The Eurasia Proceedings of Science, Engineering & Mathematics (EPSTEM)
ISSN: 2602-3199

The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 2018

Volume 2, Pages 233-238

ICRES 2018: International Conference on Research in Education and Science

A Comparison of the Performance of Classification Methods and Artificial Neural Networks for Electricity Load Forecasting

Gurkan TUNA
Trakya University

Ahmet VATANSEVER
Trakya University

Resul DAS
Fırat University

Abstract: Electricity load forecasting plays a key role for utility companies. Short-term and medium-term electricity load forecasting processes allow the utility companies to retain reliable operation and high energy efficiency. On the other hand, long-term electricity load forecasting allows the utility companies to minimize the risks. Long-term forecasting also helps the utility companies to plan and make feasible decisions in regard to generation and transmission investments. Since there are commercial and technical implications of electricity load forecasting, the accuracy of the electricity forecasting is important not only to the utility companies but also to the consumers. In this paper, we carry out a performance evaluation study to evaluate the accuracy of different classification approaches for electricity load forecasting. As shown with the results of the performance evaluation study, some of the investigated approaches can successfully achieve high accuracy rates and therefore can be used for short-, mid-, or long-term electricity load forecasting.

Keywords: Load-Forecasting plan, Artificial neural networks, Regression analysis, Support vector machine, Prediction techniques

Introduction

Thanks to the understanding of the future consumption provided by electricity load forecasting, utility companies obtain many benefits. By carrying out electricity load forecasting, utility companies plan well for the future, determine the required resources to ensure uninterrupted power to the consumers, utilize the generating plants efficiently, decide easily the best time with the minimum impact for maintenance of the power systems, and minimize the risks via economically viable decisions for future investments (Weron, 2006; Suganthi, & Samuel, 2012; Hernandez et al., 2014).

One of the most important decisions that utility companies must make is whether in the near future they need more generating plants or not and if yes what the type, size and location of the generating plants will be. In this way, the utility companies will be able to determine areas with growing demand and generate the power near the load (Suganthi, & Samuel, 2012; Almeshaiei, & Soltan, 2011). This also enables them to minimize the transmission and distribution infrastructures and reduce the associated losses.

Well-known prediction approaches such as artificial neural network models, and support vector machines and linear regression trees can be used for electricity load forecasting (Hastie, Tibshirani, & Friedman, 2009; Hernandez et al., 2014). An emerging approach for electricity load forecasting is the use of a combination of the well-known prediction approaches. Different from the existing studies, instead of focusing on the use of a prediction algorithm for electricity load forecasting, in this study we mainly focus on the performance comparison of the existing prediction algorithms used for electricity load forecasting. The dataset used in this

- Selection and peer-review under responsibility of the Organizing Committee of the Conference

study consists of the actual consumption of Ankara in Turkey and was obtained during the period from December 2011 to April 2013. The remainder of this paper is structured as follows. Section 2 reviews the prediction approaches used in this study. Discussion on the performance of the reviewed approaches is given in Section 3. Finally, this paper is concluded in Section 4.

Method

In this study, a dataset that consists of 12168 rows was obtained from Republic of TURKEY, Energy Market Regulatory Authority (EMRA) and used. The dataset consist of 12168 rows. Each row of the dataset consists of hour, day of week, month, year, temperature of Ankara, and electricity load. In the first step of the evaluation study, Linear Regression, Multilayer Perceptron and Support Vector Machines prediction techniques were preferred. After a pre-filtering step, the techniques were first implemented in WEKA (Waikato Environment for Knowledge Analysis) (https://www.cs.waikato.ac.nz/ml/weka/) and the accuracy of the techniques were compared. In parallel with the studies in the literature (Guerard, & Schwartz, 2010) correlation coefficient, one of the most commonly used indicators of forecasting accuracy, was used to compare the performance of the employed techniques. Correlation coefficient is basically a number between 0 and 1 and a measure of how well the predicted values from a forecast model fit with the real data (Guerard, & Schwartz, 2010). If the correlation coefficient is 0 or very low, there is not any relationship between the predicted values and the actual values. However, if the correlation coefficient is 1 or very high, there is strong relationship between the predicted values and actual values (Yan, & Su, 2009).

In the second step of the evaluation study, Artificial Neural Network was implemented in MATLAB (https://www.mathworks.com/products/matlab.html) for electricity load forecasting of Ankara. However, before carrying out the evaluation study, the dataset that consists of 12168 rows was first divided into training, validation and test datasets as listed in Table 1. Due to the required number of features, the number of hidden neurons was set to 5. Levenberg - Marquardt algorithm (Reynaldi, Lukas, & Margaretha, 2012) was preferred for the training phase.

Table 1. Percentages of training, validation and test phases for Ankara

Phase	Percentage (%)	Total Number of Rows
Training	70	8517
Validation	20	2434
Testing	10	1217

Results and Discussion

The first prediction technique used in the performance evaluation was Linear Regression (Yan, & Su, 2009). The results of Linear Regression technique for electricity load forecasting are listed in Table 2 and the forecasting errors are listed in Table 3. The second prediction technique used in the performance evaluation was Multilayer Perceptron (Popescu et al., 2009). The results of Multilayer Perceptron technique for electricity load forecasting are listed in Table 4 and the forecasting errors of Multilayer Perceptron technique are given in Table 5. Finally, the last technique used in the performance evaluation was Support Vector Machines (Steinwart, 2014). The results of Support Vector Machines technique for electricity load forecasting are listed in Table 6 and the forecasting errors of Support Vector Machines technique are listed in Table 7. As given by the results presented in Table 2, Table 4 and Table 6, when correlation coefficient is considered, Multilayer Perceptron technique achieved the highest forecasting accuracy.

Table 2. Classification results for Ankara when Linear Regression was employed

Term	Value
Correlation Coefficient	0.5955
Average Absolute Error	2650.1462
Root Mean Square Error	3299.8781
Relative Absolute Error	% 76.1906
Root Relative Squared Error	% 80.3389
Total Number of Rows	12168

Table 3. Error values for Ankara when Linear Regression was employed (Please note that only the first 10 rows were given.)

Row	Actual Value	Predicted Value (MW)	Error
1	26400	22937.41	-3462.59
2	24700	23254.628	-1445.372
3	23800	23546.02	-253.98
4	23400	23837.412	437.412
5	23300	24102.977	802.977
6	23700	24497.676	797.676
7	24300	24944.028	644.028
8	25700	25467.862	-232.138
9	29300	25965.868	-3334.132
10	31500	26515.528	-4984.472

Table 4. Classification results for Ankara when Multilayer Perceptron was employed

Term	Value
Correlation Coefficient	0.7225
Average Absolute Error	2370.6511
Root Mean Square Error	2990.0692
Relative Absolute Error	% 68.1553
Root Relative Squared Error	% 72.7962
Total Number of Rows	12168

Table 5. Error values for Ankara when Multilayer Perceptron was used (Please note that only the first 10 rows were given.)

Row	Actual Value (MW)	Predicted Value (MW)	Error
1	26400	24425.549	-1974.451
2	24700	24538.664	-161.336
3	23800	24692.135	892.135
4	23400	24897.205	1497.205
5	23300	25166.922	1866.922
6	23700	25489.836	1789.836
7	24300	25872.411	1572.411
8	25700	26325.012	625.012
9	29300	26888.114	-2411.886
10	31500	27782.122	-3717.878

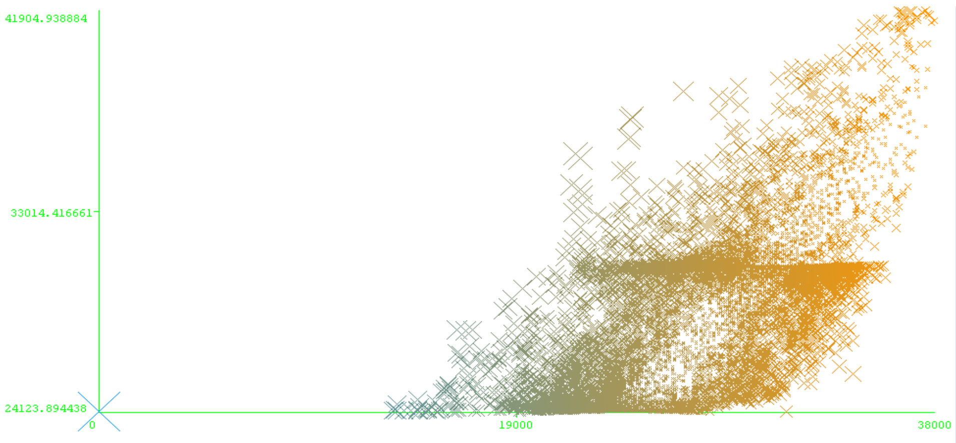

Figure 1. The actual values (X coordinate) vs predicted values (Y coordinate) when Multilayer Perceptron was employed

Table 6. Classification results for Ankara when Support Vector Machine was used

Term	Value
Correlation Coefficient	0.5938
Average Absolute Error	2640.5974
Root Mean Square Error	3310.1538
Relative Absolute Error	% 75.9161
Root Relative Squared Error	% 80.589
Total Number of Rows	12168

Table 7. Error values for Ankara when Support Vector Machine was used (Please note that only the first 10 rows were given.)

Row	Actual Value	Predicted Value (MW)	Error
1	26400	22673.061	-3726.939
2	24700	22994.054	-1705.946
3	23800	23283.813	-516.187
4	23400	23573.572	173.572
5	23300	23832.098	532.098
6	23700	24246.79	546.79
7	24300	24723.947	423.947
8	25700	25294.804	-405.196
9	29300	25834.428	-3465.572
10	31500	26436.517	-5063.483

An overview of forecasting results for Ankara is given in Figure 2. In this figure, the MSE value shows the average of the difference between the desired output and the current output of the artificial neural network. *R* represents the correlation between the actual values and the predicted values. Being the coefficient of correlation, *R* ranges from -1 to +1. If *R* is closer to +1 or -1, the two variables are related (Hastie, Tibshirani, & Friedman, 2009). On the other hand, if *R* is close to 0, it means that there is no relationship between the variables.

As shown in the error histogram shown in Figure 3, the target (actual) values are the values that the artificial neural network based model was expected to produce. The output values are values that the artificial neural network based model obtained. The error values show the margin between the target values and the output values. Results of the regression analysis for Ankara when Artificial Neural Network was employed are shown in Figure 4. The values of *R* and the results of the regression analysis show that the employed artificial neural network based model obtained satisfactory forecasting accuracy. To sum up, various statistical forecasting techniques can be used for electricity load forecasting if thousands of accurate data samples are available to be processed automatically.

	Samples	MSE	R
Training:	8517	3804642.77007e-0	8.81202e-1
Validation:	2434	3482206.47683e-0	8.91388e-1
Testing:	1217	3727131.82955e-0	8.73353e-1

Figure 2. An overview of forecasting results for Ankara

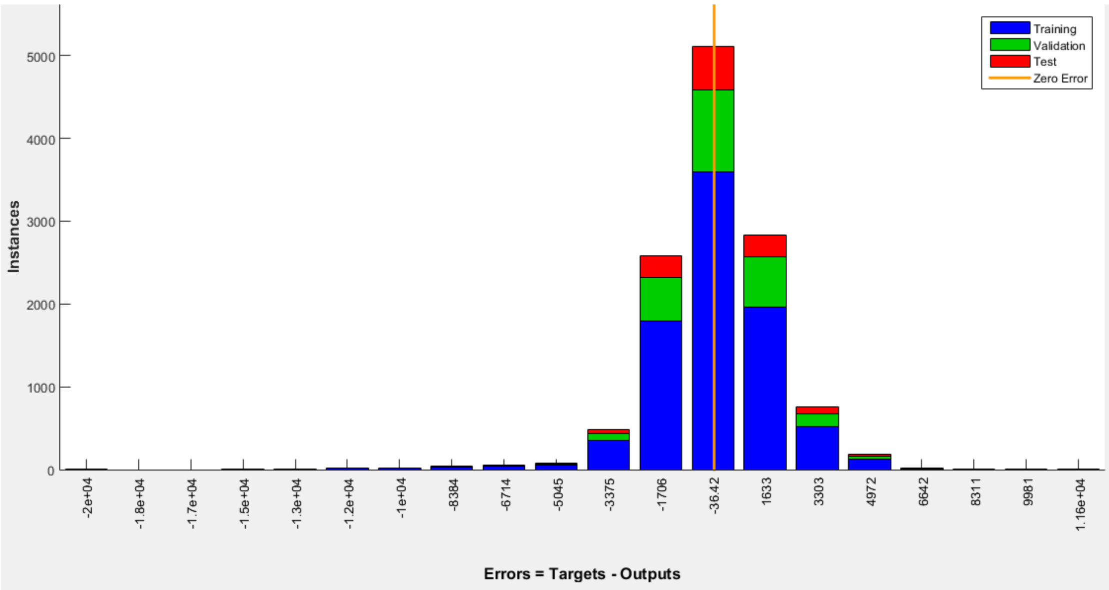

Figure 3. Error histogram for Ankara when Artificial Neural Network was employed

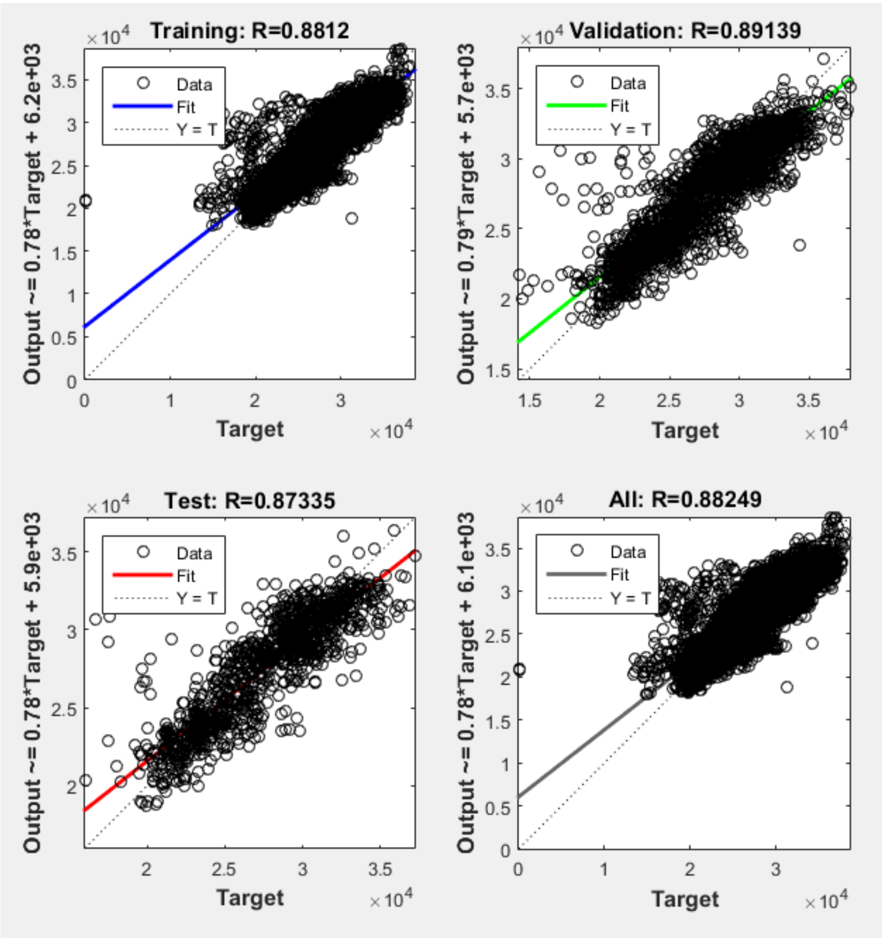

Figure 4. Results of the regression analysis for Ankara when Artificial Neural Network was employed

Conclusion

Electricity load forecasting is the predicting of electrical power required to meet the short-term, medium-term or long-term demand. It not only helps utility companies in their operation and management of the supply to their customers but also aids in planning on their capacity and operations so that all the customers can be supplied reliably with the required energy. In addition, it is an important process that contributes to the efficiency and revenues for the utility companies. Considering all these benefits, in this paper, we realized a performance evaluation study to evaluate the accuracy of different classification approaches for electricity load forecasting. The results of the performance evaluation study show that electricity load forecasting can be realized successfully if careful analysis is made and a satisfactory dataset is available. The main limitation of this study is that since seasons and other factors may affect the way customers use the power, these factors should be taken into consideration for electricity load forecasting.

References

Almeshaiei, E., & Soltan, H. (2011). A methodology for Electric Power Load Forecasting. *Alexandria Engineering Journal, 50*(2), 137-144.

Guerard, J. B., & Schwartz, E. (2010). *Quantitative corporate finance*. New York: Springer.

Hastie, T. J., Tibshirani, R. J., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer New York Inc., USA.

Hernandez, L., Baladron, C., Aguiar, J. M., Carro, B., Sanchez-Esguevillas, A. J., Lloret, J., & Massana, J. (2014). A survey on electric power demand forecasting: Future trends in smart grids, microgrids and smart buildings. *IEEE Communications Surveys & Tutorials*, *16*(3), 1460-1495.

Popescu, M. –C., Balas, V. E., Perescu-Popescu, L., & Mastorakis, N. (2009). Multilayer perceptron and neural networks. *WSEAS Trans. Cir. and Sys., 8*(7), 579-588.

Reynaldi, A., Lukas, S., & Margaretha, H. (2012). Backpropagation and Levenberg-Marquardt Algorithm for Training Finite Element Neural Network. *Proceedings of the 2012 Sixth UKSim/AMSS European Symposium on Computer Modeling and Simulation (EMS '12)* (pp. 89-94).

Steinwart, I. (2014). *Support vector machines*. Springer.

Suganthi, L., & Samuel, A. A. (2012). Energy models for demand forecasting—A review. *Renewable and Sustainable Energy Reviews, 16*(2), 1223-1240.

Weron, R. (2006). *Modeling and forecasting electricity loads and prices*. Chichester: Wiley & Sons.

Yan, X., & Su, X. G. (2009). *Linear regression analysis theory and computing*. Singapore: World Scientific.

https://www.cs.waikato.ac.nz/ml/weka/

https://www.mathworks.com/products/matlab.html

Author Information

Gurkan Tuna
Trakya University
Edirne, Turkey
Contact e-mail: *gurkantuna@trakya.edu.tr*

Ahmet Vatansever
Trakya University
Edirne, Turkey

Resul Das
Fırat University
Elazığ, Turkey