International Journal of Technical Research and Applications e-ISSN: 2320-8163,
www.ijtra.com Volume 3, Issue 2 (Mar-Apr 2015), PP. 222-225

FUZZY AND DECISION TREE APPROACH FOR FORECASTING ANALYSIS IN POWER LOAD

J. Kumaran@kumar, P. Santhosh Kumar, P. Rajkumar
Assistant Professor, Student, Student,
Department of CSE,
Pondicherry Engineering College
Pondicherry, India.
Kumaran@pec.edu, Santhoshkumar.p.91@pec.edu, Raj_dcs@pec.edu.

Abstract— The most important challenges in electric load forecasting is to find the accurate electricity load forecasting. Because, it is volatile in nature and has to be consumed immediately. Fuzzy Decision Tree is applied to predict the annual electricity requirement in India. Population and Per Capital gross domestic product (GDP) are taken as input variables and the electricity consumption is predicted output variable. Past 30 years of historical data has been used for training and 4 year of data is used for testing the fuzzy decision Tree. Comparatively to measure the accuracy of data has been made with Artificial Neural Network and Fuzzy Decision Tree model using Mean Absolute percentage error (MAPE). The results states that proposed Decision Tree model has given high performance and less error rate than the Artificial Neural Network model.

Index Terms— Load forecasting, fuzzy logic, decision tree.

I. INTRODUCTION

Nowadays, the demand for electricity is increasing tremendously. Hence, it's important to predict the power consumption based on the statistics of population growth and gross domestic product values. Neural Networks can be used to extract the patterns and detect the trends that are too complicated by various techniques. To solve the forecast problem using fuzzy decision tree model and applied this technique in real world dataset provided in the web.

A. Type of forecasting:

1) Short Term load forecasting (STLF):
Short term load forecasting is to predicting and forecast the electric power consumption using in per day. STLF is important role for utility operations and real-time control for security functions.

2) Medium Term load forecasting:
Medium term load forecasting is to predicting and forecast the electric load for one day to one year.

3) Long term load forecasting:
Long term load forecasting is to predicting and forecast the electric power consumption using more than one year. LTLF is important role in generation, transmission and distribution operations.

II. RELATED WORK

A. Artificial Neural Network

Artificial Neural Network (ANN) are real world problem in modeling complex. A neural network is a machine learning approach that is inspired by brain at particular learning task [1], [2]. ANN has fewer connection for processing elements (neurons) to solve specific problems. A neural network can be trained and adjusting the values of weights to perform particular function between elements so that a particular input leads to a specified target output [1],[3]. Therefore, the network is adopted, based on a comparison of the output and the target. In ANN model [2], [3],[4] depicts the values with the number of hidden layers and the number of neurons present in the hidden layer. This calculated value provides the error rate for the electricity consumed for the particular time period.

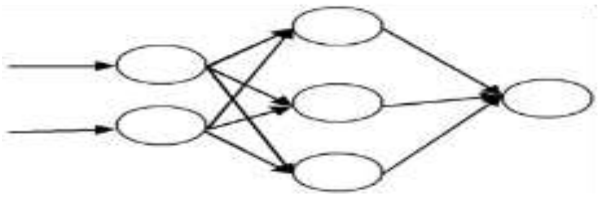
B. Fuzzy Logic

A fuzzy logic [5] system is to handle numerical data and linguistic knowledge. It is a non linear mapping of an input data vector into a scalar output. i.e. mapping numbers into numbers.

A fuzzy linear regression model [5], [6] for summer and winter seasons is developed. The estimation fuzzy problem for the model is turned out to linear optimization problem, fuzzy linear regression [7], and is solved using the linear programming based simplex method. The parameters of fuzzy, the fuzzy load for one day ahead is predicted. Results of predication of 24 hour load ahead for a utility company are presented. It has been found using such fuzzy model; a reliable operation for the electric power system could be obtained.

III. EXISTING MODEL

An ANN as a system consists of fewer connection of elements to solve particular task. These interconnecting elements in an ANN are known as neurons [8], [9]. Each neuron may have several inputs and only one output. The output of the input layer is send to the hidden layer, which in turn sent to the output layer. The multilayer perceptron's with back propagation learning algorithm [9],[10] is used in ANN model because it is suitable model for long term forecasting, which helps in the non linearity incomplete or noisy in the input data.



Input Layer Hidden Layer Output Layer
Fig. 1: Structure of ANN model

The network can be trained by adjusting the weight between the neurons[11]. Therefore, the neural network is accepted, based on a similarity of the target and the output. Back-propagation is a regular method of training multilayer artificial network. The inputs are passed to the network, then the output is calculated based on the network selected and the output is compared with the targets. If the output obtained output is not closer to the target value then the learning rule is used to adjust the weights between the neurons in order to get network outputs closer to the targets. Fig. 1 shows the structure of ANN model.

IV. PROPOSED MODEL

The development of Fuzzy Decision Trees (FDT) differs from traditional decision trees in two respects: it uses splitting criteria based of fuzzy restrictions and its inference procedures are different [12]. To adopting the following three steps, for developing proposed FDT: (i) Fuzzifying the training data; (ii) Inducing and pruning the generated tree; and (iii) Applying fuzzy rules represented by the tree for forecasting [12]. Therefore figure 2 explain the process of decision tree generation.

A. Fuzzifying training data

The analysis based on 30 years yearly electric load. The data covered GDP and population from that to extract the six sector of electric load.

Based on careful analysis of data and the information elicited from the experts of six sector. To selected the trapezoidal membership function due to its simplicity and implementation. Therefore figure 3 designed the suitable membership function and the set of linguistic terms for each of the variable.

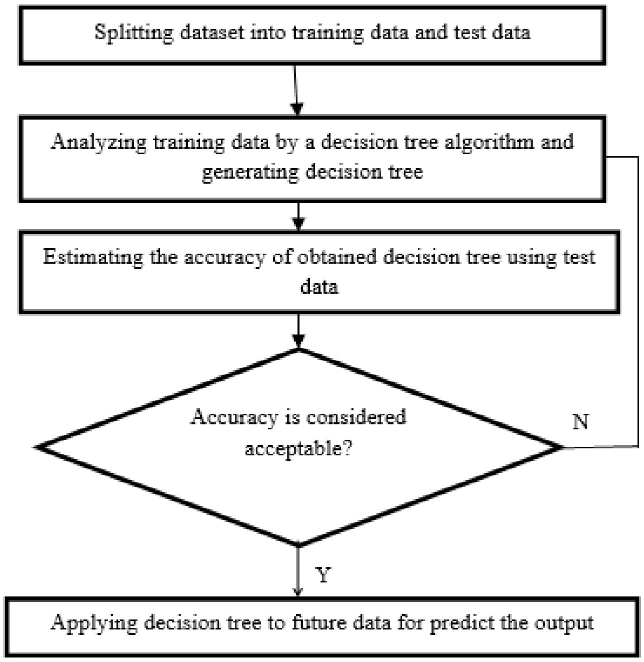


Fig.2 The Process of Decision Tree Generation

When fuzzify these rules, the result will be forecasted and to be particular instance. From analysis of load we assigned the minimum and maximum load. The forecasted load was categorized using numerical value from 0.00(minimum) to 1.00(maximum).

Given points X_1, X_2 , X_3, X_4 as shown in figure 1, the membership function for the various variable is defined by the equation.

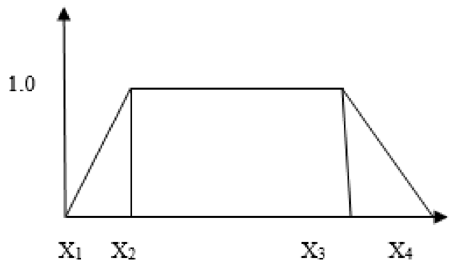


Fig. 3. Trapeziodal Memebership function Template

$$\mu_v(x) = \begin{cases} \dfrac{X_2 - x}{X_2 - X_1} & X_1 \le x \le X_2 \\ 1.0 & X_2 \le X \le X_3 \\ \dfrac{X_4 - x}{X_4 - X_3} & X_3 \le x \le X_4 \end{cases}$$

B. Inducing and Puring the fuzzy Decision Tree

The procedure FDT is comprise three elements: (i) selection of splits at every new node of the tree, (ii) a rule for determining when a node should be considered as the terminal and (iii) a rule for assigning labels to identified terminal nodes.

A number of heuristics have been proposed forbuilding and pruning fuzzy decision trees [12]. The develop method adapting [12] the simplicity and the availability of software for its implementation.The algorithm for building the FDT is presented as follows:

G=GetData();[gdp,population]
TreeBulit=F
While!TreeBulit {
XRoot=Weight
While!EndofNodes() {
Compute the data
$P_K^M = \sum|G|_{j=1} f_{k+1}(X_j^M, \mu_{vk}c(y_j))$
$P^M = \sum|D_c|_{k=1} p_k^M$
Computing the std information
$I^M = -\sum|D_c|_{k=1}(\frac{p_k^M}{p_k^M} * \log \frac{p_k^M}{p_k^M})$
Nodes I {
search the set of remaining attribute from $v\text{-}v^i$ by
Computing $I^{sM}V_i$
Selecting attribute V_i such that info gain G_i^M is maximal
}
If(X_j^M>0.0 has unique claisfication) or (v^M=v))
{
TreeBulit=T
}
}
Split M into |Di| sub nodes by making
Child M|Vpi get samples defined by $X^M|Vp^i$ and
Computing new membership using fuzzy restrictions
Leading to $M|V_p^i$ using equation
$X_j^M|V_p^i$= f1 (f0(e_j,V_p^i),X_j^M)
}

In addition to the algorithm, the parameters, variables, and functions used in develop FDT design are listed .The training data were generated from the yearly electric load for n-1 data, as well as the nth were used as a test data. In this model the entire database (E) is divided into two sets: (i)the training and (ii) test sets.

Variable/ parameter	Description
V ={V1, V2, V4, V5, V6, V7}	The set of seven fuzzy input variables
D = $D^i_1 D^i_2, D^i_3$	The set of terms defined over variable Vi.
v^i_p	The fuzzy term p for variable

	Vp,e.g. μ^{Temp}_{Low}
$u_i \in Ui$	The crisp data for the variable Vi.
D_c	The fuzzy terms for the Fload decision variable.
$E = \{e_j \| e_j = u1_j, u2_j, . . . ,un_j, y_j\}$	The training example database, each item in the database is viewed as a event
$W = \{w_j\}$	The confidence weight.
Wj	The weight of ej∈E.
M	Nodes in the fuzzy decision tree.
F_M	The set of fuzzy restrictions on the path leading to M.
V^M	The set of attributes appearing on the path leading to M.
$X^M = X^M j$	The set of memberships in M for all the training examples.
$M\|V^i_p$	The particular child of node M created by using V_i to split M.
S^M_V	The set of M's children when $V_i \in V - V_N$ is used for the split
P^M	The total example count for node M.
I^M	Information measure for node M.
$P^N k$	The example count for decision $v^c_k \in Dc$ in node M.
$G^M i = I^M - I^{SM}_{VI}$	The information gain when usingVi in M.
$\mu(.) : X \rightarrow [0.0, 1.0]$	A mapping from X to[0.0, 1.0].

The training set is further divided into two other disjoint sets. One of these sets was used for building the tree while the other was used for pruning the resulting tree. The procedure of Fuzzy decision tree utilizes the training data and added in a top-down fashion, until the tree nodes are successfully to stopping criteria are met. The resulting tree is optimized by a pruning procedure which acts in a bottom-up version, to remove irrelevant parts of the tree. To deal with missing values, an example is split into all children if the need feature value is not available. The percentage of the event with unknown values is reduced by attribute utilization.

If P_{ui}unknown denotes the total count of examples in node M with unknown values for Vi, then the split function $f_r()$ at splitting point r is defined as:

$$f_r(e_i, [V_i \text{ is } v^i_p]) = \begin{cases} 1.0 & \text{if } u^i_j \text{ unknown,} \\ |Di| & \\ \mu v^i_p(u^i_j) & \text{otherwise} \end{cases}$$

Note that |Di| is the cardinality of set Di, which is the number of elements in the set. In our model to extract the electric by GDP and Population.

V. EXPERIMENT RESULT

The 30 year historical data of power consumption, Population and GDP has been used for prediction. The historical data for population, GDP and power consumption was taken from webpage [13],[14],[15]. Table 1 shows comparison of the existing ANN model and proposed FDT forecasted values for the year 1986, 2000, 2006, and 2010. In ANN model the hidden layer eighth neurons gives the best result among other neurons. FDT model is constructed and implemented using the same input GDP and Population which is used in ANN model. To evaluate the performance of proposed FDT model results are compared with neural networks with back propagation algorithm. Mean Absolute Percentage Error (MAPE) is used for the performance metrics.

$$MAPE = \frac{\text{Actual value} - \text{Forecasted Value}}{\text{Actual value}} * 100$$

Year	Actual output (BkWh)	Forecasted Output(BkWh)		%error	
		ANN	FDT	ANN	FDT
1986	123.09	123.26	122.58	0.14	0.41
2001	316.60	314.98	316.54	0.51	0.018
2006	415.29	415.82	415.36	0.13	0.016
2010	612.64	613.59	612.35	0.16	0.047
Average absolute error				0.94	0.49

Table 1: Power Consumption: Actual Vs Forecasted

Table 1 shows power consumption of actual and forecasted value using artificial neural network and FDT model for the year 1986, 2001, 2006 and 2010 and also it shows the error rate of actual and forecasted value. Table 2 shows power consumption of future years upto 2019 using ANN and FDT model.

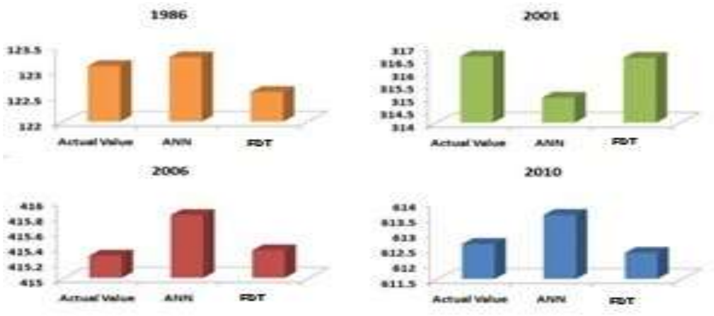


Fig. 4 Power Consumption:Actual Vs Forecasted

Table 2: Forecasted Future Power Consumption: ANN and FDT model

Year	Forecasted Output (BkWh)	
	ANN	FDT
2016	716.4	912.4
2017	758.38	979.2
2018	802.52	1048.0
2019	848.0	1118.5

VI. CONCLUSION

The power consumption of India is forecasted by using the social factors: GDP and Population. The two model FDT and the ANN model of multilayer perceptron (MLP) with back propagation network of different layer with different node have been studied and implemented using 30 years historical data and it has been tested with 4 years historical data. The proposed fuzzy Decision Tree model is implemented using the same historical data which is used in the ANN model. The result of both models is validated using the MAPE. The MAPE obtained by ANN model is 0.94 and the proposed fuzzy Decision Tree Model is 0.49. To conclude based on obtained result that the consumption using the fuzzy Decision Tree (FDT) model is efficient and more accurate than the ANN model.

REFERENCES

[1] S.Saravanan, S.Kannan, and C. Thangaraj, "Forecasting India's Electricity Demand Using Artificial Neural Network", IEEE International conference 2012.
[2] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feed forward networks are universal approximators," in the proceeding of the IEEE Transaction on Neural Networks, vol. 2, no. 5, pp. 359–366, Oct 1995.
[3] K. Hornik, "Approximation capabilities of multilayer feed forward networks," in the proceeding of the IEEE Transaction on Neural Networks, vol. 4, no. 2, pp. 251– 257, Jul 1991.
[4] M. Leshno, V. Y. Lin, A. Pinkus, and S. Schocken, "Multilayer feed forward networks with a no polynomial activation function can approximate any function," in the proceeding of the IEEE Transaction on Neural Networks, vol. 6, no. 6, pp. 861–867, Aug 1993.
[5] Rustum Mamlook, Omar Badran, Emad Abdulhadi, "A fuzzy inference model for short-term load forecasting " a Middle East University for Graduate Studies, Elsevier 2009.
[6] Manish Sankar, B. Yeganarayana, DeepakKhemani "Backpropagation learning algorithms for classification with fuzzy mean square error" Elsevier 1998.
[7] D K Ranaweera, N F Hubele and G GKarady, "Fuzzy logic for short term load forecasting" in Elsevier transacting on Electrical & Energy Systems Vol.18 No, pp.215-222,1996.
[8] A. Ghanhari, A. Naghavi, S.F. Ghaderi, and M. Sabaghian, "Artificial Neural Networks and regression approaches comparison for forecasting Iran's annual electricity load," Proceeding International Conference on Power.
[9] Bahman Kermanshahi, Hiroshi Iwamiya, Enviroinmental Engineering," Up to year 2020 load forecasting using neural nets", Elsevier 2002.
[10] ParasMandal, TomonobuSenjyu, Naomitsu Urasaki, ToshihisaFunabash,"A neural network based several-hour-ahead electric load forecasting using similar days approach"Elsevier 2006.
[11] Chih-chouchiu, deborah f. cook, jen-lung kao and yu-chaochou" Combining a neuralnnetwork and a rule-based expert system for short-term load forecasting" Elsevier 1997.
[12] Stephan Trenn "Short-term electric load forecasting in uncertain domain" in the proceeding of the IEEE Transaction on Neural Networks, Vol.19. No.5, MAY 2011.
[13] http://www.populstat.info/
[14] http://www.iasri.res.in/agridata/12data/chater1/db201213.pdf.
[15] http://www.indexmundi.com/india/gdp_(purchasing_power parity).html