ISSN:0975-9646
Saritha Chava et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (4) , 2011, 1498-1500

Practical Approach for Demand Forecasting for Electricity

Saritha Chava[1], Venkatesh Grandhi[1], Sivaram Jaladi[1], Dharmaiah Devarapalli[2]
[1]*Phoenix IT solutions Ltd, Unit-1, SDF Block-1, V.S.E.Z, Visakhapatnam, India*
[2]*Vignan Institute of Information Technology,Duvvada, Visakhapatnam, India*

Abstract- **For utilities Meter Data Acquisition System (MDAS) is an overall strategy, or process, for building decision support systems and environments that support both everyday tactical decision-making and long-term business strategy. The MDAS provides a common infrastructure for receiving meter data from sub stations, Distribution Transformers, HT/LT consumers and processes the data. This data is shared with utility applications like billing system, energy audit systems, etc. The MDAS contains two subsystems - Data Acquisition Server (DAS) connected to cellular or telephone network for managing Advanced Metering Infrastructure (AMI) and the Meter Data Management System (MDMS). AMI is the infrastructure within which date and time-stamped meter data are remotely collected and transmitted to a Data Acquisition Server and then to a centralized MDMS. The DAS will use PSTN or cellular network with GPRS to connect to all data sources such as Data Concentrator Units (DCU) installed at sub stations, energy meters installed at Distribution transformers, and HT/LT consumer premises. The AMI systems transmit meter data according to configured parameters to the DAS (Data Acquisition Server) using specified protocols and data transfer structure (common data structure for all types of meters). MDMS Data Warehouse is designed for storing and managing vast amounts of historical interval meter data (5 minute, 15 minute, etc), monthly cycle meter data and monthly profiled meter data along with the required ancillary information needed to effectively "mine" and report useful information. MDMS allows utilities to utilize an enterprise-wide meter data store to link information from diverse sources and make the information accessible for a variety of user purposes such as monitoring system performance, settlement, loss analysis and historical operational reporting. In reporting system, there is a need for precision in the load / demand forecasts. The load forecasting further drives various plans and decisions on investment, construction and conservation. Since electric utilities are basically dedicated to the objective of serving consumer demands, the consumer can place a reasonable demand on the system in terms of quantity of power. Also, the utility can then plan the power purchase requirements so as to meet the demand while maintaining the merit order dispatch to achieve optimization in the use of their resources.**

Keywords: **Load Forecasting, Support Vector Machine, Time series**

I. INTRODUCTION

Electricity demand forecasting plays an important role for the management of power systems. A good electrical load forecast shows direct and significant impact on cost - generating unit start-ups and shutdowns, energy purchases, managing system demand as well as scheduling system upgrades based on predicted load growth. Load forecasting is classified as (a) short range (~day), (b) medium range (~1 week to 1 month) and (c) long range (several years). Long-term forecasts of the electricity demand are required for capacity planning and maintenance scheduling .Medium-term demand forecasts are required for power system operation and planning. Short-term load forecasts are required for the control and scheduling of power systems. Each class of load forecasting uses different models to meet the specific objectives of the application. Load forecasting plays a pivotal role in the formulation of economic, reliable and secure operating strategies for the power system. The present work applies SVM load forecasting models. But SVM assures accuracy when compared to other models like linear regression, ANN and Artificial Intelligence. The new SVM learning algorithm is called Sequential Minimal Optimization (or SMO).

II. DATA AND TASK DESCRIPTION

The organizer of the load competition provides competitors the following data:

Electricity load demand recorded every day from August 1st 2008 to August 20th 2008

The task of competitors is to supply the prediction of maximum monthly values of electrical loads and evaluation of submissions would mainly depend on the error metric of the results:

$$M=\frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t-F_t}{A_t}\right|$$

Here n=20

Where, A and F are the real and the predicted value of maximum daily electrical load, t is the day of the year and n is the number of days in between 1st Aug 2008 to 20th August 2008. The goal of the competition is to forecast electrical load with minimum MAPE (Mean Absolute Percentage Error).

III. METHODS

Support vector machines (SVM) are a group of supervised learning methods that can be applied to classification or regression. We briefly introduce support vector regression (SVR) method which can be used for time series prediction.

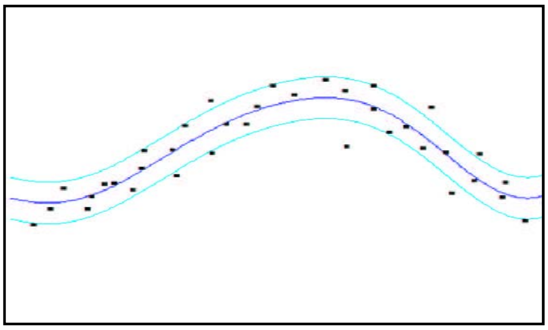

Fig1: Support Vector Machine

SVR is a more general and flexible method for working on regression problems. For experiments in this paper, we use the software SMO-Reg which is a library for support vector machines.

In SVM, different types of kernels are available. The kernel represents a legitimate inner product in feature space. The training set is linearly separable in the feature space. This is called the "Kernel trick". The different kernel functions are explained below.

1] *Polynomial:* A polynomial mapping is a popular method for non-linear modelling. The second kernel is usually preferable as it avoids problems with the hessian becoming Zero.

$$K(x,x')=(x,x')^d$$

$$K(x,x')=((x,x')+1)^d$$

2] *Gaussian Radial Basis Function*: Radial basis functions with Gaussian form.

$$K(x,x')=\exp\left(\frac{(||x-x'||^2}{2\sigma^2}\right)$$

3] *Exponential Radial Basis Function*: A radial basis function produces a piecewise linear solution when discontinuities are acceptable.

$$K(x,x')=\exp\left(\frac{||x-x'||}{2\sigma^2}\right)$$

4] *Multi-Layer Perception*: The long established MLP, with a single hidden layer, also has a valid kernel representation.

$$K(x,x')=\tanh(\rho(x,x')+e)$$

The best prediction accuracy was obtained using the RBF kernel, resulting in a MAPE of 1.32%. From this point onwards, all experiments were performed using the RBF kernel.

The interface design for the load forecasting should be easy, convenient and practical. The users could easily define what they want to forecast, whether through graphics or tables. The output should also be with the graphical and numerical format.

IV. EXPERIMENTS

In Section III, we have described the SVM technique. In this section we are explaining the process of preparing datasets to build SVM models.

Different data encodings affect the selection of modelling schemes. Here, we are discussing the schemes in detail.

Calendar attributes

We discussed about the daily periodicity of the load demand. Also, as we mentioned earlier, the load demand during night times is lower compared to day time. Therefore, encoding this information (daytime and night-time) in the training entries might be useful to model the problem. Actually, among literatures of the load forecasting, many other works have used the calendar information (time, dates or holidays) to model the load forecasting problem.

Season

Another possible feature is the Season data. This is quite a straightforward choice, since load demand and season have a causal relation. Load Forecasting will vary with seasons - winter, summer and rainy seasons. This criterion is also used to predict the load demand. However, to include the season in the training entries, we considered the season's data of August 1st 2008 to August 20th 2008.

Time series

Besides the day time, night time, weekdays, holidays and season, we considered to encode the attributes: the past load demand based on interval. This is to introduce the concept of time-series into our models. To be more precise, if y_i is the target value for prediction, the vector x_i includes several previous target values as y_{i-1}......... y_{i-n} attributes. In the training phase all yi are known but for future prediction, y_{i-1}......... y_{i-n} can be values from previous predictions. For example, after obtaining an approximate load on Feb 2011 and if n=12 then it is used with loads of Jan 2010 to Jan 2011 for predicting data of future months. The same method is continued to arrive at an approximate load of March 2011. We know that the past load demand could affect and imply the future load demand. Therefore, considering inclusion of time series information in the models might help to better forecast the load demand.

V. IMPLEMENTATION AND RESULTS

With different schemes of model construction, a series of experiments are conducted on the datasets. We prepared SVM models for load forecasting. When training an SVM model, user can select the desired parameters. The performance of SVM model will be influenced by the selected parameters. So, in order to get a "good" model, these parameters should be selected properly. Some important parameters are:

1) Cost of error Y.
2) The width of the e-insensitive tube.
3) The mapping function
4) Load of previous days included for one training set data.

In our experiments, as we mentioned earlier, for each training data we simply include the maximum load of the previous 20 days. In addition, we consider only the Radial Basis Function (RBF), which is one of the most commonly used mapping functions. The RBF function has the property that *k (x, x') = exp (-σ |x - x'|^2)*. Note that σ is a parameter associated with the RBF function which has to be tuned. Also, we fix e=0.01. For experiment, selection of parameters is a repeated process. So, we decided to use parameters automatically by using business logic.

According to the performance of the validation set, we try to infer the proper values of Y and X. Due to different characteristics of the data encoding schemes; we employ two procedures for the validation.

For time-series-based data sets, we extract the data entries of August 1st 2008 and August 20th 2008 to form the validation set and to evaluate the models. The performance is decided by averaging the errors of these two validations. For non time-series models, we simply conduct 10-fold cross validations to infer the parameters. That is, we randomly divide the training sets into 10 sets. Using each set as a validation set, we then train a model on the rest. The performance of a model would be the average of the 10 validating predictions. With this procedure, proper Y and X are selected to build a model for the future prediction. We then evaluate the performance by forecasting the load demand up to August 25th 2008 and for selected number of future days. Here, we check a condition to restrict the number of predictions based on number of inputs in training set. Figure 2 shows the past and future predicted electricity load curve in different colours by using time series data sets. In graph, Red curve shows the actual data and violet curve shows the prediction data. In this model, we achieved good accuracy. So, for our problem the SVM is the best model for predicting the load.

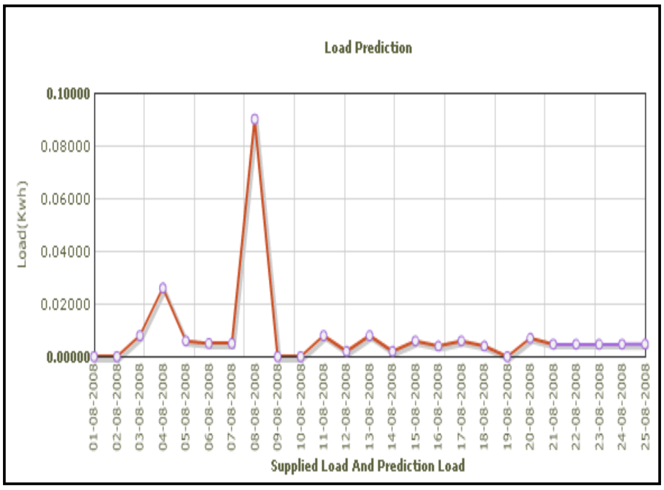

Fig: 2.Past and Future Load Curve

VI. CONCLUSIONS

In this study, the methods for forecasting electric loads on a power system were discussed. Emphasis is on load forecasting - which is important for real time operation and control of power systems by utilities. Deterministic, probabilistic and stochastic forecasting models were examined. The performance of the discussed model is dependent on the characteristics of electric loads and on the assumption that electric load patterns are basically invariant with time. The load forecasting methodology that we adopted in this study proved its accuracy during the last one year of implementation. The peak demand occurrence at day or night in the future will depend on the season.

This load forecasting model will play an important role in the operation and planning of electric power distribution. So, it is good to have a forecasting model in each power station with feed back to the central control room in order to run the system with higher efficiency and better reliability.

ACKNOWLEDGEMENTS

The authors would like to thank Phoenix IT Solutions Ltd, Unit-1, SDF Block-1, V.S.E.Z, Visakhapatnam, India for providing the opportunity to research on *Practical Approach for Demand Forecasting for Electricity* using data mining techniques.

REFERENCES

1.Takeshi Haida, Shoichi Muto: "Regression Based Peak Load Forecasting Using A transformation Technique", IEEE Trans. on Power Systems, Vol. 9, No.4, November 1994, pp. 1788 - 1794

2. H.L.Willis, J.E.D.Northcote-Green: "Comparison tests of fourteen distribution load forecasting methods", IEEE Trans. on Power Apparatus and Systems, Vol. PAS-103, No.6, June 1984, pp. 1190 - 1197

3.K. Jabbour, J. F. V. Riveros, D. Landsbergen, and W. Meyer. Alfa: Automated load forecasting assistant. IEEE Transactions on Power Systems, 3(3):908-914, August 1988.

4.H.L.Willis, R.W.Powell, D.L.Wall: "Load transfer coupling regression curve fitting for distribution load forecasting", IEEE Trans. on Power Apparatus and Systems, Vol. PAS-103, No.5, May 1984, pp. 1070 - 1076

5. H.L.Willis: "Load forecasting for distribution planning-error and impact on design", IEEE Trans. on Power Apparatus and Systems, Vol. PAS-102, No.3, March 1983, pp. 675 - 6865.S.

6. Rahman and O. Hazim. A generalized knowledge-based short-term load-forecasting technique IEEE Transactions on Power Systems, 8(2):508-514, May 1993.

7.Takeshi Haida, Shoichi Muto: "Regression Based Peak Load Forecasting Using A transformation Technique", IEEE Trans. on Power Systems, Vol. 9, No.4, November 1994, pp. 1788 - 1794

8. H.L.Willis, J.E.D.Northcote-Green: "Comparison tests of fourteen distribution load forecasting methods", IEEE Trans. on Power Apparatus and Systems, Vol. PAS-103, No.6, June 1984, pp. 1190 - 1197

9. H.L.Willis, R.W.Powell, D.L.Wall: "Load transfer coupling regression curve fitting for distribution load forecasting", IEEE Trans. on PowerApparatus and Systems, Vol. PAS-103, No.5, May 1984, pp. 1070 -1076

10. H.L.Willis: "Load forecasting for distribution planning-arror and impaction design", IEEE Trans. on Power Apparatus and Systems, Vol. PAS-102, No.3, March 1983, pp. 675 - 686

11. H.N.Tram, H.L.Willis, J.E.D. Northcote-Green, C.L.Brooks: "Load forecasting data and data base development for distribution planning",IEEE Trans. on Power Apparatus and Systems, Vol. PAS-102, No.11,November 1983, pp. 3660 - 3666

12. Mo-yuen Chow, Hahn Tram: "Methodology of Urban Re-Development Considerations in Spatial Load Forecasting", IEEE Trans. on Power Systems, Vol. 12, No.2, May 1997, pp. 996 -1001

13. E. - C. Yeh, Hahn Tram: "Information Integration in Computerized Distribution System Planning", IEEE Trans. on Power Systems, Vol. 12, No.2, May 1997, pp. 1008 - 1013

14. C. W. Gelling, R. W. Taylor: "Electric Load Curve Synthesis - A Computer Simulation of an Electric Utility Load Shape", IEEE Power Engineering Society Summer Meeting, 1980 or Talukdar & Gelling:"Load Management", IEEE Press, pp. 111-116

15. J. H. Broehl: "An End-Use Approach to Demand Forecasting", IEEE Power Engineering Society Summer Meeting, 1980 or Talukdar & Gellings: "Load Management", IEEE Press, pp. 117-121

16. Hong-Tzer Yang, Chao-Ming, Ching-Lien Huang: "Identification of ARMAX Model for Short Term Load Forecasting: An Evolutionary Programming Approach", IEEE Trans. on Power Systems, Vol. 11,No.1, February 1996, pp. 403 - 40