ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 9, Issue 6, December 2019

Business Intelligence in Electrical Power Field

Medhat A. Rostum

***Abstract*— Electricity is one of the important requirements in our daily life. Electric load forecasting process plays an important role in forecasting future electric load demand and peak load by understanding the past. This allows meeting the predicted electric load demand using a least-cost plan. Several researchers proved that, an increase in forecasting error of electricity demand by 1% leads to an increase of 10 million pounds per year in operating costs. Thus accurate electric load forecasts are needed for power system security and reliability. It also improves energy efficiency and reliable operation of a power system. In opposite, inaccurate electric load forecasts may increase operating costs. This paper presents a state–of- the–Art system that can accurately forecast Electrical Load demand in future based on available collected data from the past.**

***Index Terms*— Business Intelligence, Load Forecasting, k-means algorithm.**

I. INTRODUCTION

Due to population growth, industry development and rising living requirements; Electricity consumption has increased. So power generating stations need to generate sufficient and reliable electricity to meet its consumption. To meet the electric load demand in the future, a load forecasting process should be used to predict what is the power generating stations need to generate for sufficient and reliable electricity in the future by analyzing the historical data. The results of the load forecasting process can be used in electricity generation such as generator maintenance scheduling, proper planning, purchasing of generating units and energy reservation, etc[1-2].

The load forecasting process can be classified into three categories: Short term, Medium term, and Long term load forecasting.

- Short term load forecasting (STLF) predicts the load demand from one hour to a few days, It is used hourly to forecast peak load and redundancy in production. Also it can be used in the daily operation of a power station such as unit start-up, scheduling of generation capacity, scheduling of fuel and coal purchases, etc. So with accurate STLF, electric utility would be able to operate in a reliable and secure manner [3, 6-8, 12-15, 20-25].
- Medium term load forecasting (MTLF) predicts the load demand from a month to one year, and it can be used in scheduling maintenance for devices and equipment [4].

Manuscript received: 17 November 2019
Manuscript received in revised form: 16 December 2019
Manuscript accepted: 01 January 2020
Manuscript Available online: 10 January 2020

- Long term load forecasting (LTLF) means forecasts for a period longer than a year up to ten years or up to several decades, and it is the first step in power station planning where any power station needs about five years from planning and designing to be constructed and to be ready for entering service[2,5].

This paper presents a study of short term load forecasting (STLF) method to forecast the future electric load demands. A variety of methods have been developed for short term load forecasting. These methods can be divided into:

- Conventional methods such as Regression model [6] and Time Series models including exponential smoothing methods [7-9], Box-Jenkins' ARIMA method [12].
- Artificial Neural Networks (ANNs) methods [13-16]. These methods are capable of learning non-linear relationships between variables.
- Support Vector Machine (SVM) method [20-22].
- Hybrid methods can be used to overcome the drawbacks of using an individual model and benefit from the advantages of each model to improve the prediction accuracy [23-25].

The following are some term definitions.

A. *Business Intelligence*

Business Intelligence (BI) is a set of processes and technologies for extracting useful information from large data sets, integrating data, storing it, analyzing it, and then reporting on it. It allows decision makers to make better business decisions [26-27].

B. *QlikView*

QlikView is a tool used for BI that simplifies the analysis of data for everyone. It enables users to retrieve data from multiple data sources. QlikView file isn't a database in itself, It contains its own data repository that is updated every time the source data are refreshed [28].

C. *R language*

R Language is a statistical data analysis tool for statistical computing and graphics that is made by the R project. It has excellent tools for graphics and data visualization [29-30].

II. SYSTEM TOOLS

In this paper, all codes are written in R Language integrated with QlikView. Once R Language and the necessary packages are installed, Visual Basic script is used to call R from QlikView application and performs the desired R functions on the selected QlikView data.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 9, Issue 6, December 2019

III. APPROACH

The aim of this study is to forecast the electric load demand, measure the accuracy for each fitted model and select the best one.

The lack of smart energy data in Egypt makes smart energy data more challenging. So the amount of electricity in KW of Ireland was monitored over time ' on a daily basis between 14/07/2009 and 31/12/2010' via more than 5600 meter IDs. So we need to extract, transform and load about 3 million rows of smart meter data information from relational database to QlikView.

The future electric load demand could be forecasted for each meter ID, but it requires an enormous amount of computation and time to obtain the final results. To overcome this problem, the first step after executing the extract, transform and load "ETL" process is classification of the data by grouping similar meter IDs into a few numbers of clusters. Afterwards, for each cluster a predictive model is created to forecast the electric load demand based on past values. Finally, the accuracy for each predictive model is evaluated by measuring the Forecast Error "$e_t = y_t - f_t$" between the actual value 'yt' and forecasted value 'ft ', and the best model is selected based on high accuracy. In this paper, the following approaches to build an appropriate model are used.

A. *Time Series Models*

Time series data is defined as a sequential collection of observations, taken at regular intervals of time. These intervals can be divided into short term intervals e.g.(hourly, daily, weekly) or long term intervals e.g.(monthly, quarterly or yearly).

The difference between Time Series models and other models is that, with other models the independent (response) variable is predicted from other dependent (predictor) variables. We assume that there is no autocorrelation, but Time Series models use the previous observations to predict future ones. So these models are used to predict future observations based on understanding the past [29].

Time Series Models include the following approaches:

- Exponential Smoothing Models [10].
    - Simple Exponential Smoothing Model
    - Holt-linear Exponential Smoothing Model
    - Holt-Winter Exponential Smoothing Model
- Autoregressive AR(p) model[11].
- Autoregressive Integrated Moving Average, ARIMA (p,d,q) Model[11].
- Seasonal Autoregressive Integrated Moving Average, ARIMA (p,d,q) (P,D,Q)s Model[11].

B. *Data Mining Models*

Data mining attempts to extract useful information from large data sets to discover meaningful and useful patterns. By analyzing such large data sets, such patterns enable to make business decisions [31]. It uses a combination of statistical methods, artificial intelligence and machine learning methods.

Data Mining Models include the following approaches:

- Multi Regression Model (mr)[29].
- Artificial Neural Networks (mlp)[13,16,29].
- Support Vector Machines (SVM)[17-19,29].
- Random forests for regression [29].
- Multivariate Adaptive Regression Splines (MARS) [29].

IV. PROPOSED BI SYSTEM ARCHITECTURE

- Execute the extract, transform and load "ETL "process.
- Classify the data by grouping similar objects into a small ' K ' number of clusters using K-means clustering.
- Create a predictive model using electric load forecasting approaches.
- Measure the accuracy for each fitted model and select best model according to high accuracy.
- Use the best model to forecast the electric load demand for each cluster in the future based on the previous data.

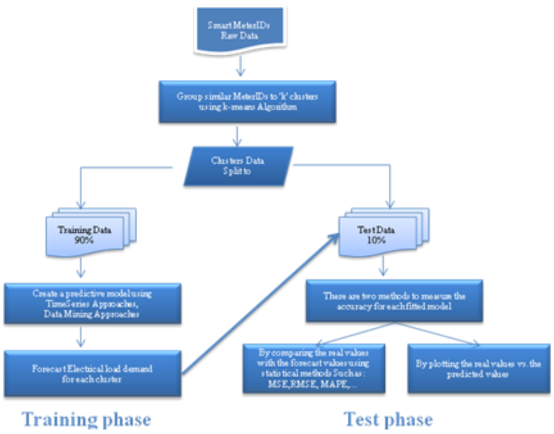

Fig. 1: BI System Architecture

V. ELECTRIC LOAD FORECASTING PROCESS

The following steps may be followed:

A. *Executing the extract, transform and load" ETL" process.*

- **Load Data**

We need to load about 3 million rows of smart meter data information from Relational databases to QlikView.

- **Transform Data**

After loading data, the electricity usage on many different time scales: Daily, Weekly, Monthly, Quarterly, Yearly, etc had to be measured. So we need to extract different time scales from the calendar information to be used in forecasting the energy consumption for the next day/month/year given

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 9, Issue 6, December 2019

the energy used in the previous N days/months/years.

B. Grouping similar Meter IDs into a few 'K' number of clusters.

- **The k-Means algorithm**

K-means Algorithm is one of the simplest and popular clustering algorithms. This algorithm is used to classify the data by grouping those objects into K number of groups based on attributes 'features'. the grouping is made by minimizing the sum of squares of distances between data and the centroid of the corresponding cluster[32].

○ Features

K-Means algorithm is used to group similar meter IDs into a few number of groups based on several features such as:

- The total energy over time in kilowatt(kw),
- The average energy over daily, weekly, monthly and yearly terms.
- The average energy over business days (BD) and weekend (WE).

○ Normalization

K-Means algorithm is a distance based technique, so all features need comparable ranges. This means that, they need to be normalized. We apply "min-max" normalization to the data to put it on the same scale, between [0,1] interval.

$$\text{Xnormalized} = \frac{X - \min(X)}{\max(X) - \min(X)}$$

○ Creating 'K' number of clusters

we must specify smart meters data "all input features", number of 'k' clusters and maximum number of iterations to group similar meter IDs into 'k' number of clusters.

○ Denormalization

After grouping similar meter IDs into 'k' number of clusters, we need to denormalize these features back into the original ranges.

Xdenormalized=

Xnormalized [max(X)-min(X)]+min(X).

C. Creating a predictive model to forecast the electric load demand for the selected cluster.

- **Training and Testing data sets.**

Selected cluster data are split to 90% training data and 10% testing data sets before modeling. The observations in the training data set are used to fit a proper model while the observations in the test data set are used to measure the accuracy of the fitted model by comparing these observations with the observations in the training data set.

- **Creating a proper model using Time Series models, Data Mining models.**

The observations in the training data set are used to fit a proper model. We hope to create an appropriate model that can be useful in the prediction of new cases.

D. Measuring the accuracy for each fitted model and selecting best model.

This paper has several types of models to deal with the prediction of the response; The regression relationship that has been established on a training data set can be useful in the prediction of new cases. To measure the accuracy for each fitted model, there are two steps:

- Plot both actual observations and forecast values.
- Measure the prediction on a new test data set using some statistical methods such as: SSE, MSE, RMSE, ME, MAE, MPE, MAPE [33-34].

In each of the following definitions, 'yt' is the actual value, 'ft' is the forecasted value and 'n' is the size of the test data set.

1. Sum of squared error (SSE): is measured as sum of square errors,
$$SSE = \sum (yt - ft)^2$$
2. Mean of squared error (MSE): is measured as an average of square errors,
$$MSE = SSE/n.$$
3. Root mean square error (RMSE): is estimated as root of calculated MSE,
$$RMSE = \sqrt{MSE}$$
4. Mean error(ME): is defined as average error values of observed minus forecast,
$$ME = \frac{\sum (yt - ft)}{n}$$

For a good forecast, The mean error should be close to zero; as it indicates a bias in the forecasts. The drawback of this method is that the positive and negative error values can cancel each other.

5. Mean absolute error (MAE): is defined as the average of absolute error values of observed minus forecasted data,
$$MAE = \frac{\sum |yt - ft|}{n}$$

For a good forecast, the obtained MAE should be as small as possible.

6. Mean percentage error (MPE): is represented as the average of the PE values.
$$MPE = \frac{\sum \frac{yt - ft}{yt}}{n} * 100$$
7. Mean absolute percentage error (MAPE): It expresses the forecast error in percentage terms.
$$MAPE = \frac{\sum \left| \frac{yt - ft}{yt} \right|}{n} * 100$$

Both, Root mean square error (RMSE) and Mean absolute percentage error (MAPE) are the most popular measures in load forecasting [34].

VI. EXPERIMENTS AND RESULTS

There are seven Time Series models namely:(AR(p), ARIMA(p,d,q), SARIMA(p,d,q)(P,D,Q)s, Simple Exponential, Holt-Linear Exponential Smoothing,

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 9, Issue 6, December 2019

Holt-Winter Exponential Smoothing, ETS model) and five Data Mining models(mr, mlp, svm, mars, random forest) to compare.

We can classify types of energy users as a household or a business according to their average daily electricity consumption in KWH. In general, the low average daily electricity consumption indicates a household rather than a business. In opposite, the bigger average daily electricity consumption indicates a business rather than household. The amount of average daily electricity consumption indicates the size of the business.

Several clusters such as clusters (1, 2, 14, 15, 17, 18, 19, 21, 26,27) with bigger average daily electricity consumption indicate a business rather than a household. The pattern of these clusters are similar.

Since we wanted to work on each cluster separately, So cluster (18) is taken into account for future investigation as an example for these types of clusters.

A. For cluster (18), Measure the prediction on a new test data set using some statistical methods.

Table 1. For cluster (18), the accuracy for each fitted Time Series model

Time Series Models	SSE	MSE	RMSE	MAPE
AR(p)	14347.69	256.69	16.3	9.355
ARIMA	33347.28	617.54	24.85	16.18
SARIMA	18128.44	335.711	18.32	10.32
Simple EXP.	34049.7	630.55	25.11	16.51
Holt-Linear	29135.48	539.54	23.22	15.39
Holt-Winter	19227.47	356.06	18.86	10.97
ETS	15011.03	277.98	16.67	9.79

Table 2. For cluster (18), the accuracy for each fitted Data mining model

Data Mining Models	SSE	MSE	RMSE	MAPE
mr	4766.39	93.45	9.66	5.4
SVM	3689.13	72.33	8.5	4.62
mlp	4776.81	93.66	9.67	5.42
Randomf.	6331.15	124.14	11.14	6.46
mars	3694.88	72.44	8.51	4.74

From table1, we found that AR (p) model has RMSE=16.3 and MAPE=9.3 but SVM model in Table2 has RMSE=8.5 and MAPE=4.6.

Whenever a decrease in the forecast error, means that more accurately you are able to predict what happens to one variable based on the knowledge you have of the other variables. Then by comparing the Experimental Results for Data Mining models in Table2 vs. the Experimental Results for Time Series models in Table1, it is noted that the forecast error for Data Mining models is low thence the accuracy for Data Mining models is very high and better than the accuracy for Time Series models.

B. For cluster (18), Plot both actual observations and forecast values.

Also if both actual observations and forecast values for cluster (18) are plotted, we will obtain the plot of the predicted values (red) coming from Time Series models vs. the real values (black) in the test data set as in figures (2:5). Also the plot of the predicted values(red) coming from Data Mining models vs. the real values(black) in the test data set are shown in figures(6:10).

In Time Series models, we have 536 smart energy data rows from " 14/07/2009 to 31/12/2010" for each cluster. We split this data to 90% training data and 10% testing data sets. So we have 482 rows used for data training from " 14/07/2009 to 07/11/2010" and 54 rows used for data testing from " 08/11/2010 to 31/12/2010".

In Data Mining models, "CasesSeries" function is used to shift and lag the data column, pulling up the previous 24 records into the current record. So we have 512 smart energy data rows from " 07/08/2009 to 31/12/2010" for each cluster. By splitting this data to 90% training data and 10% testing data sets, There will be 461 rows used for data training from" 07/08/2009 to 10/11/2010 " and 51 rows used for data testing from " 11/11/2010 to 31/12/2010".

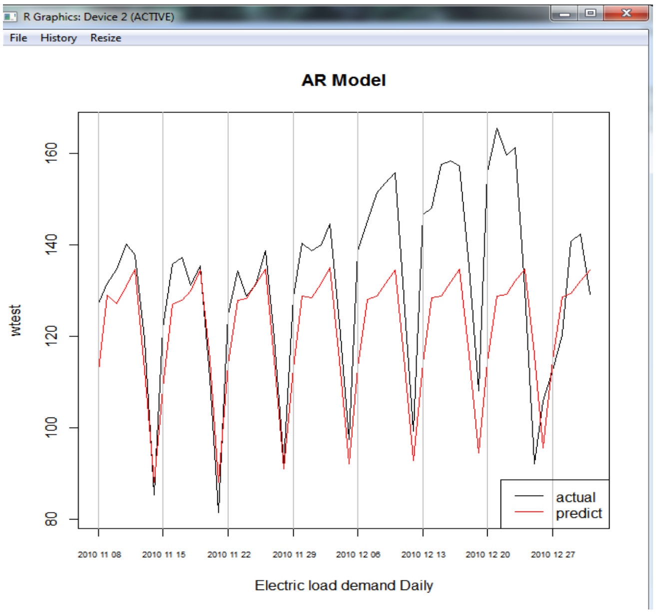

Fig.2: Plot of the predicted values (red) coming from AR model vs. the real values (black) for cluster 18.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 9, Issue 6, December 2019

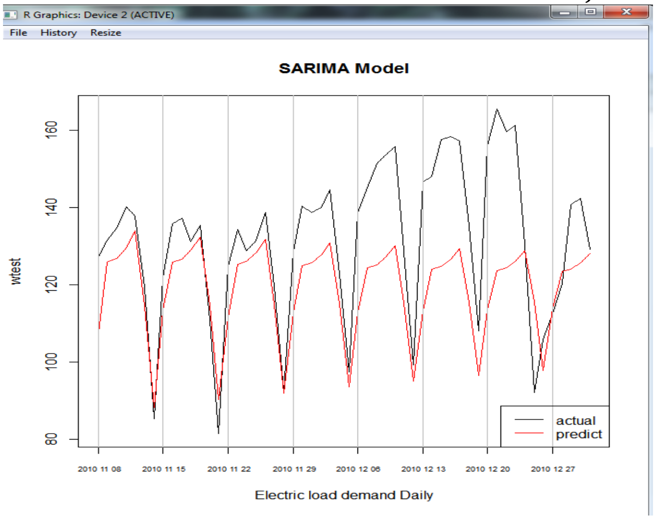

Fig.3: Plot of the predicted values (red) coming from Seasonal ARIMA model vs. the real values (black) for cluster 18.

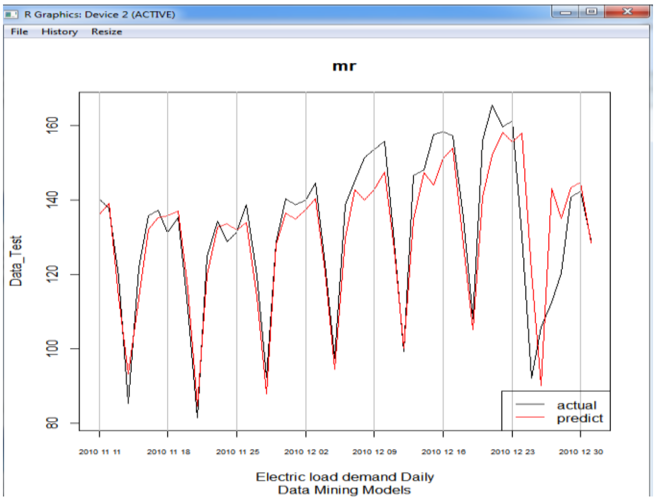

Fig.6: Plot of the predicted values(red) coming from multiple regression (mr) model vs. the real values(black) for cluster 18.

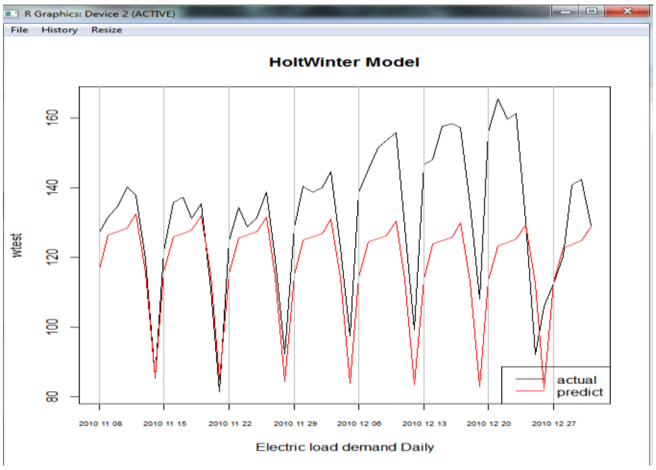

Fig.4: Plot of the predicted values(red) coming from HoltWinter model vs. the real values(black) for cluster 18.

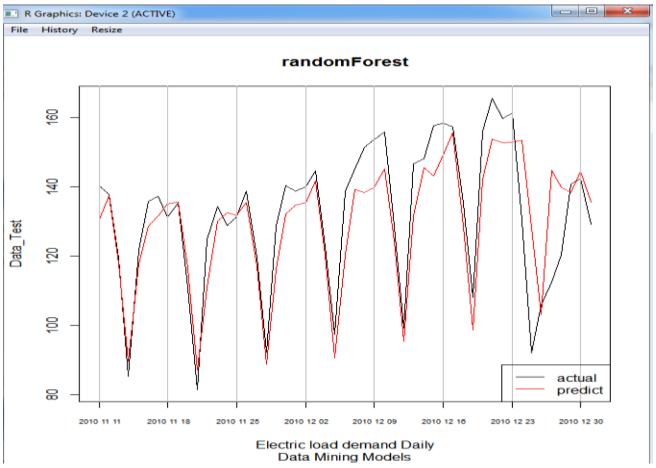

Fig.7: Plot of the predicted values(red) coming from random forest model vs. the real values(black) for cluster 18.

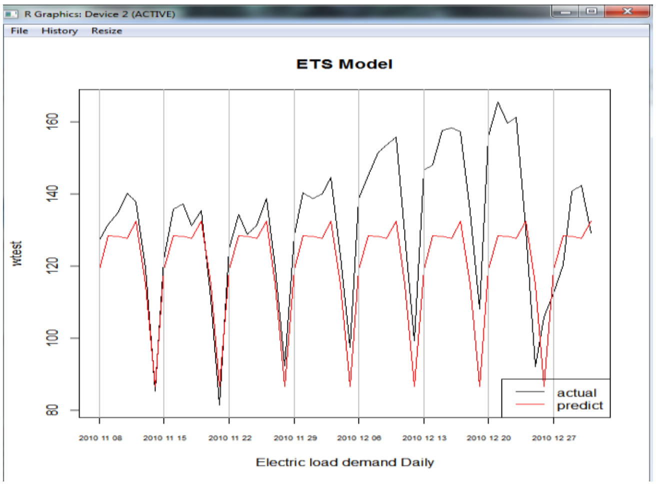

Fig.5: Plot of the predicted values(red) coming from ETS model vs. the real values(black) for cluster 18.

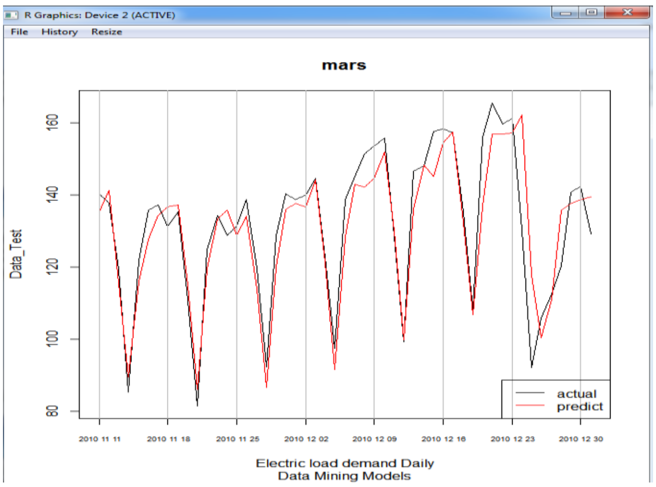

Fig.8: Plot of the predicted values (red) coming from mars model vs. the real values(black) for cluster 18.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 9, Issue 6, December 2019

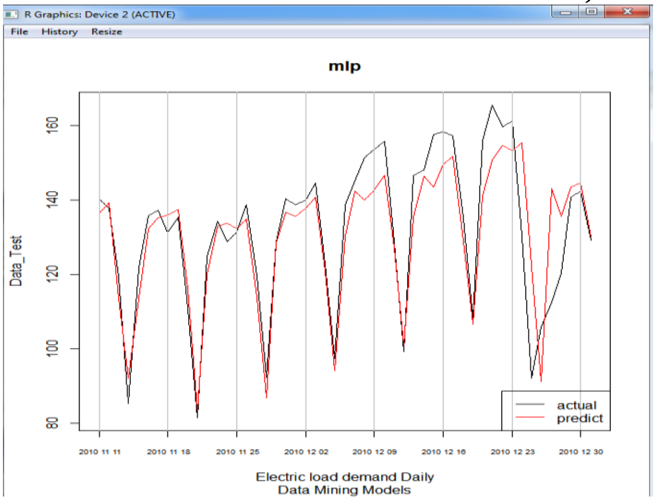

Fig.9: Plot of the predicted values (red) coming from multilayer perceptrons (mlp) model vs. the real values (black) for cluster 18.

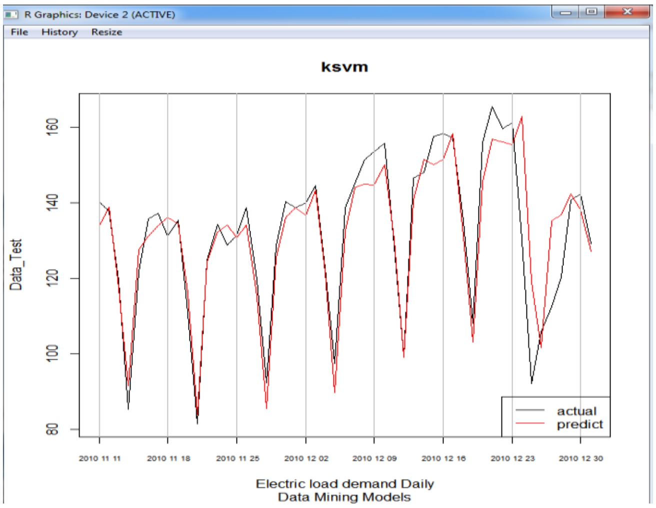

Fig.10: Plot of the predicted values (red) coming from Support Vector Machine (svm) model vs. the real values (black) for cluster 18.

From the plot of the predicted values(red) coming from Data Mining model(using SVM approach) vs. the real values(black) for cluster(18), as shown in fig.10, it can be said that the prediction seems already pretty good as the forecasted values are following the actual values in its trends.

Then from the two steps above(Statistical methods and Plot method) for measuring the accuracy of each fitted model, Data Mining model(using SVM approach) produces predicted values with low forecast error 'Residuals' and high accuracy. The high accuracy of SVM approach depends on selecting the best hyper-parameters (σ, ε and C) to improve the performance of the support vector regression (SVR) and this is extremely crucial for successful forecasting.

Also, Several clusters such as clusters(3, 5, 6, 7, 8, 9, 10, 12,13,16,20,22,24,25,29) with low average daily electricity consumption indicate a household rather than a business. The pattern of these clusters are similar. Cluster (5) is taken into account for future investigation as an example for these types of clusters.

C. For cluster (5), Measure the prediction on a new test data set using some statistical methods.

Table 3. For cluster (5), the accuracy for each fitted Time Series model

Time Series Models	SSE	MSE	RMSE	MAPE
AR(p)	2244	41.5	6.4	10.6
ARIMA	2026	37.5	6.1	10
SARIMA	1839	3.4	5.8	9.2
Simple EXP.	1997	36.9	6	10.6
Holt-Winter	1573	29.1	5.3	8.5
ETS	2154	39.8	6.3	10.53

Table 4. For cluster (5), the accuracy for each fitted Data mining model

Data Mining Models	SSE	MSE	RMSE	MAPE
mr	316.4	6.2	2.4	3.8
SVM	381	7.4	2.7	3.97
mlp	336.9	6.6	2.5	3.84
Randomf.	511.2	10	3.1	4.97
mars	401.2	7.8	2.8	4.65

From table 3, it is found that Holt-Winter model has RMSE=5.3 and MAPE=8.5 but mr model in Table4 has RMSE=2.4 and MAPE=3.8.

D. For cluster (5), Plot both actual observations and forecast values.

If both actual observations and forecast values for cluster(5) are plotted, we will obtain the plot of the predicted values(red) coming from Time Series models vs. the real values(black) in the test data set as in figures(11:14). Also the plot of the predicted values(red) coming from Data Mining models vs. the real values(black) in the test data set are displayed in figures(15:19).

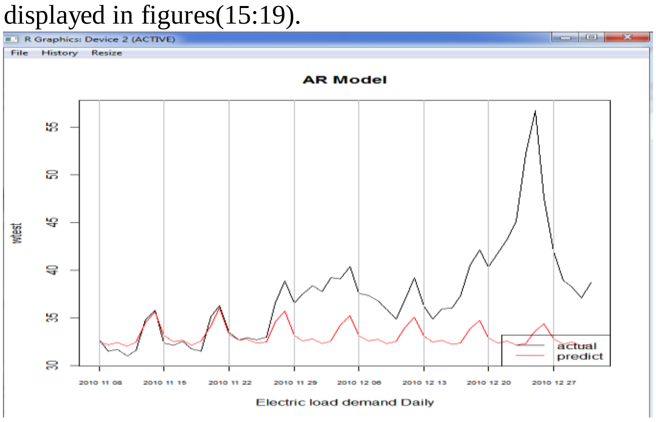

Fig.11: Plot of the predicted values (red) coming from AR model vs. the real values (black) for cluster 5.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 9, Issue 6, December 2019

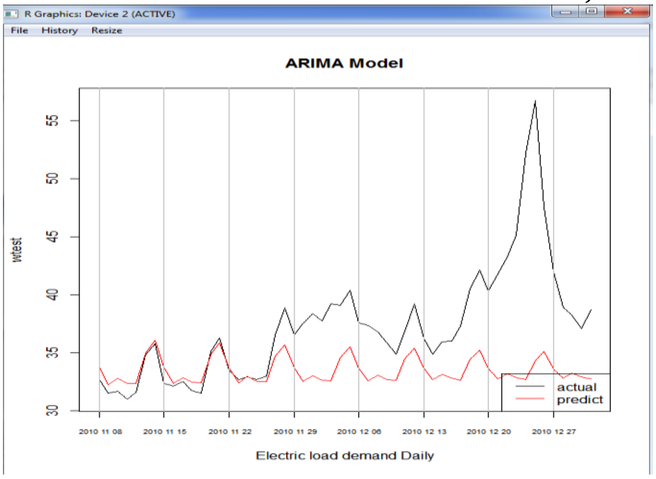

Fig.12: Plot of the predicted values (red) coming from ARIMA model vs. the real values (black) for cluster 5.

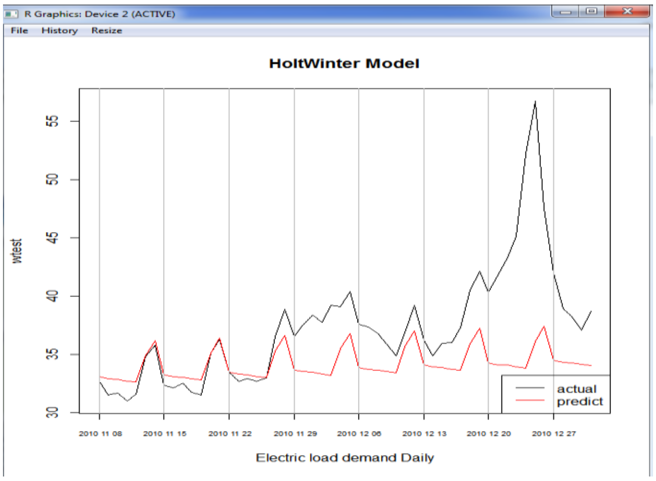

Fig.13: Plot of the predicted values(red) coming from HoltWinter model vs. the real values(black) for cluster 5.

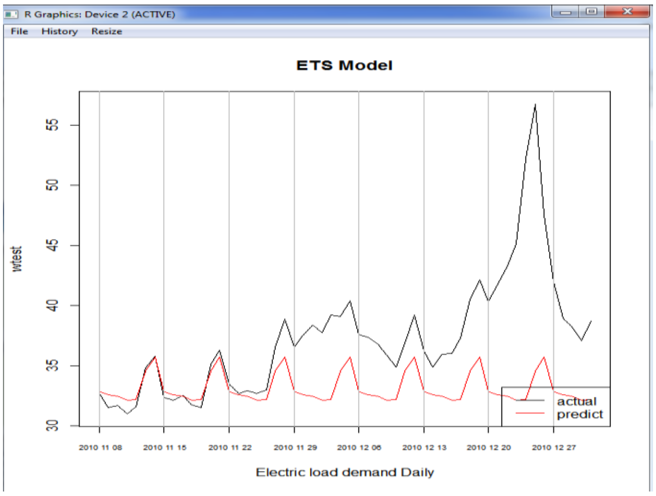

Fig.14: Plot of the predicted values (red) coming from ETS model vs. the real values (black) for cluster 5.

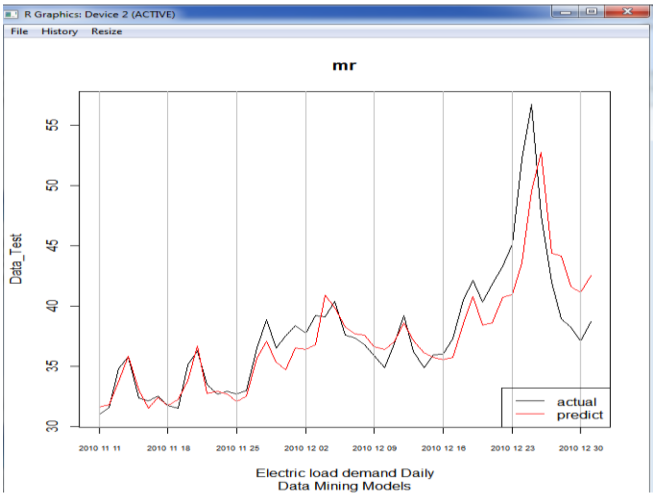

Fig.15: Plot of the predicted values (red) coming from multiple regression (mr) model vs. the real values(black) for cluster 5.

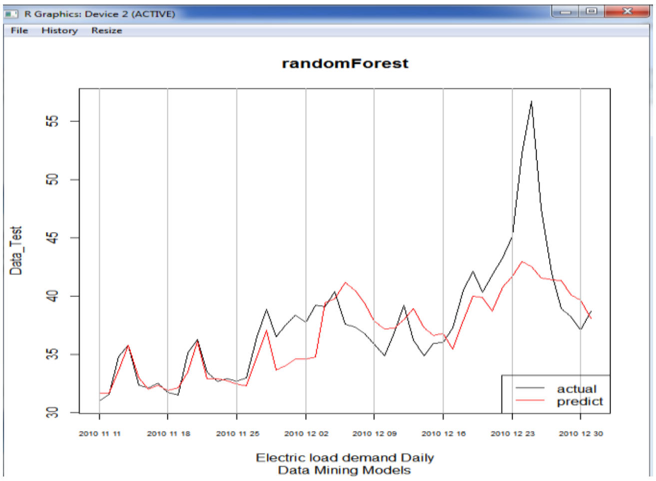

Fig.16: Plot of the predicted values(red) coming from random forest model vs. the real values(black) for cluster 5.

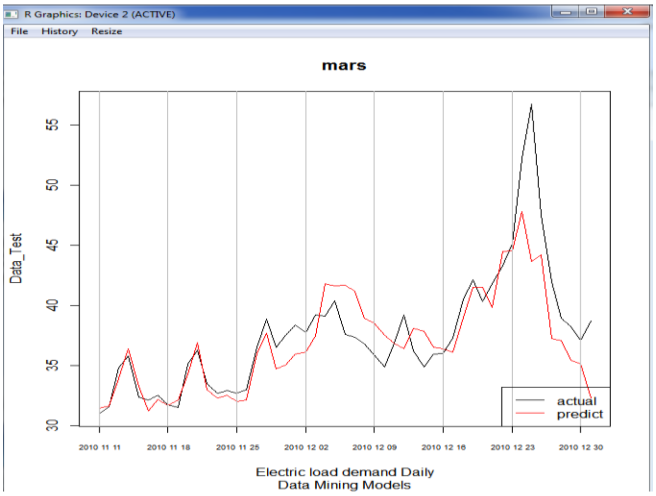

Fig.17:Plot of the predicted values(red) coming from mars model vs. the real values(black) for cluster 5.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 9, Issue 6, December 2019

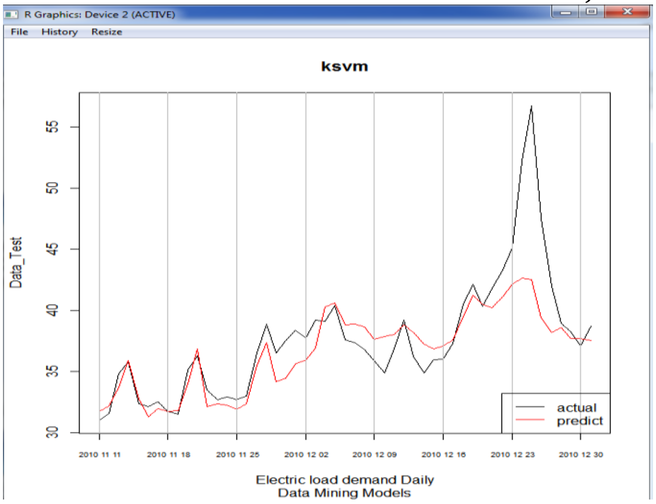

Fig.18: Plot of the predicted values (red) coming from SVM model vs. the real values (black) for cluster 5.

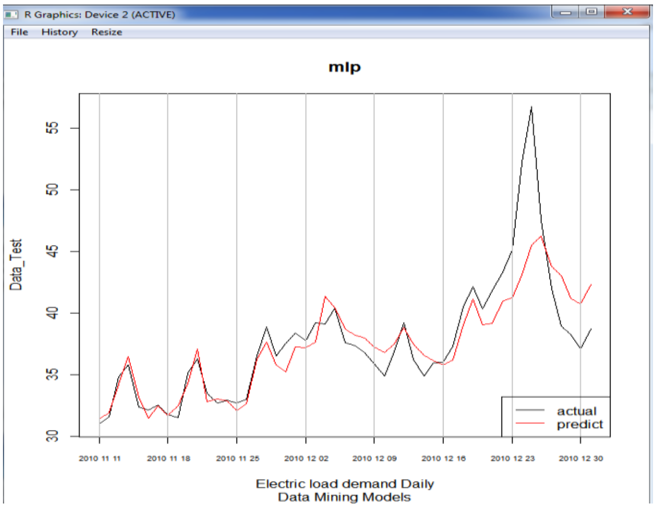

Fig.19: Plot of the predicted values (red) coming from MLP model vs. the real values (black) for cluster 5.

By comparing the results in table3 with those in table4, for cluster(5) and from plots of the predicted values(red) vs. the real values(black), as shown in figures(11:19), it can be concluded that the predicted values coming from Data Mining models(using mr approaches) are more accurate than the predicted values coming from Time Series models.

The key to the implementation of Data Mining models was given by "CasesSeries" Function, This function allows shifting and lagging of a selected data column, pulling up the previous records into the current record.

VII. CONCLUSION AND FUTURE WORK

From the Experimental Results, it can be concluded that, for a good forecast one needs really good parameter estimates. The proper selection of the model orders in ARIMA models is very important for good forecasting. Choosing suitable number of inputs and hidden neurons is very important in constructing an appropriate neural network model (mlp) to produce small error in both training and test data. Choosing the best hyper-parameters in svm is very important for successful forecasting.

For future work, to improve forecast accuracy we can use Hybrid methods to overcome the drawback of using individual models and benefit from the advantages of each model.

REFERENCES

[1] Milos STOJANOVIĆ, Milos BOŽIĆ, Zoran STAJIĆ, Marko MILOŠEVIĆ," LS-SVM model for electrical load prediction based on incremental training set update", PRZEGLĄD ELEKTROTECHNICZNY, ISSN 0033-2097, R. 89 NR 4/2013, 194-198.

[2] Rob J. Hyndman and Shu Fan," Density Forecasting for Long-Term Peak Electricity Demand", IEEE TRANSACTIONS ON POWER SYSTEMS, 2009, pages: 1-12.

[3] G. Gross, F. D. Galiana, 'Short-term load forecasting', Proceedings of the IEEE, 1987, 75(12), 1558 – 1571.

[4] Pituk Bunnoon," Mid-Term Load Forecasting Based on Neural Network Algorithm: a Comparison of Models", International Journal of Computer and Electrical Engineering, Vol. 3, No. 4, August 2011.

[5] Wagdy MANSOUR, Mohamed MOENES, Hassan MAHMOUD, Ahmed GHAREEB," LONG TERM LOAD FORECASTING FOR THE EGYPTIAN NETWORK USING ANN AND REGRESSION MODELS",21st International Conference on Electricity Distribution Frankfurt, 6-9 June 2011.

[6] Papalexopoulos A. D., Hesterberg T. C., A regression-based approach to short-term load forecasting, IEEE Transactions on Power System, 5 (1990), No. 4 1535–1550.

[7] Christianse, W.R. Short term load forecasting using general exponential smoothing. IEEE Trans.Power Apparatus Syst. 1971, 90, 900–911.

[8] Park, J.H.; Park, Y.M.; Lee, K.Y. Composite modeling for adaptive short-term load forecasting. IEEE Trans. Power Syst. 1991, 6, 450–457.

[9] J. W. Taylor and P. E. McSharry, "Short-Term Load Forecasting Methods: An Evaluation Based on European Data", IEEE Transactions on Power Systems, vol. 22, pp. 2213-2219, 2007.

[10] Rob J Hyndman, "Forecasting based on state space models for exponential smoothing", 29 August 2002.

[11] G.E.P. Box, G. Jenkins, "Time Series Analysis, Forecasting and Control", Holden-Day, San Francisco, CA, 1970.

[12] Huang S. J., Shih K. R., Short-term load forecasting via ARMA model identification including non Gaussian process considerations, IEEE Transactions on Power System, 18(2003), No. 2, 673–679.

[13] D.C. Park, M.A. El-Sharkawi, R.J. Marks II, L.E. Atlas & M.J. Damborg, "Electric load forecasting using an artificial neural network", IEEE Transactions on Power Engineering, vol.6, pp.442-449 (1991).

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 9, Issue 6, December 2019

[14] Lee, K.Y., Cha, Y.T., and Park, J.H. 1992. "Short Term Load Forecasting using an Artificial Neural Network". IEEE Transactions on Power Systems. 7(1):124-132.

[15] S.J. Kiartzis, A.G. Bakirtzis, V. Pertridis, "short-load forecasting using neural networks", Electric Power Systems Research 33(1995) 1-6.

[16] Markos Markou, Elias Kyriakides and Marios Polycarpou," 24-Hour Ahead Short Term Load Forecasting Using Multiple MLP ", pages: 1-6.

[17] Cortes, Corinna , and Vladamir V apnik . "Support Vector Networks," Machine Learning20 (1995): 273–297.

[18] V. Vapnik, Statistical Learning Theory, (book) Wiley, NY, 1998.

[19] Alexandros Karatzoglou, David Meyer, Kurt Hornik," Support Vector Machines in R", Journal of Statistical Software, April 2006, Volume 15, Issue 9.

[20] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," Advances in Neural Information Processing Systems, vol. 9, no. October, pp. 155–161, 1997.

[21] B.-J. Chen, M.-W. Chang and C.-J. Lin, "Load Forecasting Using Support Vector Machines: A Study on EUNITE Competition 2001," IEEE Transactions on Power Systems, vol. 19, no. 4, pp. 1821–1830, 2004.

[22] Hong WC (2009) Electric load forecasting by support vector model. Appl Math Model 33:2444–2454. doi:10.1016/j.apm.2008.07.010.

[23] A.A. Desouky, M.M. Elkateb, 'Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA', IEE Proceedings of Generation, Transmission and Distribution, 2000, 147(4), 213 - 217.

[24] Hong WC (2009) Hybrid evolutionary algorithms in a SVR-based electric load forecasting model. Int J Electr Power Energy Syst 31:409–417. doi:10.1016/j.ijepes.2009.03.020.

[25] Hongzhan NIE, Guohui LIU, Xiaoman LIU, Yong WANG, "Hybrid of ARIMA and SVMs for Short-Term Load Forecasting", Energy Procedia 16 (2012) 1455 – 1460.

[26] Greg Nelson," Paper BI 001Introduction to the SAS® 9 Business Intelligence Platform: ATutorial", pages: 1-12.

[27] Danny Stoltenberg Stjerne," INTRODUCTION TO BUSINESS INTELLIGENCE ", pages: 1-18.

[28] Miguel García, Barry Harmsen," QlikView 11 for Developers",(book) Copyright © 2012 Packt Publishing.

[29] Adler, J.: R In a Nutshell: A Desktop Quick Reference. Sebastopol, CA: O'Reilly Media, 2009.

[30] Gentleman, R., Ihaka, R., 1996. R: A language for data analysis and graphics. The Journal of Computational and Graphical Statistics 5 (2), 491-508.

[31] Johannes Ledolter," DATA MINING AND BUSINESS ANALYTICS WITH R" Published by John Wiley & Sons, Inc.

[32] Teknomo, Kardi. K-Means Clustering Tutorials. http:\\people.revoledu.com\kardi\tutorial\kMean\Last Update: July 2007.

[33] Marin Matijaš," ELECTRIC LOAD FORECASTING USING MULTIVARIATE META-LEARNING ", Zagreb, 2013.