A Comparative Study of Machine Learning Algorithms for Short-term Electrical Load Forecasting

Wesam Yehia
Computer Science Department, Faculty of Computers and Information, Menoufia University, Shebin Elkom 32511, Egypt,
wesam.yehai89@gmail.com

Mahmoud Hussein
Computer Science Department, Faculty of Computers and Information, Menoufia University, Shebin Elkom 32511, Egypt,
fci_3mh@yahoo.com ,

Ashraf B. Elsisi
Computer Science Department, Faculty of Computers and Information, Menoufia University, Shebin Elkom 32511, Egypt,
ashrafelsisim@yahoo.com

Abstract— **Electrical load forecasting is an important field, where it can be used to estimate the amount of electricity needed, the price of electricity, and the number of generators to be used. Such forecasting can be performed using machine learning approaches. Therefore, in this paper, we compare four machine learning algorithms that can be used to predict electrical load. These algorithms are Support Vector Machine (SVM), Least Square Support Vector Machine (LSSVM), Gradient Boosting Machines (GBM), and Random Forest Regression (RF). The comparison is done on three months hourly recorded data set that is publicly available from Pennsylvania Jersey Maryland (PJM) company. Our contribution is the identification of which algorithm has a better accuracy and less execution time. This enables electricity companies concerned with load forecasting to choose an algorithm from the existing ones as fast as possible without wasting time searching for an appropriate algorithm. The results show that RF achieves the best accuracy, and also has the least execution time.**

Keywords— Machine Learning, Electrical Load Forecasting, SVM, LSSVM, GBM, RF

I. INTRODUCTION

Load forecasting is a technique used by power companies to predict the power or energy needed to balance the supply and load demand at all the times [1]. Accurate load forecasting can decrease the fluctuating behavior between energy generation and consumption [2], where the number of generators and the amount of electricity needed can be estimated.

A reliable forecasting of electrical load demand could help to avoid problems resulted from unexpected loads, and give vital information to make decisions on energy generation and purchase, especially market-based dynamic pricing strategies. Furthermore, accurate prediction would have a significant impact on operation management, e.g., preventing overloading and allowing an efficient energy storage [3].

Machine learning can handle time series data (that is sequence of observations taken sequentially in time) [4] by reconstructing the time series data to look like a supervised learning problem by using previous time steps as input variables and use the next time step as the output variable. Thus, machine learning algorithms can be used for load forecasting where load data is a time series data.

There are different and many algorithms for electrical load forecasting, so we provide this study to enable electricity companies concerned with load forecasting to choose an algorithm from the existing ones as fast as possible and without wasting time searching for an appropriate algorithm to load predictions.

In this paper, we provide a comparative study between four machine learning algorithms SVM [5], LSSVM [6], GBM [7], and RF [8] in terms of forecasting accuracy and the time taken by them to perform the prediction.

To compare the four algorithms, we use Mean Absolute Percentage Error (MAPE), and execution time as the evaluation metrics for the algorithms on a publicly available dataset. The dataset is taken from PJM company [9]. The dataset consists of 2208 observations. It is hourly-recorded representing electricity load of Western United States for three months from October 2017 to December 2017. We executed the algorithms ten times and each time we calculate their MAPE and execution time. After calculating the average of the ten iterations, we found that the RF Algorithm achieves the best accuracy, and also has the least execution time (i.e., the RF is the best candidate for performing load forecasting).

The rest of the paper is divided into three sections. In the first section, we explain general steps of forecasting in machine learning, and give an overview about the four algorithms we compare. In the second section, we provide the comparison between the four proposed algorithms. The last section explains the main conclusions of this study, and the future work.

II. ELECTRICAL LOAD FORECASTING

There are four types of forecasting according to the forecasting period [10]. The first type is the very short-term forecasting whose period is less than a day [11]. The second type is the short-term forecasting whose period ranges from one day to a week [12]. The third type is medium forecasting ranges from more than one week to year ahead [13]. The fourth is the long-term forecasting whose period is more than one year ahead [14].

In this paper, we focus on short-term load forecasting as such type of forecasting is very critical for the continues supply of electricity to the customers. In the following, we present the forecasting steps in general, and describe a set of algorithms that can be used to perform the forecasting.

A. Forecasting Steps

There are five basic steps of forecasting [15] which are :

1. Problem definition: define what we want to forecast and who are concerned with this forecasting.

2. Gathering information: collect historical data required for making the forecasting.

3. Preliminary(exploratory)analysis: analyze the data to identify its patterns (trends, seasons…etc.).

4. Choosing and fitting models: choose the appropriate models for the data and fitting them.

5. Using and evaluating a forecasting model: use the chosen model for making forecasting and evaluate it.

We use these steps to perform short-term load forecasting and compare the different algorithms. Such that, First, we define the problem which is predicting the next 50 electrical load values that is considered as short-term load forecasting. Second, we get the data from an electricity company called PJM in Pennsylvania. Third, we don't need to analyze data in our paper. Fourth, we choose the four proposed algorithms (SVM, LSSVM, RF and GBM) and fitting them to the data. Fifth, we use the four algorithms and evaluate them in terms of accuracy and execution time.

B. Algorithms

There are many algorithms for short-term load forecasting but we choose SVM, LSSVM, GBM, and RF because they have acceptable accuracy and less execution time. In the following, a short overview about each of them is introduced.

1) SVM

SVM is a machine-learning algorithm used for solving regression problem and provides a proficient prediction model [5]. It is known for its kernel trick to handle nonlinear input spaces. Thus, we use SVM in our paper as a load predictor because of the nonlinearity of the load, where there is a different load value for each time step.

SVM divides the data into appropriate categories by making a hyperplane between them (see the two red lines in Figure 1 as the decision boundary and the green line as the hyper-plane). The objective is to basically consider the points that are within the decision boundary line, and the best fit line is the hyper-plane that has a maximum number of points.
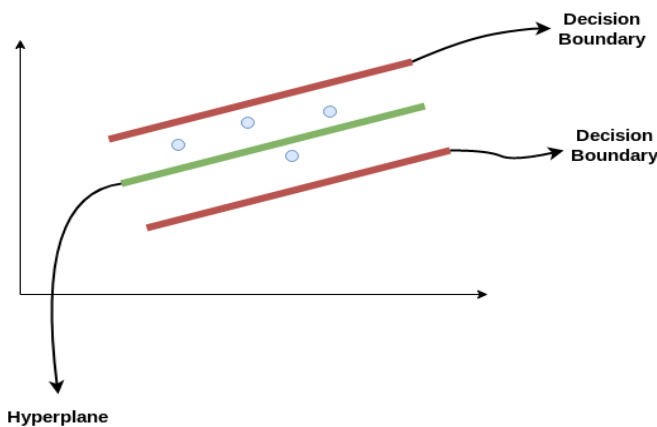

Fig. 1. *SVM*

2) LSSVM

In LSSVM [6], the input data are mapped with high-dimensional feature space with the help of kernel functions. Using kernel functions, the problem can be mapped into a linear form. LSSVM is a modification of the standard SVM, but LSSVM uses a least square cost function to construct the optimization problem. It is based on equality constraints which results in a series of linear equations instead of a quadratic programming problem that will reduce the

calculation complexity. LSSVM equation for function estimation is shown in Equation (1).

$$y(x) = \sum_{k=1}^{N} \propto_k k(x, x_k) + b \qquad (1)$$

Where α_k (w) is the weighting factor, x is the training samples and x_k is the support vectors, b represents the bias, and N is size of the training samples. LSSVM equation is explained in Figure 2.
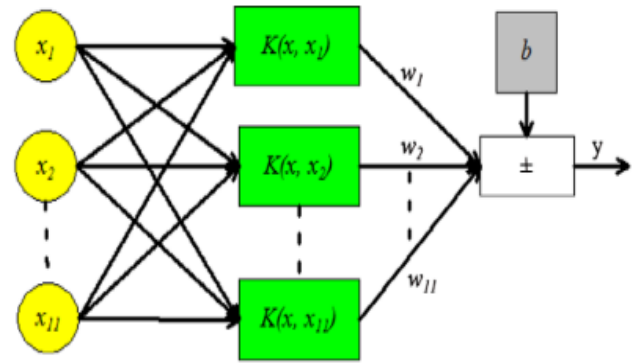

Fig. 2. *LSSVM*

3) RF

RF is a supervised learning algorithm built based on the concept of wisdom of crowds [7], where a large number of relatively uncorrelated models (trees) operating as a committee will outperform any of the individual constituent models . Thus, trees protect each other from their individual errors (i.e., different trees cannot get the same error), and this results in improving the accuracy and reducing overfitting. The general steps of RF are shown in Figure 3:

- The dataset is sampled randomly and each tree is created from a different sample of rows and at each node, a different sample of features is selected for splitting.
- Each of the trees makes its own individual prediction.
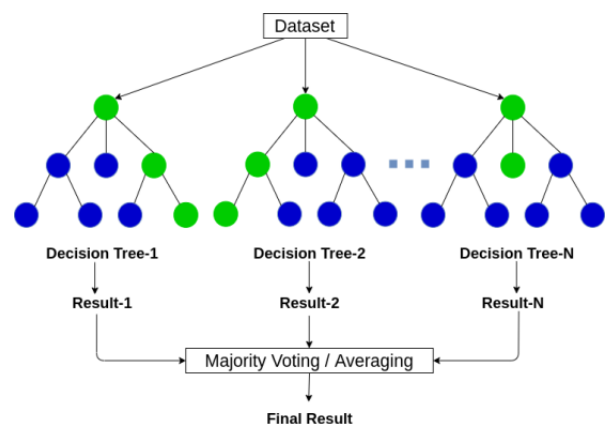- The predictions are then averaged to produce a single result.

Fig. 3. *RF*

4) GBM

GBM algorithm is a tree-based machine learning algorithm used for solving classification and regression problems [8]. The general steps of GBM are:

- It creates fixed number of decision trees which are called as weak learners or weak predictive models. These decision trees are of fixed size or depth as shown in Figure 4.
- These trees are built using residuals. The predicted value produced by each (model) is scaled by the learning rate.
- This learning rate enables the algorithm to have a more gradual and steady improvement at each step. The algorithm continues to build trees in this fashion until it has made the number of trees or additional trees fail to improve the fit .
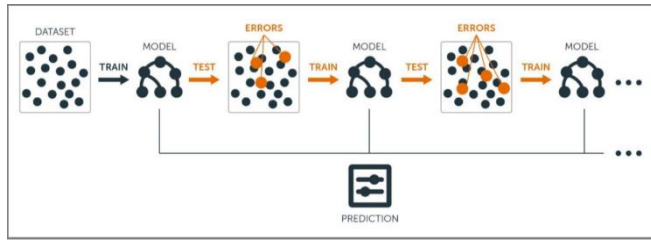

Fig. 4. *GBM*

III. RESULTS

To compare the four proposed machine learning algorithms (SVM, LSSVM, RF and GBM), we have implemented them in Python 3.8.3, and executed them on a laptop with the following specifications: Processor Intel(R) Core (TM) i3-CPU@2.4GHz, and 8GB RAM with Windows 10.

There are different evaluation metrics used to evaluate machine learning algorithms such as Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), Execution time, etc. MAE, MAPE, MSE metrics are used as accuracy indication, where the algorithm with less MAPE has more accuracy and vice versa. Execution time is the time required for a forecast model to complete its execution .

In this paper, we use execution time and MAPE to perform the comparison. MAPE allows comparison between regression models designed for predicting diverse categories of data [16]. Therefore, it is convenient because we have diverse values of load. MAPE can be calculated using Equation (2).

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{A_i - F_i}{A_i} \right| \qquad (2)$$

Where n is the number of data points, A_i represents the actual value, and F_i is the forecast value returned from the algorithm.

For the comparison, we use a dataset consisting of 2208 hourly-recorded observations representing electricity load of Western United States for three months from October 2017 to December 2017. The dataset is randomly split with the ratio of 70:30, where 70% used to train the algorithms while the 30% is used for testing them. Then, we use the algorithms to predict the next 50 hours and calculate MAPE for each algorithm by using Equation (2). In the following, we present the results obtained for each algorithm.

A. SVM

Prediction of the next 50 hours of load forecasting using SVM is shown in Figure 5. SVM model can predict future load values with MAPE of 0.14357, and execution time of 0.70 seconds.
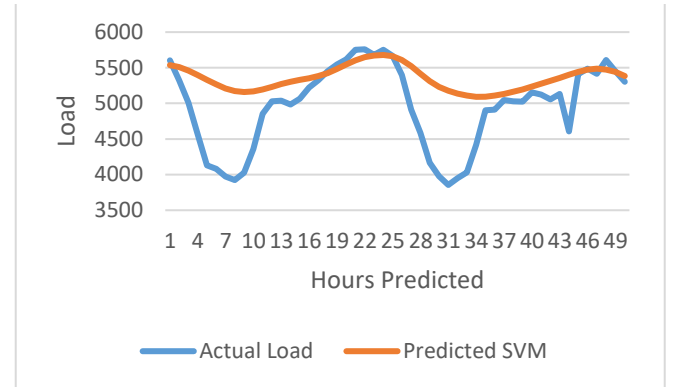

Fig. 5. SVM predicted values

B. LSSVM

Figure 6 shows the prediction of the next 50 hours when LSSVM is used. The LSSVM model's MAPE is 0.05152, while its execution time is 10.70 seconds.
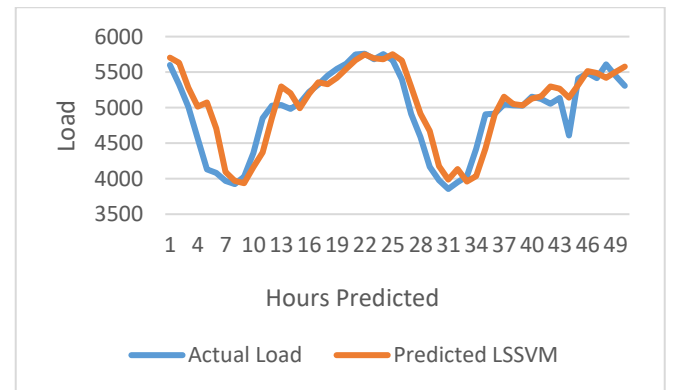

Fig. 6. LSSVM predicted values

C. RF

When RF model is used, It yielded MAPE of 0.041883 and execution time of 0.6 seconds. The prediction of the next 50 hours by using RF is shown in Figure 7.
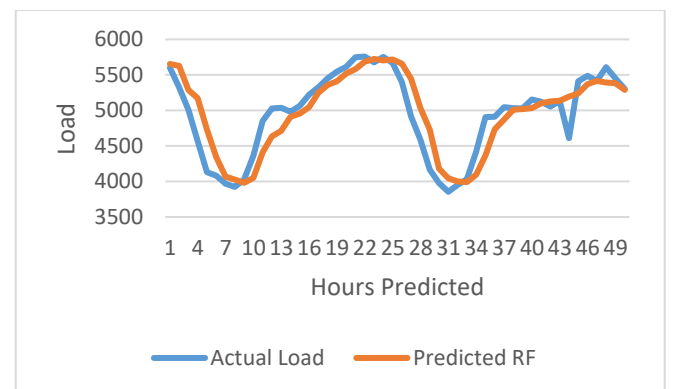

Fig. 7. RF predicted values

D. GBM

The result of using GBM for predicating the next 50 hours is shown in Figure 8. This model can predict future load

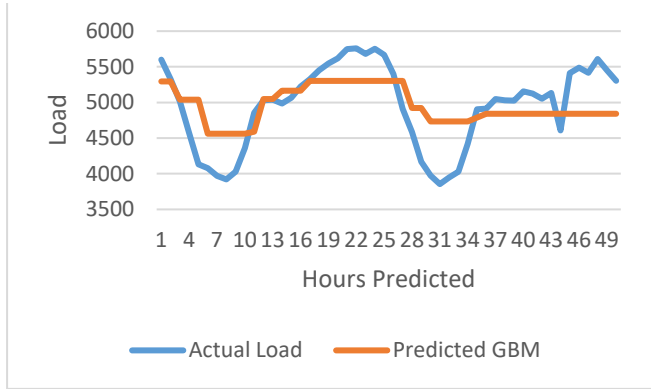values with MAPE of 0.06212, while its execution time is 1.6 seconds.

Fig. 8. GBM predicted values

E. Comparison Summary

To compute the average of MAPE and execution time for each algorithm, we executed the algorithms ten times and each time we compute MAPE and execution time for each algorithm as show Table 1 and Table 2 respectively.

Index	SVM	LLSVM	RF	GBM
1	0.14357	0.05152	0.04168	0.06212
2	0.14935	0.03915	0.03860	0.05649
3	0.15287	0.05044	0.04146	0.06065
4	0.14888	0.03612	0.03880	0.06475
5	0.14051	0.03548	0.02698	0.04399
6	0.16146	0.07164	0.03460	0.06295
7	0.14723	0.03673	0.03704	0.05566
8	0.15512	0.05228	0.03778	0.06006
9	0.15232	0.03843	0.03851	0.05442
10	0.15326	0.05422	0.04096	0.06068
Average	**0.15046**	**0.04660**	**0.03764**	**0.05818**

Table.1 MAPE for the four algorithms in each iteration

Index	SVM	LLSVM	RF	GBM
1	0.70	10.70	0.60	1.60
2	0.70	9.60	0.60	1.60
3	1.10	8.90	0.70	1.70
4	0.80	8.70	0.60	1.50
5	0.80	7.60	0.60	1.40
6	0.90	7.90	0.70	1.80
7	0.70	8.50	0.60	1.40
8	0.70	8.60	0.60	1.50
9	0.70	9.30	0.60	2.20
10	0.70	9.80	0.60	1.70
Average	**0.78**	**8.96**	**0.62**	**1.64**

Table.2 Execution time for the four algorithms in seconds

Figure 9 and Figure 10 shows the average MAPE and execution time for the ten iterations respectively. It is clear that RF has the least MAPE and hence it has the best accuracy. It also has the least execution time and hence it executes faster than other algorithms.
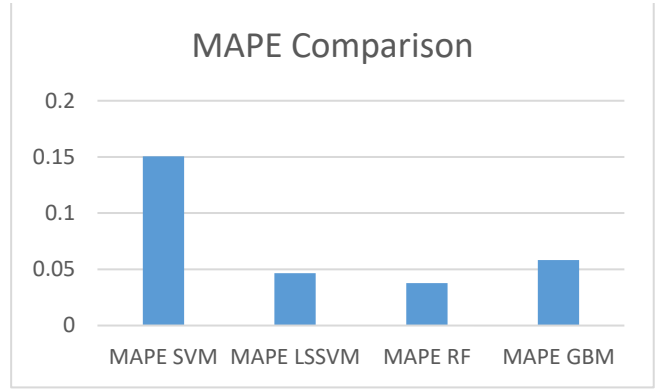

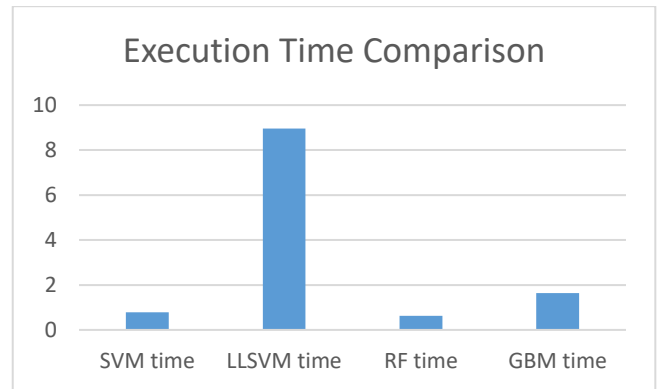Fig. 9. Average MAPE for the four algorithms

Fig. 10. Average of execution time for the four algorithms in seconds

IV. CONCLUSION AND FUTURE WORK

A reliable forecasting of electrical load demand could help to avoid problems resulted from unexpected loads, and give vital information to make decisions on energy generation and purchase, especially market-based dynamic pricing strategies. Furthermore, accurate prediction would have a significant impact on operation management. In this paper, we compare four machine learning algorithms (SVM, LSSVM, RF and GBM) for predicting the electrical load of Western United States.

We execute the algorithms ten times and each time we get different MAPE because of randomization of selecting data in training and testing phases and also different execution time. We calculate the average of MAPE for the ten iterations and found that SVM, LSSVM, RF and GBM has MAPE of 0.15046, 0.04660, 0.03764 and 0.05818 respectively. Also We calculated the average of execution time for the ten iterations and found that SVM, LSSVM, RF and GBM takes of 0.78, 8.96, 0.62 and 1.64 seconds respectively. From the values of the average of the ten iterations, its clears that RF algorithm has the least MAPE and the least execution time compared to the other three algorithms and hence RF outperforms the other three algorithms (SVM, LSSVM, and GBM) in terms of accuracy and execution time.

In the future we plan to compare more algorithms and also aim to combine two or more of these algorithms to get less MAPE and hence more accuracy. Furthermore, as we only have used one dataset to compare the algorithms, we plan to

use other electrical load datasets to confirm the results achieved in this paper.

REFERENCES.

[1] Anwar, Tahreem & Sharma, Bhaskar & Chakraborty, Koushik & Sirohia, Himanshu. (2018). Introduction to Load Forecasting. International Journal of Pure and Applied Mathematics. 119. 1527-1538.

[2] Ghulam Hafeez, Nadeem Javaid, Muhammad Riaz, Khalid Umar, Zafar Iqbal, and Ammar Ali, An Innovative Model Based on FCRBM for Load Forecasting in the Smart Grid,2020, Springer International Publishing

[3] Brownlee, Jason. "Time Series Forecasting as Supervised Learning." Machine Learning Mastery, 14 Aug. 2020, machinelearningmastery.com/time-series-forecasting-supervised-learning/.

[4] D. Felice, M., Yao, X.: Short-term load forecasting with neural network ensembles: a comparative study [application notes]. IEEE Comput. Intell. Mag. 6(3), 47–56 (2011).

[5] Ayub, Nasir & Javaid, Nadeem & Mujeeb, Sana & Zahid, Maheen & Khan, Wazir & Khattak, Muhammad. (2019). Electricity Load Forecasting in Smart Grids Using Support Vector Machine. 10.1007/978-3-030-15032-7_1.

[6] Ailing Yang, Weide Li, Xuan Yang,Short-term electricity load forecasting based on feature selection and Least Squares Support Vector Machines,Knowledge-Based Systems,Volume 163,2019,Pages 159-173,ISSN 0950-7051.

[7] Liu, Y., Wang, Y., and Zhang, J. (2012, September). New machine learning algorithm: Random forest. In International Conference on Information Computing and Applications (pp. 246-252). Springer, Berlin, Heidelberg.

[8] Richard S. Zemel and Toniann Pitassi, A Gradient-Based Boosting Algorithm for Regression Problems, NIPS'00: Proceedings of the 13th International Conference on Neural Information Processing SystemsJanuary 2000 Pages 675–681.

[9] https://www.pjm.com/. [Last Accesed: August 2021]

[10] Hahn, H., Meyer-Nieberg, S., Pickl, S.: Electric load forecasting methods: tools for decision making. Eur. J. Oper. Res. 199(3), 902–907 (2009) .

[11] Taylor, J.W.: An evaluation of methods for very short-term load forecasting using minute-by-minute British data. Int. J. Forecast. 24(4), 645–658 (2008) .

[12] D. Felice, M., Yao, X.: Short-term load forecasting with neural network ensembles: a comparative study [application notes]. IEEE Comput. Intell. Mag. 6(3), 47–56 (2011) .

[13] Pedregal, D.J., Trapero, J.R.: Mid-term hourly electricity forecasting based on a multi-rate approach. Energy Convers. Manag. 51(1), 105–111 (2010).

[14] Filik, ¨U.B., Gerek, ¨O.N., Kurban, M.: A novel modeling approach for hourly forecasting of long-term electric energy demand. Energy Convers. Manag. 52(1), 199– 211 (2011)

[15] R. J. Hyndman, "Forecasting: An Overview," Int. Encycl. Stat. Sci., pp. 536–539, 2011, doi: 10.1007/978-3-642-04898-2_256.

[16] Tayman, J., Swanson, D.A. On the validity of MAPE as a measure of population forecast accuracy. Population Research and Policy Review 18, 299–322 (1999).