International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 3 Issue: 8
ISSN: 2454-4248
10 – 13

Classification of Data Mining Techniques for Electricity Load Forecasting

Rajnikant Pandya
Student - Research Scholar (Computer Science)
Gujarat Vidyapith-Ahmadabad
Gujarat, India
pansweet@rediffmail.com

Dr. Ajay Parikh
Professor & Head, Department of Computer Science.
Gujarat Vidyapith-Ahmedabad
Gujarat, India
hodcs@gujaratvidyapith.org

***Abstract*—** The global requirement for electricity is increasing daily with the expansion of infrastructure and the addition of new appliances. Precise power consumption forecasting is of major importance for government to define the future power requirement of a given region. This is one of the most scientifically and technologically challenging problem.

In this paper, we present the review of available literatures of some data mining techniques and algorithms used by other researchers for short term and long term electric load forecasting. Data mining operations are performed on historical and secondary data to develop a prediction model which is capable of predicting electricity consumption. The work that has been done by various researchers in this field has been reviewed. The review has outlined a number of promising techniques like ARIMA, MLR, ANN, FLN, SVM, Gradient boosted tree, K-NN, Holt-winter, Very Fast Decision Tree, Fuzzy logic and Genetic algorithms etc that have been used to understand the relationships of various factors on electric power prediction.

Keywords- *Data Mining, Short term load forecasting, Long term load forecasting, Prediction, Accuracy*

*****

I. INTRODUCTION

Right now we are fighting with the problems of insufficient power supply because of heavy usages and inadequate production and planning of power requirement all over the world. In this context governments are trying a lot to accomplish this challenge. If company knows in advance, how much power is needed in particular areas then it is beneficial for distribution and production. Electricity load Forecasts are the elementary factors considered by utilities while planning power generation, capacity building via power infrastructural development. Hence, rationalizing the importance bestowed to this area of research since the last few years. Accurate load forecasting becomes more complex, yet more important for management of electricity requirement.

II. BACKGROUND STUDY OF ELECTRIC POWER PREDICTION.

A review of the literature shows that there are numerous studies for load forecasting. Electric load forecasting can be divided in to two categories- short term load forecasting and long term load forecasting. Short term load forecasting predict the load demand for one hour to several day and long term load forecasting predict the load demand for one month to year.

The various techniques and algorithms are used like Classification of time series, ANN(Artificial neural network), Cross Validation Technique- 2-fold,k-fold,Simple statistics based algorithm, Machine learning algorithm (Naive bayesian, K-Nearest Neighbor, Support vector),Time series based predictor, Fuzzy Neural Network Technology (FNN), FLN(Functional-Link Network),Hybrid Neural Model, Very fast decision tree (VFDT), CRISP-DM Model, various form of ARIMA(Auto Regressive Integrated Moving Average), MLR(Multiple Linear Regressions) etc.

III. ELECTRICITY LOAD FORECASTING

There are many studies that support the applicability of data mining techniques for electricity load forecasting. There are two categories for electricity load forecasting like

A. Short term load forecasting

B. Long term load forecasting

Other than these there are many mining techniques are used for various types of forecasting.[7][12][13][14] [16] uses fuzzy set methodology, genetic algorithm and ANN for commercial load forecasting, to reduce the predicting average error, estimate and predict electricity demand, prediction of electric power requirements to meet with actual demand and prediction of electric energy consumption. [10][11] uses supervised and unsupervised learning, intrinsic and extrinsic cluster analysis methodologies for electricity consumption analysis using meter reading data, daily electricity load and prediction of electric power requirement based on weather conditions.[9] uses theories of incremental learning and prepare streaming process engine for smart power grid system

 ISSN: 2454-4248
10 – 13

created in JAVA for electricity demand forecasting for next day.

A. Short term load forecasting

V. Uher, R. Burget, M.K. Dutta and P. Mlynek et al [2] carried out study during the period 1-1-2011 to 31-12-2014 has surfaced out a research issue of precise prediction of electric power requirement to match with dynamically growing consumption need day by day. These researchers describe the method based on input data for power consumption only. No other influences means additional parameters were included.

To meet with actual consumption, they used five machine learning methodologies, Local polynomial regression, Neural net, Gaussian process, Linear regression and Polynomial regression and could achieve RMSE(Root Mean Square Error) 0.632, 0.711, 0.716, 1.083 and 1.305 respectively and RMSE(%) 5.77, 6.49, 6.54,9.89 and 11.92 % respectively.

M.M.Ismail and M.A.M. Hassan et al [4] Carried out study of short-term load forecasting issues and predict electric power requirement to meet with real consumption and reduce the RMSE using different model. The data are used for this research are daily peak load in MW. The data collected from north Cairo of Electricity Distribution Company for the year 2008 and 2009.

They used different algorithms for this research like Time series analysis, ANN, ANFIS, Stochastic, ARIMA, REGARIMA, Multiple regression etc. They have used multiple regression model for (t power 5) and (t power 12) and achieved the RMSE 150.24 and 130.36. They also used ARIMA(1,0,1) and ARIMA(0,1,1) and achieved the RMSE 70.37 and 32.29 respectively.

Then they create a new model of combining ARIMA and multiple regressions called REGARIMA methodology and it compared with ANFIS and ANN methodologies and could achieve the result RMSE 32.29, 45.58 and 46.96 MAE 10.65, 25.73 and 27.03 and MPE 1.22, 1.74, 1.689 respectively.

B. koo and his team et al [5] carried out the study using the data for the period of 2007 to 2008 for precise prediction of short term electric load forecasting with an objective to reduce MAPE. The data to classify is daily load data and this data is collected in 1- hour interval of the total electric generation of Korea. They have suggest to improve the accuracy for forecasting, to take additional data such as weather , temperature , event of the day for implementing these used methodology.

They used K-NN methodology for classification and Holt-winter and Seasonal ARIMA (SARIAM [Seasonal Auto Regressive Integrated Moving Average]) for forecasting and could achieved total MAPE 3.645615 and 3.192888 respectively.

V.H. Kher and S.K.Joshi et al [15] have taken the issues of short term load forecasting for precise prediction of electricity requirement to meet with actual consumption. They collect the load data of GEB, Gujarat for two month from dispatch center of Gujarat Electricity Board. Weather parameters are not included. In the prediction, they include only hour of the day and previous hour data and forecast electric power requirement of one hour ahead.

They used ANN methodology and develop new model called FLN [Functional-Link Network] for forecasting electric power requirement of one hour ahead. The result of FLN with learning rate=0.6 and momentum=0.3 of a period of one day. The percent error [% Error= {(A-P)/A}*100] shows the accuracy of forecasting. The percent errors vary between -1.1630 to 1.2766.

B. Long term load forecasting

H.K. Mohamed and his team et al [8] have taken the issues of long term electricity load forecasting to meet the actual consumption with an objective to reduce MAPE1 [Mean Absolute Percentage Error] and MAPE2 [Maximum Absolute Percentage Error]. The data collected from different sites of Egyptian electricity sector [20, 21, and 22]. The parameters include in this study are year, month, total power generation, total consumption, load factor, losses, population, GDP, price, temperature, humidity and peak load. All of these parameters are collected in excel spreadsheet. To enhance the result, they have suggested that, used other factors such as land use and season to see their effects in this methodologies.

They used NCBI [Natural Cluster Based Interpolate] methodology for detecting missing values, Histogram methodology for detecting odd values. For prediction, they used Regression and Neural network methodologies and could achieve MAPE1 and MAPE2 for different 10 cases. The lowest MAPE1 and MAPE2 out of these are 1.746% and 4.107% respectively.

C.P. Ronald Reagan and S.R. Sari et al [17] have taken the issues of precise prediction of long term electric power requirement for Tamil Nadu to meet with actual demand with an objective to minimize/reduce MAPE. They collect data of Tamil Nadu state for the period of 1983 to 2013. The parameters include in dataset are total electricity consumption, population, Gross state domestic product, per capita income and yearly peak demand. They Develop fuzzy-neural approach model and make the comparison with some approaches, MLRM [Multiple Linear Regression Method], GA-MLRM [Genetic Algorithm- MLRM], SA- MLRM [Simulated Annealing- MLRM] and could achieve MAPE(%) average for year 2006 to 2012 are 2.00%, 1.94% and 1.86% respectively.

 ISSN: 2454-4248
10 – 13

TABLE I: Summary of electricity load forecasting techniques

Techniques	Advantages	Limitations
Multiple Linear Regression	• Determines the relation influence of one or more predictors' variable to the criterion values. • Provides better prediction from multiple predictors. • It avoid depending on a single predictor.	• Only looks at linear relationship between dependent and independent variables. • Assumes that data are independent.
Artificial Neural Network	• Few traditional statistical training required. • Capable to detect complex nonlinear relationships between dependent and independent variables • Capacity to detect all feasible interaction among predictors.	• Its "black box" nature • Heavy computational load. • Experimental nature of model development .
ARIMA	• Its theory is well constructed and unrevealed. • Steady prediction of time-varying trends and seasonal patterns. • Less parameter required. • Sufficient no of parameters regressed and averaged.	• Implicit seasonal indications only. • Complex to interpret coefficients • Risk of over fitting or discriminate.

IV. REVIEW ANALYSIS

B.koo, V.Kher and I. abdual and his team[5][15][6] have taken the issues of predicting electric demand for short term load forecasting in the year 2012 [5] and 2016[5][6] and could achieve MAPE 3.645615using holt-winter,3.19288 using SARIMA[5], 1.11833% and 1.07158% using SARIMA[6], and -1.1630 to1.2766 using FLN[15].

H.K. Mohamed, K.kandananond, H.owda and C.P. Ronald and his team [8][13][16][17] have taken the issues of predicting electric demand for long term load, in the year 2006 researcher[8] could achieve MAPE 1.746% using regression and NN, in 2011[13] achieve 2.80981% using ARIMA, 3.2604527% using MLR and 0.996% using ANN, in 2014[16][17] could achieve 2% using MLRM, 1.94%using GA-MLRM and 1.86% using SA-MLRM.

V. CONCLUSION AND SUGGESTIONS

The present research provides the knowledge of what an existing system of electricity load forecasting and how many ways we can implement it. The issues of accurate prediction and rate of error in previous research are not conclusive and are still required enhancement and improvement. For this, it is possible to adopt alternative approach with various parameters to address these issues. In the future, accuracy of prediction will increase further by hybrid algorithms and combining parameters like electricity consumption, whether information, population and adding new one like use of electronics devices in particular region. So we suggest that, combine parameter which previously used and add new like use of electronics devices in particular region with the increase in size of data set and dividing in monthly time period to increase the accuracy of prediction.

ACKNOWLEDGMENT

I acknowledge my sincere and profound gratitude to my guide, Dr. Ajay Parikh, for his valuable guidance, dedicated concentration and support throughout this work. I also acknowledge my sincere gratitude to authorities of Gujarat Vidyapith, Ahmedabad for their help and support. I am also thankful to my friends for their cooperation.

REFERENCES

[1] Noorollah Karimtabar, Sadegh Pasban and Siavash Alipour "Analysis and predicting electricity energy consumption using data mining techniques- A case study I.R. Iran - Mazandaran province, 2nd International Conference on Pattern Recognition and Image Analysis (IPRIA 2015) March 11-12, 2015

[2] Vaclav Uher, Radim Burget, Malay Kishore Dutta, Petr Mlynek "Forecasting Electricity Consumption in Czech Republic", IEEE, 2015, pp. 262-265.

[3] Joaquim L. Viegas, Susana M., R. Melicio, V. M. F. Mendes "Electricity demand profile prediction based on household characteristics", IEEE,2015.

[4] Mohamed M Ismail and M A. Moustafa Hassan "Artificial Neural Network Based Approach Compared With Stochastic Modelling For Electrical Load Forecasting", Proceedings of International Conference on Modelling, Identification & Control (ICMIC) Cairo, Egypt, 31st Aug.- 2ndSept. 2013, pp. 112-118.

[5] Bon-Gil Koo, Min-Seok Kim, Kyu-Han Kim, Hee-Tae Lee , June-Ho Park and Cheol-Hong Kim "Short-Term Electric Load Forecasting Using Data Mining Technique", IEEE Proceedings of7th International Conference on Intelligent Systems and Control (ISCO 2013), 2012, pp.153-157.

[6] Intan Azmira binti Wan Abdul Razak, Md. Shah bin Majid, Hasimah Abd. Rahman and Mohammad Yusri Hassan "Short Term Load Forecasting Using Data Mining Technique", 2nd IEEE International Conference on Power and Energy (PECon 08), December 1-3, 2008, Johor Baharu, Malaysia, pp.139-142.

[7] G. Lambert-Torres, Member, D. da Silva Filho and C.H.V. de Moraes "Intelligent Data Mining in Power Distribution Company for Commercial Load Forecasting", IEEE, 2007, pp. 1-6.

[8] Hoda K. Mohamed, Soliman M. El-Debeiky, Hassan M. Mahmoud, Khaled M. El Destawy "Data Mining for Electrical Load Forecasting In Egyptian Electrical Network", IEEE, 2006, pp. 460-465.

[9] Hang Yang, Huajun Chen, Cai Yuan, Fang Lianhang "An Intelligent System for Forecasting the Trend of Consume Electricity", IEEE International Congress on Big Data, 2014, pp. 677-682.

[10] Daswin De Silva, Xinghuo Yu, Damminda Alahakoon, and Grahame Holmes "A Data Mining Framework for Electricity Consumption Analysis From Meter Data", IEEE Transactions On Industrial Informatics, Vol. 7, No. 3, August 2011,pp. 399- 407.

[11] B.D. Pitt, D. *S* . Kirschen "Application of Data Mining Techniques to Load Profiling", IEEE,1999,pp. 131-136.

[12] LIU Chengshui and YI Hongmei "Research on Power System Load Forecasting Model Based on Data Mining Technology", IEEE International Conference on Intelligent System Design and Engineering Application, 2010, pp. 240-243.

[13] Karin Kandananond **"**Forecasting Electricity Demand in Thailand with an Artificial Neural Network Approach". *Energies* **2011**, *4*, 1246-1257; doi: 10.3390 /en4081246 ISSN 1996-1073

[14] Ali Azadeh, Morteza Saberi, S.F. Ghaderi, A. Gitiforouz, V. Ebrahimipour "Improved Estimation of Electricity Demand Function by Integration of Fuzzy System and Data Mining Approach". ENERGY CONVERSION AND MANAGEMENT · AUGUST 2008

[15] V.H. Kher and S.K. Joshi "Short-Term Electrical Load Forecasting Using Artficial Neural Network". at: https://www.researchgate.net/publication/2287169 60.

[16] Hasan M. H. Owda , Babatunji Omoniwa, Ahmad R. Shahid, Sheikh Ziauddin "Using Artificial Neural Network Techniques for Prediction of Electric Energy Consumption" 6-Dec-2014

[17] C.P. Ronald Reagan, S.R. Sari "Long Term Load Forecasting in Tamil Nadu Using Fuzzy-Neural Technology". International Journal of Engineering and Innovative Technology (IJEIT) Volume 3, Issue 9, March 2014, pp 5-8.

[18] Han J., Kamber M. Data Mining: Concepts & Techniques, Morgan & Kaufmann, 2000.

[19] ISSDA, "Data from the Commission for Energy Regulation." [online]Available: http://www.ucd.ie/issda/data/commissionforenergyregulationcer/

[20] Ministry of Electricity and Energy -Egyptian Electricity Transmission Company, National Control Central for Electricity, "report for statistics data from 1981-2004".

[21] Ministry of Electricity and Energy-- Egyptian Electricity holding Company, "Annual reports of electric statistics", Cairo -Egypt, 2001-2004.

[22] Ministry of Electricity and Energy -Egyptian Electricity Transmission Company, "The difference between using the ordinary calendar or the summer calendar", 2004.