have been employed in the suggested model. The next stage is data preprocessing, following the data selection process. The cost inputs that are supplied as an input to the prediction model are as follows:

$$C = \begin{bmatrix} c(J_1E_1) & c(J_2E_1) & \cdots & c(J_pE_1) \\ c(J_1E_2) & c(J_2E_2) & \cdots & c(J_pE_2) \\ \vdots & \vdots & \vdots & \vdots \\ c(J_1E_q) & c(J_2E_q) & \cdots & c(J_pE_q) \end{bmatrix} \tag{1}$$

where J denotes the prediction period, and E represents the number of days considered for prediction. The dimension of the matrix is $p \times q$, where J_p is the pth day of prediction and E_q is the qth period in a day. The demand input data is the same as the pricing information and shows the data in terms of the number of days. The demand input dataset may be shown as follows:

$$e = \begin{bmatrix} e_m(J_1E_1) & e_m(J_2E_1) & \cdots & e_m(J_pE_1) \\ e_m(J_1E_2) & e_m(J_2E_2) & \cdots & e_m(J_pE_2) \\ \vdots & \vdots & \vdots & \vdots \\ e_m(J_1E_q) & e_m(J_2E_q) & \cdots & e_m(J_pE_q) \end{bmatrix} \tag{2}$$

3.2 Pre-processing

Data preprocessing is a technique of data mining used to turn unprocessed data into an appropriate and understandable format. There is a lack of cohesion, error, incompleteness and no clear behavior in the data obtained from the actual world. We thus need to change this data into an understandable format before using it. Data preparation techniques on raw data must be employed to achieve the objective. Data undergo a range of data purification and optimization techniques during the preprocessing phase:

• The first step is to clean up data. The missing values, smoothing noise and data incoherence have been filled or deleted from the set.
• Conflicts between data are addressed while data of various formats are integrated, and outliers have been eliminated.
• The transformation of data is also a key step in normalizing the values of the variables between a similar interval to make them comparable.
• Data are reduced by removing irrelevant or duplicate information in the data reduction stage.

3.2.1 Eliminating the Missing Data

Eliminating missing values is a key task in data preparation. The missing information might diminish predicting accuracy and lead to a false result. Therefore, before expecting, it is necessary to address missing data values. For this objective, many approaches are available. One of them is to remove data rows. As the name indicates, data sets omit the rows with missing values. Predicting algorithms have been used to estimate missing data values using the current value of information. Researchers progress from fundamental to complicated ways to pick the optimum approach. The first strategy has been employed in this paper, as the data set has extremely low numbers of missing values. So, the rows with missing data have been eliminated using Python's Pandas library function instead of increasing the computing cost.

3.2.2 Eliminating Outliers

The outliers are the values in a given dataset that differ considerably from other data sets. These outsourcing systems might come from public vacation days like Christmas or exceptional events in which the energy demand is higher than ordinary days in SG's power demand statistics. Such data sets should be excluded during the prediction since the typical day's electricity use has to be predicted. Data observations these days might lead to a badly trained model for the prediction method. Z-score is the methodology often employed by researchers[30] to eliminate such values.

$$z = (y - n)/t \quad (3)$$

Z-score has been computed from equation (3). $n \, and \, t$, denotes the mean and standard deviation, respectively. For each sample in the database, this score is obtained. A limit is established to eliminate the outliers. There is no complex rule to choose the limit, as the removal of outliers has been balanced. As information from the data set causes information loss and the outliers are prone to mislead the predicted values. Equilibrium has to be established. Based on the context and the available dataset, the value of the chosen outlier will change.

3.2.3 Normalization

In ANN, data has to be standardized before transferring it to the prediction model. The values of many variables are not within the same range in data collection. Their values have to be normalized to make these values comparable. Uncommon input may lead to a network that is not conditioned. Data normalization also plays a key role in steady weight and other parameter integration. A min-max scaler has been utilized to standardize the data. The data range varies from zero to one.

$$y_N = (y - y_{least})/(y_{maximum} - y_{least}) \quad (4)$$

y_N denotes the normalized data. y is the input value from the data set. $y_{maximum}$ and y_{least} are the highest and lowest values of the data set, respectively.

3.3 Proposed MV-LSTM Algorithm

Once the data preprocessing step has been completed, the data will now be sent to the prediction step. MV-LSTM has been proposed in this paper. MV-LSTM has the same functioning and fundamental architecture as LSTM; the only variation is that numerous variables are employed in MV-LSTM. The MV optimization algorithm has been utilized to set hyperparameters. An RNN variant is an LSTM model. This is a prominent predictive time series framework that has been employed in several sectors and uses data with long-term prediction effectively.

As the LSTM is considered a variation of RNN, its internal design resembles that of the RNN. The inner layer connectivity allows the information to be transmitted back and forth, making it a great alternative for time series prediction. The prediction model based on RNN sets rules apart and predicts future data points. The backpropagation function makes it conceivable that it also plays a vital role in weight updating. Existing RNN confronts the constraint of gradient disappearance; LSTM overcomes that. Information with lengthy dependencies has been handled effectively. There are generally three gates, i.e., input gate, output gates and forgotten gate. The inclusion of the forgetting gateway is the reason for the gradient issue to be mitigated in LSTM [14]. The memory cell is another remarkable contribution to the LSTM. In all three gates, LSTM uses the following equations as an activation function:

$$g_u = \beta[X_g \times (H_{u-1} + y_u) + c_g] \quad (5)$$

$$j_u = \beta[X_j \times (H_{u-1} + y_u) + c_j] \quad (6)$$

$$Ot_u = \beta[X_{ot} \times (H_{u-1} + y_u) + c_{ot}] \quad (7)$$

Equations (5)-(7) denote the forgetting, input, and output gate functions. β is the sigmoid function, X is the weight function, y represents the input value, H is the secret state, and c is the partial vector. The size of the batch is an important RNN training factor. It shows the number of samples utilized during the weighting update of the system. To build the group, the following equations are used:

$$y\_Ba(j) = y\_tr(s), y\_tr(s + sl - p), \ldots, y\_tr(s + sl - p) \quad (8)$$

Where $s = random(b, c)$- random number between $b \, and \, c$ used as a reference value to select the sample available in the dataset. j is the sample number of the group.

In predicting algorithms, their hyperparameters considerably check the reliability. The model parameters should thus be specially designed. The values of the batch size, duration and window size settings have been optimized in this article. Scholars prefer algorithms motivated by nature because of their high efficiency. Nevertheless, these techniques are stochastic and require a careful choice of the input variables, for example, population density, production, and elevation. If the individual variables are not perfectly calibrated, they fall into local optimum or request additional computing time. The MV optimization approach has been employed for LSTM optimization in this study due to these difficulties. This algorithm has no tuning variables, and it is simple to execute.

In this study, three design parameters must be set, i.e., batch size, periods and window size. The population size is equivalent to 50 (the algorithm is run several times to get the optimal population size). The key step is determining the objective function because the solutions have to be assessed on this basis. The aim here is to reduce the model's prediction error. Two alternatives are chosen after population evaluation: best and worst. These solutions are then used in their values to update the present population and establish the next generation. These parameters are sent to the prediction model when the optimum values have been obtained.

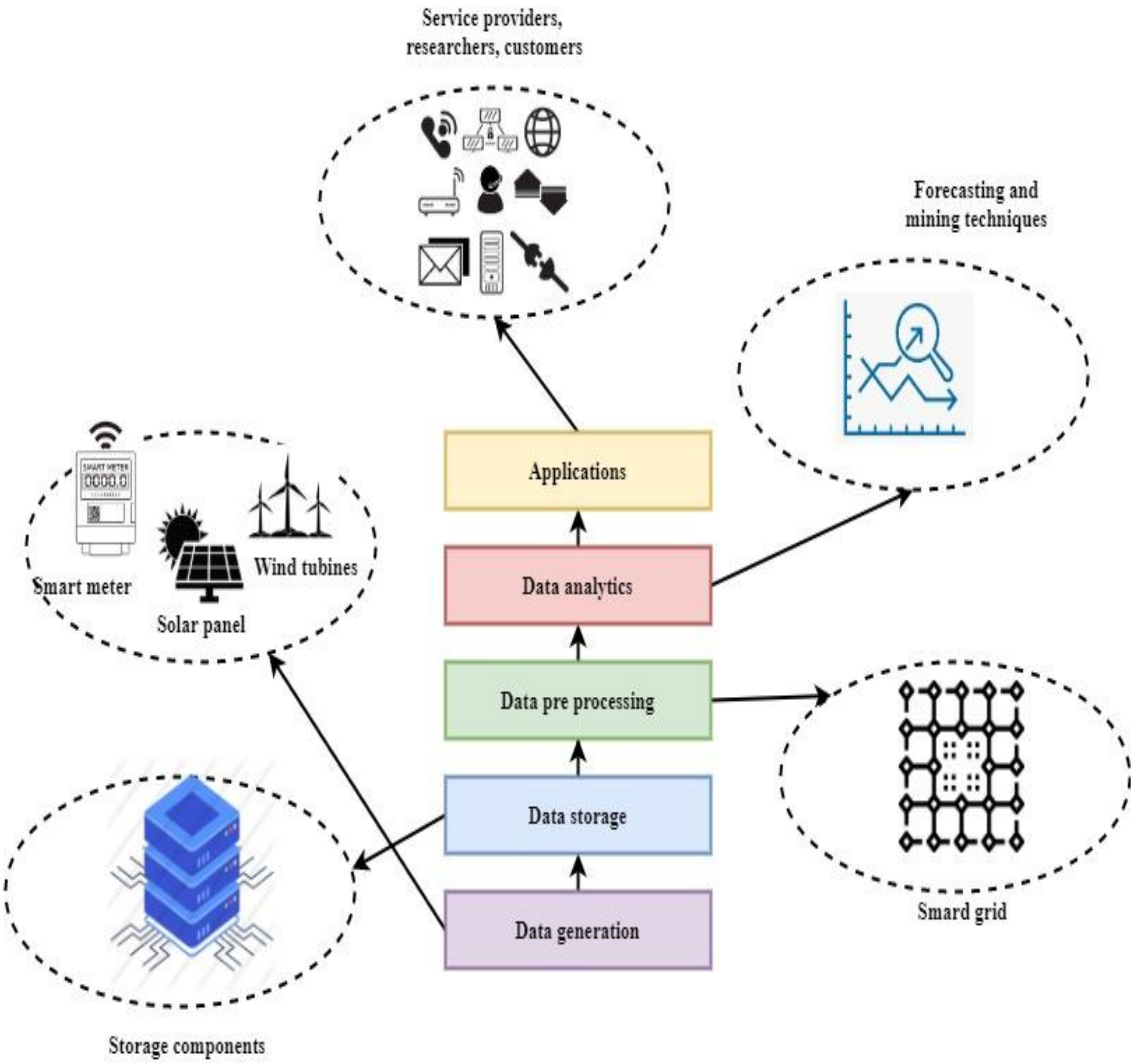


Figure 1 Context of Big Data

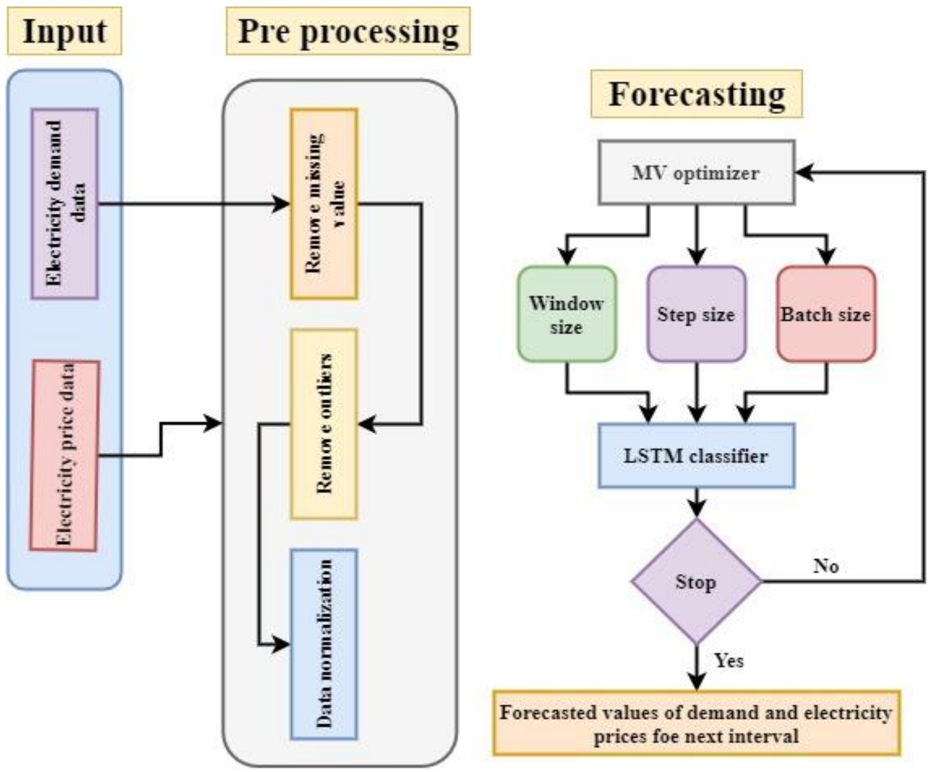


Figure 2 Electricity price and demand prediction model employed with Big Data using MV-LSTM framework

4 RESULTS AND DISCUSSION

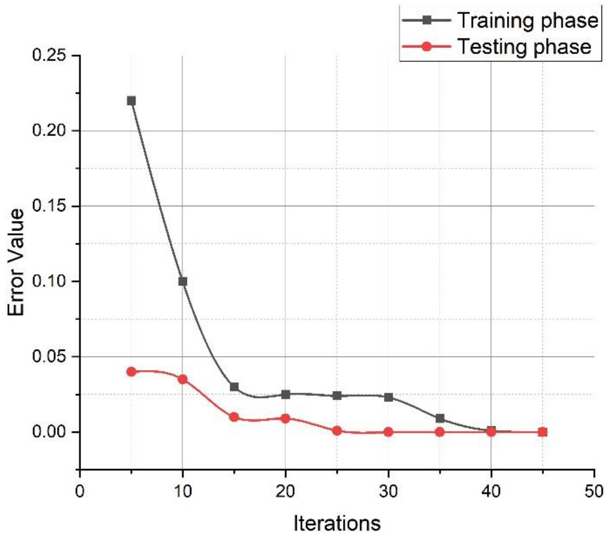
Two demand and price data sets have been applied to the proposed prediction models. The simulations have been performed using an Intel Core i5 processor with 8 GB of RAM and a 64-bit Python 3.6 operating system. The simulation tool used is Matlab. This section discusses the simulation findings of the projected levels of demand and the price of electricity.

The MV-LSTM is used for prediction purposes. The MV-LSTM parameters have been tuned using the suggested optimization methodology, and the Adom Optimizer has been utilized inside the template as a local optimizer. The maximum value of iterations is set at 45. The optimizer modifies the parameters and improves the model's predictability.

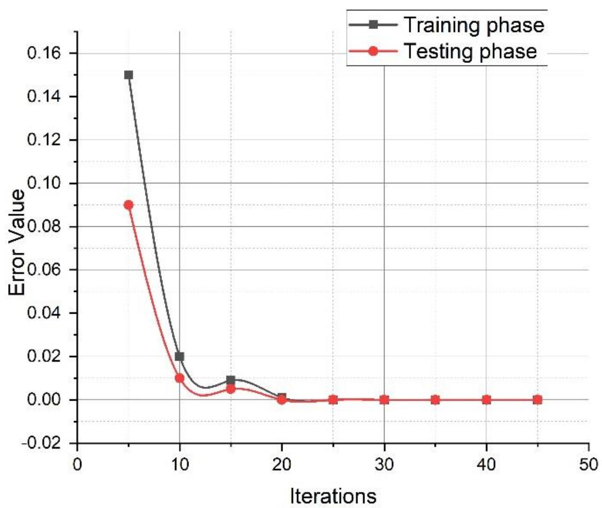
Figure 3 depicts the error value for demand and cost prediction during the training and testing phase in the proposed framework. The error value has been assessed for an increasing number of iterations. The model is evaluated with RMSE for its accuracy. In addition, during the initial iteration, the accuracy of the test and train reduces, with high error values. But, as the number of iterations increases, the error value decreases and the accuracy is improved significantly. The test and train errors in electricity demand and pricing are constant as the number of iterations reaches 50. Thus, it has been concluded that the proposed predictive algorithm (in the training phase) and the model are trained properly to estimate the actual data. Table 1 provides simulation parameters to analyze the performance of the proposed MV-LSTM.

Table 1 Simulation parameters to analyze the performance of the proposed MV-LSTM

Parameters	Values
Number of datasets taken	2
Processor	Intel core i5
RAM size	8GB
Operating system	64-bit Python version 3.6
Maximum number of iterations	45
Gate activation function	Sigmoid
Initial learning rate	0.009
Gradient limit	1
Number of layers	3
Maximum period	750
Least batch size	50



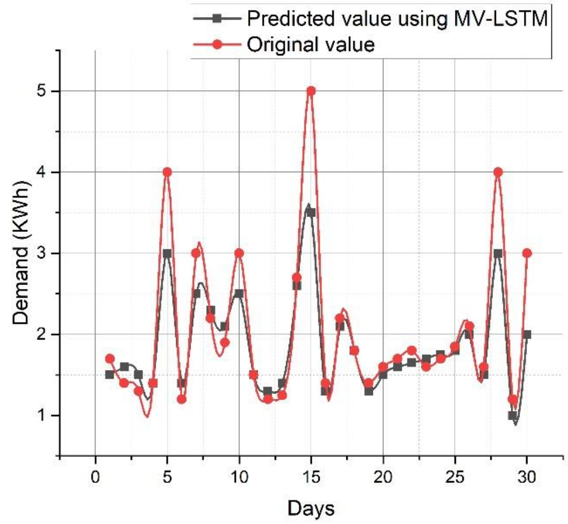
(a) Demand prediction



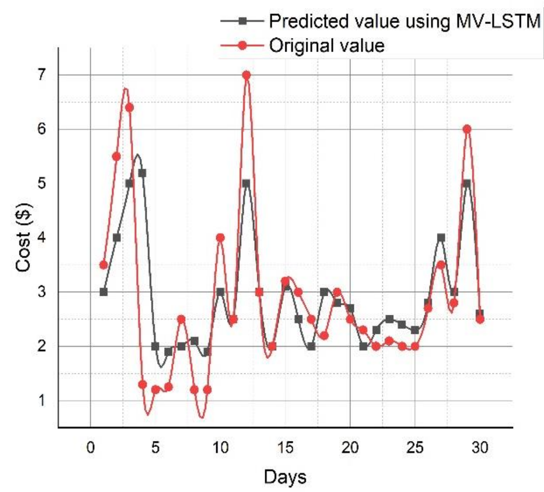
(b) Cost prediction

Figure 3 Error value for demand and cost prediction during the training and testing phase in the proposed framework

Figure 4 compares the predicted value using the proposed MV-LSTM and the original value for demand and cost in 30 days. There is a fluctuation in the actual demand and cost values concerning the number of days in a month. The real electricity demand has got the least value of 1KWh on the 29th day. The predicted value of demand on the same day has reached a value of 1.1KWh, closer to the original value. Except for the days in which the electricity demand is huge, the prediction of the proposed MV-LSTM has achieved performance closer to the actual value. Similarly, the cost prediction shows a higher deviation only when the actual cost values are higher.



(a) Demand prediction



(b) Cost prediction

Figure 4 Comparison of predicted value using the proposed MV-LSTM and original value for demand and cost in 30 days

A comparison of prediction types in terms of error value for demand and cost prediction has been shown in Figure 5. SVM, LSTM and the proposed MV-LSTM have been considered for comparison. It has been observed that the error value is maximum for the SVM prediction. LSTM has got moderate performance with a low error rate than SVM but a higher error value than the proposed method. The proposed MV-LSTM has the best prediction accuracy, achieving values closer to the actual value for both demand and cost prediction. Also, the

RMSE value is the least for the proposed MV-LSTM prediction scheme.

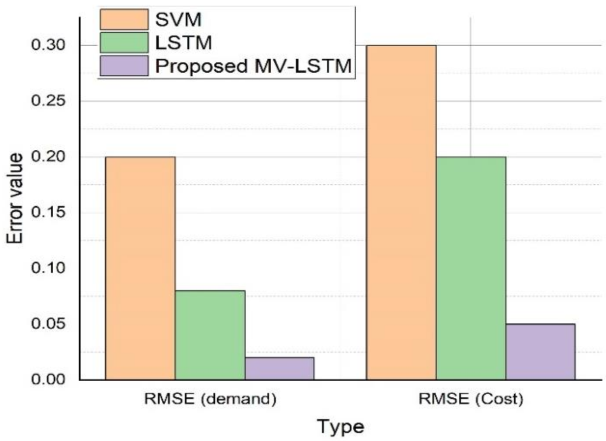


Figure 5 Comparison of prediction types in terms of error value for demand and cost prediction

5 CONCLUSION

An electricity price and demand prediction have been employed with Big Data using multiple variables, long short-term memory (MV-LSTM). Therefore, the suggested model employs several factors as input and estimates future energy consumption and cost values. The parameter setting is needed since the predictive model's performance depends on these parameter values. Inaccurate forecasts can lead to the selection of unsuitable values. Including an optimization, mechanism enhances predictability with minimal effort by the user. Data gets preprocessed and improved by the z-score procedure for efficient predictions of missing values and outliers. In addition, before prediction, the data is standardized. The key step is determining the objective function because the solutions have to be assessed on this basis. SVM, LSTM and the proposed MV-LSTM have been considered for error rate comparison. It has been observed that the error value is maximum for the SVM prediction. LSTM has got moderate performance with a low error rate than SVM but a higher error value than the proposed method. The proposed MV-LSTM has the best prediction accuracy, achieving values closer to the actual value for both demand and cost prediction. Also, the RMSE value is the least for the proposed MV-LSTM prediction scheme.

References

[1] Gungor, V. C., Sahin, D., Kocak, T., Ergut, S., Buccella, C., Cecati, C., & Hancke, G. P., “Smart grid and smart homes: Key players and pilot projects”, IEEE Industrial Electronics Magazine,vol. 6, no.4, pp.18-34, 2012.

[2] Monteiro, V., Pinto, J. G., & Afonso, J. L., “Operation modes for the electric vehicle in smart grids and smart homes: Present and proposed modes”, IEEE Transactions on Vehicular Technology, vol.65, no.3, pp.1007-1020, 2015.

[3] Monteiro, V., Ferreira, J. C., Melendez, A. A. N., Couto, C., & Afonso, J. L. ,“Experimental validation of a novel architecture based on a dual-stage converter for off-board fast battery chargers of electric vehicles”, IEEE Transactions on Vehicular Technology, vol.67, no.2, pp.1000-1011, 2017.

[4] Monteiro, V., Pinto, J. G., & Afonso, J. L., “Experimental validation of a three-port integrated topology to interface electric vehicles and renewables with the electrical grid”, IEEE Transactions on Industrial Informatics, vol.14, no.6, pp.2364-2374, 2018.

[5] Gungor, V. C., Sahin, D., Kocak, T., Ergut, S., Buccella, C., Cecati, C., & Hancke, G. P., “Smart grid and smart homes: Key players and pilot projects”, IEEE Industrial Electronics Magazine, vol.6, no.4, pp.18-34, 2012.

[6] Zaballos, A., Vallejo, A., & Selga, J. M., “Heterogeneous communication architecture for the smart grid”, IEEE network, vol.25, no.5, pp.30-37, 2011.

[7] Yu, X., Cecati, C., Dillon, T., & Simoes, M. G., “The new frontier of smart grids”, IEEE Industrial Electronics Magazine,vol. 5, no.3, pp.49-63, 2011.

[8] Siano, P., “Demand response and smart grids—A survey”, Renewable and sustainable energy reviews, vol.30, pp.461-478, 2014.

[9] Gelazanskas, L., & Gamage, K. A.,“Demand side management in smart grid: A review and proposals for future direction”, Sustainable Cities and Society, vol.11, pp.22-30, 2014.

[10] Aslam, S., Iqbal, Z., Javaid, N., Khan, Z. A., Aurangzeb, K., & Haider, S. I., “Towards efficient energy management of smart buildings exploiting heuristic optimization with real time and critical peak pricing schemes”, Energies, vol.10, no.12, 2017.

[11] Khan, A., Javaid, N., & Khan, M. I., “Time and device based priority induced comfort management in smart home within the consumer budget limitation”, Sustainable cities and society, vol. 41, pp.538-555, 2018.

[12] Kim, B. G., Zhang, Y., Van Der Schaar, M., & Lee, J. W.,“Dynamic pricing and energy consumption scheduling with reinforcement learning”, IEEE Transactions on Smart Grid, vol.7, no.5, pp.2187-2198, 2015.

[13] Lu, R., Hong, S. H., & Zhang, X., “A dynamic pricing demand response algorithm for smart grid:

reinforcement learning approach", Applied Energy, vol.220, pp.220-230, 2018.

[14] Wen, Z., O'Neill, D., & Maei, H., "Optimal demand response using device-based reinforcement learning", IEEE Transactions on Smart Grid, vol.6, no.5, pp.2312-2324, 2015.

[15] Cai, H., Shen, S., Lin, Q., Li, X., & Xiao, H., "Predicting the energy consumption of residential buildings for regional electricity supply-side and demand-side management", IEEE Access, vol.7, pp.30386-30397, 2019.

[16] Mosbah, H., & El-Hawary, M. E., "Optimization of neural network parameters by Stochastic Fractal Search for dynamic state estimation under communication failure", Electric Power Systems Research,vol. 147, pp.288-301, 2017.

[17] Rendon-Sanchez, J. F., & de Menezes, L. M., "Structural combination of seasonal exponential smoothing forecasts applied to load forecasting", European Journal of Operational Research, vol.275, no.3, pp.916-924, 2019.

[18] Xing, B., & Gao, W. J., " Overview of computational intelligence", Artificial Intelligence: Concepts, Methodologies, Tools, and Applications, IGI Global, pp. 12-31, 2017.

[19] Ren, Y., Suganthan, P. N., Srikanth, N., & Amaratunga, G, "Random vector functional link network for short-term electricity load demand forecasting", Information Sciences, vol.367-368, pp.1078-1093, 2016.

[20] Kezunovic, M., Xie, L., & Grijalva, S. ,"The role of big data in improving power system operation and protection",2013 IREP Symposium Bulk Power System Dynamics and Control-IX Optimization, Security and Control of the Emerging Power Grid, 2013.

[21] Jaradat, M., Jarrah, M., Bousselham, A., Jararweh, Y., & Al-Ayyoub, M. ,"The internet of energy: smart sensor networks and big data management for smart grid", Procedia Computer Science, vol.56, pp.592-597, 2015.

[22] Mi, J., Wang, K., Liu, B., Ding, F., Sun, Y., & Huang, H., "A multiobjective evolution algorithm based rule certainty updating strategy in big data environment", GLOBECOM 2017-2017 IEEE Global Communications Conference,2017.

[23] Xu, C., Wang, K., Li, P., Xia, R., Guo, S., & Guo, M., "Renewable energy-aware big data analytics in geo-distributed data centers with reinforcement learning", IEEE Transactions on Network Science and Engineering, vol.7, no.1, pp.205-215, 2018.

[24] Moghaddass, R., & Wang, J., "A hierarchical framework for smart grid anomaly detection using large-scale smart meter data", IEEE Transactions on Smart Grid, vol.9,no.6, pp.5820-5830, 2017.

[25] Munshi, A. A., & Yasser, A. R. M. "Big data framework for analytics in smart grids", Electric Power Systems Research, vol.151, pp.369-380, 2017.

[26] Munshi, A. A., & Mohamed, Y. A. ,"Cloud-based visual analytics for smart grids big data", 2016 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), 2016.

[27] Guo, H., Liu, J., & Zhao, L., "Big data acquisition under failures in FiWi enhanced smart grid", IEEE Transactions on Emerging Topics in Computing, vol.7, no.3, pp.420-432, 2017.

[28] Hou, W., Ning, Z., Guo, L., & Zhang, X., "Temporal, functional and spatial big data computing framework for large-scale smart grid", IEEE Transactions on Emerging Topics in Computing, vol.7, no.3, pp.369-379, 2017.

[29] Alamaniotis, M., Gatsis, N., & Tsoukalas, L. H., "Virtual Budget: Integration of electricity load and price anticipation for load morphing in price-directed energy utilization", Electric Power Systems Research, vol.158, pp.284-296, 2018.

[30] Amjady, N., & Daraeepour, A., " Mixed price and load forecasting of electricity markets by a new iterative prediction method", Electric power systems research, vol.79, no.9, pp.1329-1336, 2009.