Optimizing Parameters of Artificial Intelligence Deep Convolutional Neural Networks (CNN) to improve Prediction Performance of Load Forecasting System

F M Butt¹,², L Hussain³,⁴, S H M Jafri¹, K J Lone³, M Alajmi⁵, I Abunadi⁶, F N Al-Wesabi⁷ and M A Hamza⁸

¹ Department of Electrical Engineering, Mirpur University of Science & Technology, Mirpur 10250, Azad Kashmir, Pakistan
² Department of Electrical Engineering, University of Azad Jammu and Kashmir, Chehla Campus, Muzaffarabad 13100, Azad Kashmir, Pakistan
³ Department of Computer Science & IT, University of Azad Jammu and Kashmir, King Abdullah Campus, 13100 Muzaffarabad, Azad Kashmir, Pakistan
⁴ Department of Computer Science & IT, University of Azad Jammu and Kashmir, Neelum Campus, 13230 Athmuqam, Azad Kashmir, Pakistan
⁵ Department of Computer Engineering, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia
⁶ Department of Information Systems, Prince Sultan University, P.O.Box No. 66833 Rafha Street, Riyadh 11586
⁷ Department of Computer Science, College of Science & Art at Mahayil, King Khalid University, Saudi Arabia.
⁸ Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, AlKharj, Saudi Arabia

corresponding: L Hussain (lall_hussain2008@live.com), F N Al-Wesabi (falwesabi@kuu.edu.sa)

Abstract. Load Forecasting is an approach that is implemented to foresee the future load demand projected on some physical parameters such as loading on lines, temperature, losses, pressure, and weather conditions etc. This study is specifically aimed to optimize the parameters of deep convolutional neural networks (CNN) to improve the short-term load forecasting (STLF) and Medium-term load forecasting (MTLF) i.e. one day, one week, one month and three months. The models were tested based on the real-world case by conducting detailed experiments to validate their stability and practicality. The performance was measured in terms of squared error, Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Mean Absolute Error (MAE). We optimized the parameters using three different cases. In first case, we used single layer with Rectified Linear Unit (ReLU) activation function. In the second case, we used double layer with ReLU – ReLU activation function. In the third case, we used double layer with ReLU – Sigmoid activation function. The number of neurons in each case were 2, 4, 6, 8, 10 and 12. To predict the one day ahead load forecasting, the lowest prediction error was yielded using double layer with ReLU – Sigmoid activation function. The number of neurons in each case were 2, 4, 6, 8, 10 and 12. To predict the one week ahead load forecasting demands, the lowest error was obtained using single layer ReLU activation function. Likewise, to predict the one month ahead forecasting using double layer with ReLU – Sigmoid activation function. To predict ahead one-week load forecasting demands, the lowest error was obtained using single layer ReLU activation function. Moreover, to predict ahead three months forecasting using double layer ReLU – Sigmoid activation function produced lowest prediction error. The results reveal that by optimizing the parameters further improved the ahead prediction performance. The results also show that predicting nonstationary and nonlinear dynamics of ahead forecasting require more
complex activation function and number of neurons. The results can be very useful in real-time implementation of this model to meet load demands and for further planning.

1. Introduction

The phenomenon of load forecasting in the last few years has captivated growing attention out of academics, scholars and practitioners, being a very significant fragment in management modernization across the electric power networks and systems. Load forecasting of power holding high accuracy has an ability to relieve the conflict between supply and demand of power along with providing a firm foundation to sustain then stability and reliability in power grid. Having said that, electric load happens to be a haphazard, non-stationary sequence, that is regulated by multiple factors that include the factors related to weather, time, season in addition to day and random out-turns, that escort to the way where load forecasting becomes a challenging matter of investigation [1].

Currently, the methodologies available for the phenomenon of the load forecasting can be classified into dual parts including classical mathematical and statistical technique and perspectives (approaches) which find their roots to artificial intelligence. Many of the theories related to load forecasting are observed having the base on time series examination along with auto regression frameworks, which includes the vector auto regression model (VAR) model [2,3] auto regressive moving average (ARMA) model [4–6], and so forth. The prediction methodologies related to the smoothing of time series have been criticize and condemn by the researchers considering the weakness inside that comes out of the nonlinear fixing capability. With the expansion of electricity merchandise, the demand of load forecasting with high accuracy is becoming more severe and efficient. Consequently, artificial intelligence, that includes support vector architecture along with neural network, obtains escalating recognition by the scholars [7], implemented STLF by adopting Artificial Neural Network. The instances holding real data depicted effectiveness observed in the suggested technique with the demonstration of making the use of artificial neural network that can decrease and lesson the errors in load forecasting, in comparison to variety of other existing techniques. Yu and Xu [8], Suggested a suitable blended approach that was constructed on an up graded back propагative neural network meant for ST load forecasting of gas.

In the past, researchers developed various traditional methods to predict the electric load forecasting. The traditional methods are established includes counting stochastic time series [9], knowledge based methods [10], multiple linear regression [11], and exponential smoothing [12]. Traditional methods are seen often functioning imperfectly along nonlinear forecasting in addition to having STLF, MTLF and LTLF as a nonlinear complication.

The artificial intelligence (AI) based methodologies, as an illustration the artificially functioning neural networks [13–16], fuzzy logistic techniques [17], Bayesian neural system [18], expert systems structures [19], and support vector apparatus [20–23], are used in handling wide number of forecasting issues. Researchers are developing various tools based on AI methods for accurate load forecasting due to nonlinear and nonstationary behaviour of these signals. These methods require more powerful optimization strategies for improving the prediction capabilities. Moreover, deep learning networks can be further optimized by using different set of optimization functions and number of neurons. The combination of optimization functions and neurons can play a vital role in recognizing the complex behavior of forecasting time series and can predict ahead forecasting demands.

The prediction performance of the load forecasting systems is still a challenging task. In the previous studies, the researchers utilized traditional methods which require still optimization to improve the prediction performance. This study is specifically aimed to optimize the parameters of machine and deep learning algorithms, which could efficiently improve the prediction performance.

The aim of this study is to improve the ahead prediction of 24 hours, one week, one month and three months load forecasting by optimizing the CNN model using single layer with ReLU activation function and double layers with ReLU–Sigmoid and ReLU–ReLU combinations of activation functions. In each of these conditions, we also changed the number of neurons changing from 2, 4, 6, 8, 10 and 12 neurons. We aimed to investigate that whether which layer and sigmoid function combination is
appropriate to accurately predict the ahead STLF and MTLF. As the longer forecasting load demands may require a more complex and combination of activation functions and more neurons than a simplest activation function, layer, and number of neurons. We also aimed that how the number of neurons impact the load forecasting prediction. So, this study aimed to optimize the CNN model for improving the prediction performance than merely using th CNN with default parameters.

2. Methods Section

2.1. Description of Materials

In this analysis, we used dataset of electric power consumption collected on hourly basis between 1st January 2017 to 31, May 2020 (41 months) from Islamabad Power station. We split the dataset into two sets: train and test with respect to testing period i.e. for model 1, we used the data of 24 hours as a test data and previous data was considered as training data. Similarly, for model 2, we used last one week as test set and for model 3 and 4 last one- and three-months data use as test set and previous data was considered as training data. Then we trained the DNN with respect to three cases as follow.

Case 1. Single hidden layer with ReLU Activation Function. For five input variables, the ReLU is utilized to do the forecast. The neurons are selected randomly until the varied MSE is obtained.

Case 2. For two hidden layers of ReLU Activation Function

In this case, again the neuron number varied until a small MSE is achieved.

Case 3. ReLU-Sigmoid Combination

For Sigmoid-ReLU combination, and two hidden layers the neurons are selected randomly until a better MSE forecast is obtained. In layer 1, ReLU is utilized, whereas in layer 2, sigmoid is used.

until a better MSE result are obtained. ReLU is used in layer 1 and sigmoid is used in layer 2.

After that, we evaluated and compared the performance of each model with different error metrics on all test sets.
Figure 1. Schematic diagram regarding the flow of study

The Figure 1 reflect the schematic diagram of our proposed study. After ready the load forecasting time series data, it was pre-processed and divided into training and testing data. We applied three models for one day, one week, one month and three months and activation functions for each model are selected according to the case 1, 2 and 3. The evaluation of each model was tested based on different error methods.

2.2. Prediction Methods with Parametric Optimization

2.2.1. Deep Learning

Deep learning (DL) known as structural learning is a part of family of machine learning methods which is based on artificial neural networks (ANNs) with representation learning. ANN works like a human brain do. In ANN, the neuron plays a vital role in performing any sort of activity and forward
data to different neurons. In the DL, the adjective “deep” refer to use multiple layers in the network. ANNs are inspired to process information and communicate information via distributed nodes in biological systems.

2.2.2. Convolutional Neural Network (CNN)

There are many applications of convolution neural network, for example image fragmentation, object location, face recognition, object detection, video grouping, image inscribing and depth assessment [24–30]. Neural system showed signs of improvement results by utilizing convolutional neural system in machine vision. At that point specialists acquired this thought in modern zones that this innovation gets well known. In 1998 [31] proposed plan that present Convolution Neural Network resembles LeCun plan which are improved structure than [32] engineering of the CNN, CNN gets well known by winning the ILSVRC2012 rivalry [24].

![CNN Architecture for load forecasting](image)

**Figure 2.** CNN Architecture for load forecasting

The Figure 2 reflects the schematic flow of our CNN algorithm. The load forecasting time series as 1D is utilized as input to our CNN model, further processed in different layers of CNN model. The output is generated in term of different error metrics.

2.3. Activation Functions

Activation function is used as a decision-making parameter to calculate and obtain different most relevant features from the given data. If the obtained value of output from activation function is zero, it represents the absence of feature and if obtained value is one then it shows presence of that feature in the data. While working with computational networks, the output can be defined by the activation function of a node which shows the output from that specific node is single input or a set of inputs. In digital computing, activation function depends upon the received input and it is represented in the form of either “ON” (1) or “OFF” (0), like the behaviour of linear perceptron as used in neural networks. However, in case of non-linear activation function, various nontrivial problems can be solved by using small number of nodes. In ANN earlier mentioned function is also termed as transfer function. Activation function plays vital role in the training of multi-layer neural network model and adjusting the weights. In this work, we have optimized the following activation functions in our CNN model by using a rectified linear unit and non-linear sigmoid function for hidden layer in the model:

2.3.1. Sigmoid Function

The sigmoid function can be defined as

\[
f(x) = \frac{1}{1 + e^{-x+\alpha}}
\]
Here x is used as input to the neuron, α denote the offset parameter. The sigmoid function computes the value as zero. The functionality of the sigmoid function is same is step function. For larger x, the sigmoid function output is near to 1.

2.3.2. Rectified Linear Unit (ReLU)
Rectified linear unit mathematically can be defined as
\[ R(x) = \max(0, x) \]
Where x represents the input to the neuron.
ReLU and linear functions are almost similar to one another. Moreover, ReLU has following two distinct advantages:

i) As compared to other activation functions, ReLU is more efficient for training purpose.

ii) It overcomes the vanishing gradient problem. However, the zero output is considered as the limitation of ReLU.

2.3.3. Exponential Linear Unit (ELU)
The ELU can be defined as
\[ f(x) = \alpha(e^x - 1), x < 0, \text{otherwise} f(x) = x \]
Where, α is the selected parameter and x shows the input.
For positive values of x, ELU function acts like the ReLU function, but in case of negative values, this function is bounded by a fixed value of α is -1. That is why it is beneficial for pushing the mean activation function of neurons closer to zero which is helpful in learning and make it more robust to noise.

2.4. Optimization of CNN Activation Functions
Proper training data can play a vital role in the development of efficient, more robust and decent neural networks. There exist several parameters that can greatly affect the performance of neural networks, including the activation functions, selection of number of layers, number of neurons and the availability of load data. To optimize the parameters, various cases were utilized as discussed in our previous section in order to investigate the performance of our proposed methodology.

3. Experimental Section
In this study, we optimized the deep CNN models based on layers, activation functions and number of neurons used in each layer to predict the ahead load forecasting demands. The study is specifically aimed to see the impact of optimizing the parameters to predict the ahead load forecasting demands.

In this study, we optimized the CNN model with different activation functions and number of neurons. We used ReLU and Sigmoid both functions with neurons 2, 4, 6, 8 10 and 12 neurons to predict 24 hours, one week, one month, and 3 months load ahead load forecasting. We also computed the ahead of forecasting with single and double layer. The performance was evaluated in terms of R², MSE, RMSE, MAE and MAPE. The Table 1 reflect the ahead load forecasting for one day with three different cases using optimization functions, layers and numbers of neurons used.

Case 1: To compute the one day ahead forecasting, with single layer and ReLU activation function, i) using number of neurons as 4, ahead prediction performance was obtained with R² (0.8193), MSE (199650.62), RMSE (446.82), MAE (355.73) and MAPE (2.41); ii) number of neurons as 6, the performance was obtained with R² (0.8367), MSE (1180397.79), RMSE (424.73), MAE (338.48) and MAPE (2.35); iii) number of neurons as 8, ahead prediction performance was obtained with R² (0.8259), MSE (192332.88), RMSE (438.56), MAE (347.83) and MAPE (2.42); iv) with number of neurons as 10, R² (0.8029), MSE (217788.08), RMSE (466.68), MAE (376.48) and MAPE (2.59); v) number of neurons as 12, R² (0.8013), MSE (219568.5), RMSE (468.58), MAE (382.86) and MAPE (2.64).

Case 2: To compute the one day ahead forecasting, with double layer and ReLU - ReLU activation functions, i) using number of neurons as 4, ahead prediction performance was obtained with R² (0.8132), MSE (206408.68), RMSE (454.32), MAE (354.79) and MAPE (2.47); ii) number of neurons as 6, the performance was obtained with R² (0.8178), MSE (201367.34), RMSE (424.73), MAE (372.26) and MAPE (2.57); iii) number of neurons as 8, ahead prediction performance was obtained with R² (0.8259), MSE (192332.88), RMSE (438.56), MAE (347.83) and MAPE (2.42).
MSE (218874.21), RMSE (467.84), MAE (395.70) and MAPE (2.69); iv) with number of neurons as 10, \( R^2 \) (0.8159), MSE (203377.73), RMSE (450.97), MAE (375.47) and MAPE (2.61); v) number of neurons as 12, \( R^2 \) (0.8060), MSE (214312.89), RMSE (462.94), MAE (382.82) and MAPE (2.64).

Case 3: To compute the one day ahead forecasting, with double layer and ReLU - Sigmoid activation functions, i) using number of neurons as 4, ahead prediction performance was obtained with \( R^2 \) (0.8304), MSE (187358.92), RMSE (432.85), MAE (327.62) and MAPE (2.23); ii) number of neurons as 6, the performance was obtained with \( R^2 \) (0.8313), MSE (186375.26), RMSE (431.71), MAE (330.64) and MAPE (2.25); iii) number of neurons as 8, ahead prediction performance was obtained with \( R^2 \) (0.8414), MSE (175220.93), RMSE (418.59), MAE (344.36) and MAPE (2.37); iv) with number of neurons as 10, \( R^2 \) (0.7964), MSE (225016.51), RMSE (474.36), MAE 398.98) and MAPE (2.73); v) number of neurons as 12, \( R^2 \) (0.8031), MSE (217597.19), RMSE (466.47), MAE (369.21) and MAPE (2.56).

Table 1. Prediction of the next one day ahead load forecasting with parameters optimization by choosing different neurons, and layers and activation function combinations

| Neurons | \( r^2 \) | MSE      | RMSE     | MAE      | MAPE   |
|---------|---------|----------|----------|----------|--------|
| Case 1: Single layer ReLU |
| 4       | 0.8193  | 199650.62| 446.82   | 355.73   | 2.41   |
| 6       | 0.8367  | 180397.79| 424.73   | 338.48   | 2.35   |
| 8       | 0.8259  | 192332.88| 438.56   | 347.83   | 2.42   |
| 10      | 0.8029  | 217788.08| 466.68   | 376.48   | 2.59   |
| 12      | 0.8013  | 219568.35| 468.58   | 382.86   | 2.64   |
| Case 2: Double layer ReLU-ReLU |
| 4       | 0.8132  | 206408.68| 454.32   | 354.79   | 2.47   |
| 6       | 0.8178  | 201367.34| 448.74   | 372.26   | 2.57   |
| 8       | 0.8019  | 218874.21| 467.84   | 395.70   | 2.69   |
| 10      | 0.8159  | 203377.73| 450.97   | 375.47   | 2.61   |
| 12      | 0.8060  | 214312.89| 462.94   | 386.82   | 2.64   |
| Case 3: Double layer ReLU-Sigmoid |
| 4       | 0.8304  | 187358.92| 432.85   | 327.62   | 2.23   |
| 6       | 0.8313  | 186375.26| 431.71   | 330.64   | 2.25   |
| 8       | 0.8414  | 175220.93| 418.59   | 344.36   | 2.37   |
| 10      | 0.7964  | 225016.51| 474.36   | 398.98   | 2.73   |
| 12      | 0.8031  | 217597.19| 466.47   | 369.21   | 2.56   |

Case 1: Load Forecast 24 hours ahead
Figure 3. One day ahead load forecasting with parametric optimization a) Single layer ReLU activation function, b) Double layer ReLU – ReLU activation function, c) Double layer ReLU – Sigmoid activation function

Figure 3 (a–c) reflect the one day ahead prediction performance with different number of neurons used. The forecasting showed the difference of actual load with different predicted load on number of neurons i.e. 2, 4, 6, 8, 10 and 12. In Figure 3 (a), the highest prediction was obtained with neuron 6 followed by neuron 4, 8, 10 and 12 respectively. The other prediction measures are reflected in pictorial representation accordingly.

The Table 2 reflect the ahead load forecasting for three months with three different cases using optimization functions, layers and numbers of neurons used.

Table 2. Prediction of the next three months ahead load forecasting with parameters optimization by choosing different neurons, and layers and activation function combinations

| Neurons | r²     | MSE      | RMSE   | MAE    | MAPE   |
|---------|--------|----------|--------|--------|--------|
| 4       | 0.9793 | 219282.58| 468.28 | 345.14 | 2.82   |
| 6       | 0.9797 | 215362.27| 464.07 | 340.38 | 2.76   |
| 8       | 0.9794 | 217950.60| 466.85 | 345.61 | 2.82   |
| 10      | 0.9788 | 224270.61| 473.57 | 352.08 | 2.89   |
| 12      | 0.9800 | 211554.75| 459.95 | 338.26 | 2.76   |
| 4       | 0.9794 | 218400.40| 467.33 | 345.60 | 2.83   |
| 6       | 0.9792 | 220562.81| 469.64 | 351.56 | 2.90   |
| 8       | 0.9789 | 223658.92| 472.93 | 346.11 | 2.82   |
Figure 4. Three months ahead load forecasting with parametric optimization a) Single layer ReLU activation function, b) Double layer ReLU – ReLU activation function, c) Double layer ReLU – Sigmoid activation function
Figure 4 (a-c) reflect the one month ahead prediction performance with different number of neurons used. The forecasting showed the difference of actual load with different predicted load on number of neurons i.e. 2, 4, 6, 8, 10 and 12. In Figure 4 (c), the highest prediction was obtained with neuron 6 followed by neuron 8, 10, 12, 4 and 8 respectively. The results reveal that for more complex and longer forecasting a complex combination of activation function and more neurons are desired to learn and predict more accurately.

Table 3. Comparison of result with the existing studies

| Method          | Day-ahead | Week-ahead | Month ahead | Three months |
|-----------------|-----------|------------|-------------|--------------|
| RDNN [33]       | 5.66%     | 7.55%      | -           | -            |
| SOM-SVM [34]    | 6.05%     | 8.03%      | -           | -            |
| Capula DBN [35] | 5.63%     | 7.26%      | -           | -            |
| LSTM [36]       | -         | 9.34%      | -           | -            |
| NP-ARMA [37]    | -         | -          | 4.83%       | -            |
| CNN [38]        | -         | -          | 5.08%       | -            |
| LSTM [39]       | -         | -          | 4.83%       | -            |
| XGBoost [40]    | -         | -          | 10.08%      | -            |
| DT [41]         | -         | -          | -           | 3.02%        |
| RNN [41]        | -         | -          | -           | 2.91%        |
| gate-RNN [41]   | -         | -          | -           | 3.23%        |
| CNN with default parameters [42] | 5.62% | 9.79% | 5.10% | - |

This work

i) Single layer (ReLU), 6 neurons
   2.35% 1.95% 2.78% 2.76%

ii) Double layer (ReLU-ReLU), 4/6/10 neurons
   2.47% 1.98% 2.84% 2.75%

iii) Double layer (ReLU-Sigmoid), 4/6/10 neurons
   2.23% 1.99% 2.78% 2.78%

4. Conclusion

An accurate electric load forecasting is indispensable due to its application in decision making and operations of power grid. Accurate forecasting system is very helpful in developing an optimal market plan to enhance economic benefits of energy management. A stable and robust load forecasting system is desired. In this study, we optimized the parameters of deep convolutional neural network model using single and double layer with ReLU and ReLU-ReLU, ReLU-Sigmoid activation functions and changed the number of neurons from 2, 4, 6, 8, 10 and 12. To predict the one day ahead load forecasting, the lowest prediction error was yielded using double layer with ReLU – Sigmoid activation function yielding MAPE (4.23). The results revealed that proposed method can be implemented more efficiently in predicting accurate load forecasting system. This system will help for an efficient energy management and to manage the operations and load demands in an intelligent way. This system will help to provide efficient system to meet the local demands of our country.

References
[1] M. Božić, M. Stojanović, Z. Stajić, D. Tasić, A New Two-Stage Approach to Short Term Electrical Load Forecasting, Energies. 6 (2013) 2130–2148. https://doi.org/10.3390/en6042130.
[2] R.-H. Liang, Y.-K. Chen, Y.-T. Chen, Volt/Var control in a distribution system by a fuzzy optimization approach, Int. J. Electr. Power Energy Syst. 33 (2011) 278–287. https://doi.org/10.1016/j.ijepes.2010.08.023.
[3] T. Niknam, M. Zare, J. Aghaei, Scenario-Based Multiobjective Volt/Var Control in Distribution Networks Including Renewable Energy Sources, IEEE Trans. Power Deliv. 27 (2012) 2004–2019. https://doi.org/10.1109/TPWRD.2012.2209900.
C.-M. Lee, C.-N. Ko, Short-term load forecasting using lifting scheme and ARIMA models, Expert Syst. Appl. 38 (2011) 5902–5911. https://doi.org/10.1016/j.eswa.2010.11.033.

S.S. Pappas, L. Ekonomou, D.C. Karamousantas, G.E. Chatzarakis, S.K. Kotisakia, P. Liatsis, Electricity demand loads modeling using AutoRegressive Moving Average (ARMA) models, Energy. 33 (2008) 1353–1360. https://doi.org/10.1016/j.energy.2008.05.008.

S.S. Pappas, L. Ekonomou, P. Karampelas, D.C. Karamousantas, S.K. Katsikas, G.E. Chatzarakis, P.D. Skafidas, Electricity demand load forecasting of the Hellenic power system using an ARMA model, Electr. Power Syst. Res. 80 (2010) 256–264. https://doi.org/10.1016/j.epsr.2009.09.006.

N. Kandil, R. Wamkeue, M. Saad, S. Georges, An Efficient Approach for Shorter Term Load Forecasting using Artificial Neural Networks, in: 2006 IEEE Int. Symp. Ind. Electron., IEEE, 2006: pp. 1928–1932. https://doi.org/10.1109/ISIE.2006.295867.

F. Yu, X. Xu, A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network, Appl. Energy. 134 (2014) 102–113. https://doi.org/10.1016/j.apenergy.2014.07.104.

S. Shenoy, D. Gorinevsky, S. Boyd, Non-parametric regression modeling for stochastic optimization of power grid load forecast, in: 2015 Am. Control Conf., IEEE, 2015: pp. 1010–1015. https://doi.org/10.1109/ACC.2015.7170865.

M.S. Kandil, S.M. El-Debeify, N.E. Hasanien, Long-term load forecasting for fast developing utility using a knowledge-based expert system, IEEE Trans. Power Syst. 17 (2002) 491–496. https://doi.org/10.1109/TPWRS.2002.1007923.

G. Dudek, Pattern-based local linear regression models for short-term load forecasting, Electr. Power Syst. Res. 130 (2016) 139–147. https://doi.org/10.1016/j.epsr.2015.09.001.

W. Christiaanse, Short-Term Load Forecasting Using General Exponential Smoothing, IEEE Trans. Power Appar. Syst. PAS-90 (1971) 900–911. https://doi.org/10.1109/TPAS.1971.293123.

M.Q. Raza, A. Khosravi, A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings, Renew. Sustain. Energy Rev. 50 (2015) 1352–1372. https://doi.org/10.1016/j.rser.2015.04.065.

L.C.P. Velasco, C.R. Villegas, P.N.C. Palahang, J.A.A. Dagaang, Next day electric load forecasting using Artificial Neural Networks, in: 2015 Int. Conf. Humanoid, Nanotechnology, Inf. Technol. Control. Environ. Manag., IEEE, 2015: pp. 1–6. https://doi.org/10.1109/HNICEM.2015.7393166.

L. Hernández, C. Baladrón, J. Aguilar, L. Calavia, B. Carro, A. Sánchez-Esguevillas, F. Pérez, Á. Fernández, J. Lloret, Artificial Neural Network for Short-Term Load Forecasting in Distribution Systems, Energies. 7 (2014) 1576–1598. https://doi.org/10.3390/en7031576.

J. Buitrags, S. Asfour, Short-Term Forecasting of Electric Loads Using Nonlinear Autoregressive Artificial Neural Networks with Exogenous Vector Inputs, Energies. 10 (2017) 40. https://doi.org/10.3390/en10010040.

L. Suganthi, S. Iniyan, A.A. Samuel, Applications of fuzzy logic in renewable energy systems – A review, Renew. Sustain. Energy Rev. 48 (2015) 585–607. https://doi.org/10.1016/j.rser.2015.04.037.

D. Niu, H. Shi, D.D. Wu, Short-term load forecasting using bayesian neural networks learned by Hybrid Monte Carlo algorithm, Appl. Soft Comput. 12 (2012) 1822–1827. https://doi.org/10.1016/j.asoc.2011.07.001.

Kwang-Ho Kim, Jong-Keun Park, Kab-Ju Hwang, Sung-Hak Kim, Implementation of hybrid short-term load forecasting system using artificial neural networks and fuzzy expert systems, IEEE Trans. Power Syst. 10 (1995) 1534–1539. https://doi.org/10.1109/59.466492.

F. Kaytez, M.C. Taplamacioglu, E. Cam, F. Hardalac, Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines, Int. J. Electr. Power Energy Syst. 67 (2015) 431–438. https://doi.org/10.1016/j.ijepes.2014.12.036.

A. Selakov, D. Cvijetinović, L. Milović, S. Mellon, D. Bekut, Hybrid PSO–SVM method for short-term load forecasting during periods with significant temperature variations in city of Burbank, Appl. Soft Comput. 16 (2014) 80–88. https://doi.org/10.1016/j.asoc.2013.12.001.

D. Niu, S. Dai, A Short-Term Load Forecasting Model with a Modified Particle Swarm Optimization Algorithm and Least Squares Support Vector Machine Based on the Denoising Method of Empirical Mode Decomposition and Grey Relational Analysis, Energies. 10 (2017) 408. https://doi.org/10.3390/en10030408.

Y. Liang, D. Niu, M. Ye, W.-C. Hong, Short-Term Load Forecasting Based on Wavelet Transform and Least Squares Support Vector Machine Optimized by Improved Cuckoo Search, Energies. 9 (2016) 827. https://doi.org/10.3390/en9100827.

J. Günther, P.M. Pilarski, G. Helfrich, H. Shen, K. Diepold, First Steps Towards an Intelligent Laser
Welding Architecture Using Deep Neural Networks and Reinforcement Learning, Procedia Technol. 15 (2014) 474–483. https://doi.org/10.1016/j.protcy.2014.09.007.

[25] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, in: 2014 IEEE Conf. Comput. Vis. Pattern Recognit., IEEE, 2014: pp. 580–587. https://doi.org/10.1109/CVPR.2014.81.

[26] E. Shelhamer, J. Long, T. Darrell, Fully Convolutional Networks for Semantic Segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (2017) 640–651. https://doi.org/10.1109/TPAMI.2016.2572683.

[27] Y. Taigman, M. Yang, M. Ranzato, L. Wolf, DeepFace: Closing the gap to human-level performance in face verification, in: Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2014: pp. 1701–1708. https://doi.org/10.1109/CVPR.2014.220.

[28] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, L. Fei-Fei, Large-scale video classification with convolutional neural networks, Comput. Vis. Pattern Recognit. (CVPR), 2014 IEEE Conf. (2014) 1725–1732. https://doi.org/10.1109/CVPR.2014.223.

[29] Z. Ghahramani, Unsupervised Learning, in: 2004: pp. 72–112. https://doi.org/10.1007/978-3-540-28650-9_5.

[30] A. Karpathy, L. Fei-Fei, Deep visual-semantic alignments for generating image descriptions, in: 2015 IEEE Conf. Comput. Vis. Pattern Recognit., IEEE, 2015: pp. 3128–3137. https://doi.org/10.1109/CVPR.2015.7298932.

[31] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proc. IEEE. 86 (1998) 2278–2324. https://doi.org/10.1109/5.726791.

[32] T. Homma, L.E. Atlas, R.J.M.I.T. Homma, L.E. Atlas, R.J.M. II, An artificial neural network for spatio-temporal bipolar patterns: Application to phoneme classification, Neural Inf. Process. (1988) 31–40. https://doi.org/10.1 (1988) 31–40.

[33] G.M.U. Din, A.K. Marnerides, Short term power load forecasting using Deep Neural Networks, in: 2017 Int. Conf. Comput. Netw. Commun., IEEE, 2017: pp. 594–598. https://doi.org/10.1109/ICCCNC.2017.7876196.

[34] S. Fan, L. Chen, Short-Term Load Forecasting Based on an Adaptive Hybrid Method, IEEE Trans. Power Syst. 21 (2006) 392–401. https://doi.org/10.1109/TPWRS.2005.860944.

[35] T. Ouyang, Y. He, H. Li, Z. Sun, S. Baek, Modeling and Forecasting Short-Term Power Load With Copula Model and Deep Belief Network, IEEE Trans. Emerg. Top. Comput. Intell. 3 (2019) 127–136. https://doi.org/10.1109/TETCI.2018.2880511.

[36] X. Wang, F. Fang, X. Zhang, Y. Liu, L. Wei, Y. Shi, LSTM-based Short-term Load Forecasting for Building Electricity Consumption, in: 2019 IEEE 28th Int. Symp. Ind. Electron., IEEE, 2019: pp. 1418–1423. https://doi.org/10.1109/ISIE.2019.8781349.

[37] I. Shah, H. Iftikhar, S. Ali, Modeling and Forecasting Medium-Term Electricity Consumption Using Component Estimation Technique, Forecasting. 2 (2020) 163–179. https://doi.org/10.3390/forecast20009.

[38] S.H. Rafi, Nahid-Al-Masood, Short Term Electric Load Forecasting Using High Precision Convolutional Neural Network, in: 2020 Int. Conf. Comput. Electr. Commun. Eng., IEEE, 2020: pp. 1–7. https://doi.org/10.1109/ICCECE48148.2020.9223102.

[39] S.H. Rafi, N.-A.-M. Nahid-Al-Masood, Highly Efficient Short Term Load Forecasting Scheme Using Long Short Term Memory Network, in: 2020 8th Int. Electr. Eng. Congr., IEEE, 2020: pp. 1–4. https://doi.org/10.1109/iECON48109.2020.229546.

[40] R.A. Abbasi, N. Javaid, M.N.J. Ghuman, Z.A. Khan, S. Ur Rehman, Amanullah, Short Term Load Forecasting Using XGBoost, in: 2019 IEEE 28th Int. Symp. Ind. Electron., IEEE, 2019: pp. 1120–1131. https://doi.org/10.1109/ISIE.2019.8781349.

[41] F. Mehmood Butt, L. Hussain, A. Mahmood, K. Javed Lone, Artificial Intelligence based accurately load forecasting system to forecast short and medium-term load demands, Math. Biosci. Eng. 18 (2021) 400–425. https://doi.org/10.3934/mbe.2021022.