A Prediction Model and Its Affecting Parameters for Prediction of Rural Electricity Requirements

Rajnikant Pandya
Student-Research Scholar (Computer Science)
Gujarat Vidyapith-Ahmedabad
Gujarat, India

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

***Abstract*—** In today's era, data is the most important part of any organization. Apart from importance, the size of data increasing very fast, so amount of data become very large. These large amounts of data can be used very effectively with data mining technique for prediction and classification to retrieve hidden knowledge.
Demand of electricity in rural area increasing day by day, so forecasting of electricity demand for rural area become major part of power Supply Company.
In this paper, we present the study of various parameters affecting to prediction of electricity demand of rural area. Using these parameters we may achieve higher accuracy of electric power requirement by discovering knowledge from historical data. The prediction of electricity demand may be influenced by many factors, so it becomes necessary to develop predictive data mining model. For this prediction model, we have to analyze primary and secondary data. Primary data are collected from PGVCL Circle Office-Amreli and secondary data are collected from online resources and vendors of electronics devices. At the end we will conclude the result against our parameters.

Keywords- *Data Mining, Electricity demand forecasting, Prediction, Predictive model.*

1. INTRODUCTION

Data mining has ability to discover hidden knowledge, patterns and relationship among various parameters. This is very helpful in decision making process of any organization. Data mining has been successfully used in different areas including electricity demand forecasting. Electricity demand forecasting of rural area is an interesting research area which extract useful and previously unknown facts from database. In order to meet electricity demand continually and having sustained economic and overall growth of rural area, prediction of electricity demand has become very important part.
The term data mining comes with some handy technique and algorithms for prediction, classification etc. Normally data mining process divided into two parts.

1. Prediction of unknown patterns, trends and behaviors.
2. Classification.

This research falls under first part, in which we have tried to predict rural electricity demand from historical data. An accurate prediction becomes more useful in developing power supply strategy, planning and management.

2. BACKGROUND STUDY AND RELATED WORK.

There are number of studies on the relationship among power consumption and prediction of power usage rate. Some previous researcher describes methods based on consumption data only and some of researcher use additional data and apply various techniques and algorithms.
There are many studies that support the applicability of data mining techniques for electricity load forecasting.
N.Karimtabar, S. Pasban and S. Alipour et al [3] have taken the issue of precise prediction of electric power requirement to meet with actual consumption with an objective to reduce MAPE/Relative error. The datasets that were collected are related to Islamic republic of Iran. The data is recorded between 1991 and 2013. The data used for creating the dataset includes parameters population, temperatures, moisture, electricity price as independent variables and electricity consumption rate (total amount, domestic, commercial, industrial, and agriculture) as target variable.
They used regression, Neural Net and SVM (Support Vector Machine) methodologies and could achieve relative error (%) 0.9, 1.6 and 11.1 and correlation coefficient 0.996, 0.968 and 0.976 respectively.
LIU Chengshui and YI Hongmei et al [5] have taken the issues of power system load forecasting to meet with actual demand with an objective to reduce the predicting average error. In this paper the 2001-01 to 2001-05 data are selected as the training set, it take the data from June to September as the test data to validate the algorithm. The attributes of dataset are date type, historical load, temperature, humidity, wind speed, light, rain, snow, air pressure.
They used Set theory, Genetic algorithm and new ANN with attribute reduction algorithm methodologies and could achieve the average error for month of June to September using new ANN methodology are 3.05, 3.12, 2.24 and 3.16 respectively.
They also compared this average error with traditional method for June to September and these are higher than new ANN i.e. 4.21, 4.46, 3.67 and 4.48 respectively.
K. Kandananond et al [1] have taken the issues of precise prediction of electric power requirement in Thailand to meet with actual demand with an objective to reduce MAPE[Mean Absolute Percentage Error]. The energy data that were collected from Thailand. The data is recorded from 1986 to 2010. The attributes used for creating the dataset are year,

population, GDP, SET index, Export (Million baht) and Electricity Consumption (GWh). He used ARIMA (0,2,2) [Auto Regressive Integrated Moving Average], MLR [Multiple Linear Regression] and ANN [Artificial Neural Network] methodologies to reduce MAPE and could achieve MAPE 2.80981%, 3.2604527% and 0.996% respectively. In ANN model MLP (4,6,1) [Multi Layer Perceptrons] obtain the best result i.e. 0.0996% MAPE.

Ali Azadeh, Morteza Saberi, S.F. Ghaderi, A. Gitiforouz, V. Ebrahimipour et al [2] has taken the issues of estimation and prediction of electricity demand for developing countries. This study is carried out using the data from period of 1995 to 2005 of monthly electricity consumption of IRAN.

They used Genetic algorithm, ANN and proposed new integrated fuzzy algorithm methodologies and could achieve MAPE 0.014, 0.0156 and 0.02 respectively. Here MAPE result showed the negative impact of preprocessing data on fuzzy system.

Hasan M. H. Owda , Babatunji Omoniwa, Ahmad R. Shahid, Sheikh Ziauddin et al [4] are carried out study for prediction of electric energy consumption for the period of 2012 and 2013. Monthly data for electric energy consumed in Gaza strip was collected from year 1994 to 2013. The attributes of data set are energy consumption, temperature, humidity, population, and GDP.

They used ANN, Fuzzy logic and genetic algorithms and create new proposed model. This proposed model was validated using 2-Fold and K-fold cross validation techniques and could achieve MSE, RMSE, MAE and MAPE for year 2012 are 2-fold→1.21%, 1.10%, 1.21% and 120.91% and for K-fold→ 3.87%, 1.97%, 3.87% and 386.54% respectively. Same as for the year 2013 are 2-Fold→1.74%, 1.32%, 1.74% and 173.76% and for K-Fold→ 1.88%, 1.37%, 1.88% and 187.65% respectively.

3. DATA COLLECTION

Here we have collected the data from various sources. The historical data of power consumption, power generation and power loss are collected from PGVCL Circle Office-Amreli for amreli rural sub division. Weather data are collected online [6][7] and electronics device selling data are collected from various vendors of amreli city.

4. PARAMETERS FOR PREDICTION

In this research, we used Historical data of 11 years and 10 month for prediction. Our key steps of research at this point are as under

- Deciding parameter list.
- Value of each parameter
- Analyzing data
- Data cleaning
- Prediction using this data under proposed model
- Verify the actual result and the result generated by the proposed model.

Affecting parameters are extracted and list out here for reference. Here we have total 18 parameters.

Parameter Name	**Description**	**Remark**
Year	It indicates the year of data	Integer Number
Month	It indicates the month of data	Integer Number (All data are collected monthly)
Total_Power_Generation	Total sent power in MUs.	Decimal Number
Power_Consumption	Total consumed power in MUs.	Decimal Number
Power_Loss	Total power loss in MUs	Decimal Number
High_Temp	Maximum temperature in Celsius.	Integer Number
Low_Temp	Minimum temperature in Celsius.	Integer Number
Humidity	Humidity in fraction	Decimal Number
Wind Speed	Wind Speed in mile per second	Decimal Number
Precipitation	Rain in millimeter	Decimal Number
Population	Total population of specified region	Integer Number

TV	Number of Television in specified region	Integer Number
REF	Number of Refrigerator in specified region	Integer Number
WM	Number of Washing Machine in specified region	Integer Number
AC	Number of Air Conditioner in specified region	Integer Number
Oven	Number of Oven in specified region	Integer Number
Cooler	Number of Cooler in specified region	Integer Number
Atamaker	Number of Atamaker in specified region	Integer Number

Month & Year – All the data are collected monthly basis, so this parameter indicates particular month & year of all other data.

Total_Power_Generation- This is a total power sent to specified region for its consumption requirement. Value of this parameter is in MUs. [1 MU= 10,00,000 Units].

Power_Consumption- This parameter is total power consumption of specified region in MUs.

Power_Loss- This parameter indicates total power loss of specified region. This is a sent power minus consumed power.

High_Temp & Low_Temp – These parameters indicates maximum and minimum temperature of that region in Celsius.

Wind Speed – This is a speed of wind in mile per second.

Humidity- This variable will be assigned a value of humidity in environment in fraction.

Precipitation- This is a rain fall in a specified region. Rain fall measured in MM i.e millimeter.

Population- This variable indicates total population of all villages of specified region.

TV- This parameter indicates total number of television in all villages of specified region.

REF- This parameter indicates total number of refrigerators in all villages of specified region.

WM- This variable indicates total number of washing machine in all villages of specified region.

AC- This parameter indicates total number of air conditioner in all villages of specified region.

Oven- This parameter indicates total number of oven in all villages of specified region.

Cooler- This variable indicates total number of cooler in all villages of specified region.

Atamaker- This parameter indicates total number of atamaker in all villages of specified region.

Our specified region is amreli rural sub division. Sub division partition is according to feeder distribution of PGVCL circle office amreli.

5. EVALUATION OF PARAMETERS

For the development of the final model, each parameter has to be analyzed and evaluated and finds its impact on final result.

6. FUTURE WORK AND CONCLUSION

Here we have collected the most affecting parameters for the prediction of electricity demand for rural area. As a space for proposed work, in our next paper we will test each and every parameters against the model.

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] 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

[2] 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

[3] 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

[4] 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

[5] 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.

[6] For weather data collection online resource is https://www.timeanddate.com/weather/@1278715

[7] For weather data collection online resource is http://globalweather.tamu.edu

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

[9] S. Gnanapriya, R. Suganya, G. Sumithra Devi, M. Suresh Kumar "Data mining concept and technique", Data mining and knowledge engineering Vol.2 no 9, 2010

[10] ISSDA, "Data from the Commission for Energy Regulation." [online]Available: