ISSN 2278-3091
Volume 9 No.2, March - April 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse152922020.pdf
https://doi.org/10.30534/ijatcse/2020/152922020

Short Term Load Forecasting using Regression Trees: Random Forest, Bagging and M5P

Ankit Kumar Srivastava[1], Devender Singh[2], Ajay Shekhar Pandey[3]
[1]Assistant Professor of Electrical Engineering Department, IET, Dr. Rammanohar Lohia Avadh University, Ayodhya, India, aankit.srivastava32@gmail.com
[2] Professor of Electrical Engineering Department, IIT (BHU), Varanasi, India, dsingh.eee@itbhu.ac.in
[3] Professor of Electrical Engineering Department, KNIT, Sultanpur, India, ajayshekharpandey@gmail.com

ABSTRACT

Decision making in the energy market has to be based on accurate forecasts of the load demand. Therefore, Short Term Load Forecasting (STLF) is important tools in the energy market. In this paper, load forecasting using regression tree methods (Random Forest, Bagging and M5P) are used to effectively forecast the load. The usefulness of the proposed methods has been authenticated through extensive tests using real load data from the Australian electricity market. A comparison of these methods shows that there is an edge in M5P in relation to accuracy.

Key words: Bagging, Data mining, Load Forecasting, M5P, Random Forest, Regression Tree.

1. INTRODUCTION

The term short-term load forecasting (STLF) is a process to estimate the load over an interval ranging from hours to week. It plays a crucial role in expansion of the generation cost, dynamic state estimation, online scheduling and security function of energy management system (EMS) [1] and it covers the organized operation and planning of the power system. In short-term load forecasting, the frequent change of load performance is due to change in different aspects such as weather conditions [2], time factor, economic indicator, random disturbances etc. Due to the exceedingly nonlinear relations between load and other factors like temperature, humidity, wind speed, etc., non-linear techniques, for modeling as well as forecasting, plays key role in STLF.

In short-term load forecasting, a wide diversity of methodologies and procedures have been described in the literature [3}, [4]. These methodologies typically use algorithms which range from statistical domain [5], [6], [7] to Artificial Intelligence (AI) based approaches [8]. In last few decades, several efforts have been made to resolve load forecasting difficulties, by using various conventional statistics based strategies like time series models [9], weight translation method, similar day approach, AI and intelligent hybrid approaches. Several utilities and system operators use ANN based load forecasters [10], [11]. In current years, with the advancement of the artificial intelligence, researchers proposed tree based methods [12]. Regression Tree is a data-driven, multivariate, nonlinear and nonparametric technique. It also maps the relationship between input and output data algorithmically, they learn from their relationship and accumulate this learning into their constraints.

In this paper, Regression Tree STLF model's steps of designing, training and testing methods for power system operation are discussed. This methodology minimizes the modeling effort and provides high accuracy and which suggest significant benefits over the conventional approach [13], [14]. The proposed model works on learns by itself methodology. The regression tree method maps the relationship between the input variables such as previous load and weather data, and the output in a so called training process of the method. A trained regression tree makes an ideal predictor of the given the input data.

The aim of the present work is to compare different regression Trees model for STLF. Effectiveness of regression tree methods have been compared the M5P is seen to have a slight edge over the bagging and random forest based approaches.

2. SHORT TERM LOAD FORECASTING USING REGRESSION TREE

As of now one of the most generally utilized procedures for value forecasting is data mining. Data mining offers a method for breaking down data factually and concentrate to determine such principles that can be utilized for expectations. By and by it is being utilized in numerous spaces, for example, securities exchange, sports, banking segment, and so on. Researchers have now understood that data mining can be utilized as an apparatus for power load expectation also. The fundamental element of data mining is data itself. It is characterized as crude arrangement of data which can be utilized to extricate important data relying on the prerequisites of the application. Data can be put away in a composed way which is known as database. The term data mining alludes to the strategies that are utilized to separate the necessary data from the given arrangement of data that may be valuable for factual reason or making expectations by learning designs in data and relationship between various parameters. Researchers, mathematicians and specialists have thought of a wide scope of calculations for expectation issues. In specific cases, regression procedure ends up being progressively successful, though sometimes, rule based method and decision tree calculations give precise outcome with a low computational expense. This current work examines regression tree calculations and their exhibition in forecast of power load, Fig 1 shows the run of the mill structure of a decision tree.

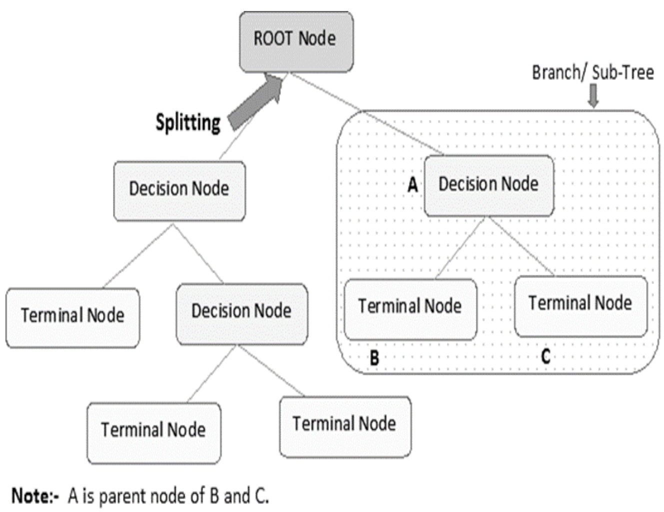

Figure 1: Typical structure of a decision tree

2.1 RANDOM FOREST

This strategy proposed by Leo Breiman is a gathering based technique which puts accentuation just on groups of decision trees. It depends on two procedures the CART (Classification and Regression Trees) and the bagging. A subspace is arbitrarily chosen from the element space. The RF is the arrangement of decision trees that develop input factors taken arbitrary to frame a hub. Every hub and the decision trees are made utilizing, the examples from the data utilizing a procedure called boot tying. Hub and decision trees are built utilizing bootstrap test from the learning test and these factors. The calculation can't to consider each factor accessible. It keeps trees from being fundamentally the same and in this manner, co-related forecasts are not created. Henceforth, the trees can be viewed as about autonomous, which lessens the danger of over-fitting and difference of expectation likewise diminishes. A ultimate choice is gotten by conglomerating over the group.

The strategy utilizes the arrangement of data, which were not utilized in bootstrap procedure to produce a blunder gauge or speculation mistake. The blunder on data, which is kept separate from pack (OOB) is utilized to give a proportion of constraining mistake. Progressive trees are included during preparing, until the OOB blunder balances out and halted when OOB mistake balances out.

2.1.1 Algorithm

1) Consider a bootstrap test of a specific size from the preparation data. (leave the size alone K)
2) Select 'K' arbitrary factors from the arrangement of factors.
3) From these 'K' factors, pick the best factor.
4) Split the above into two little girl hubs.
5) Recursively recurrent the means from 2 to 4 until a base hub size is accomplished. Along these lines an irregular backwoods tree Tn is acquired.
6) Compute the trees Tn; n=1, 2, 3,…….., N.
7) Predict the sample use:
8) $F(x)= (1/N) \sum n=1\ n=N\ Tn\ (x)$
9) $F(x)= (1/N) \sum n=1\ n=N\ Tn\ (x)$

2.1.2 Characteristics

1) Used for classification just as regression issues.
2) Very basic model and has exceptionally vigorous methodology.
3) Requires just barely any parameters to gauge the outcome.
4) It can manage the missing qualities.
5) It is quicker than different strategies and has high steadiness.
6) It has more precision as it abstains from over-fitting of data.
7) It is amazingly powerful in taking care of the enormous datasets.
8) It gives resistance to clamor as it creates uncorrelated trees

2.2 RANDOM FOREST

The Bootstrap technique is foundation of bagging. For a sample of 1000 values of (x) we, to get an estimate of the mean of the sample, Mean of the sample can be calculated directly from the sample as:

$$Mean(x) = 1/1000 * sum\ (x) \tag{1}$$

Here the sample size is small and that our calculated mean could have error. Now we can improve our estimate of error using bootstrap procedure.

1) Make sub-samples of the data randomly.
2) Calculate mean of each sub-sample.

3) Calculate the average of mean of all the sub-samples to get means of database.

2.2.1 Bootstrap Aggregation (Bagging)

Bagging applies bootstrapping technique to a data having large variance. The bootstrapping method is applied on machine learning algorithm. For a sample dataset of 2000 instances (x), Bagging algorithm works in following manner.

1) Apply random sub samples of database with replacement.
2) Apply learning procedure on each sample to create the model.
3) For a new input set, the predicted value of each model and used the mean as predicted value.

2.2.2 Bagging with Decision Trees

In the case of bagging with decision trees, the individual tree over fitting of training data has not been taken in to account. For this reason, as well as for efficiency individual tree grown deep (few training set of each node) or more classification done of child node or it allows to grow tree more deeper, trees are not pruned.

The decision trees obtained by pruning have small bias and have high variance. The only parameters when bagging decision trees is the number of samples, this can be chosen from increasing number of sample trees on each iteration until the results stopped improving further.
Large number of models may take a long time to execute but never create a problem of over fitting.

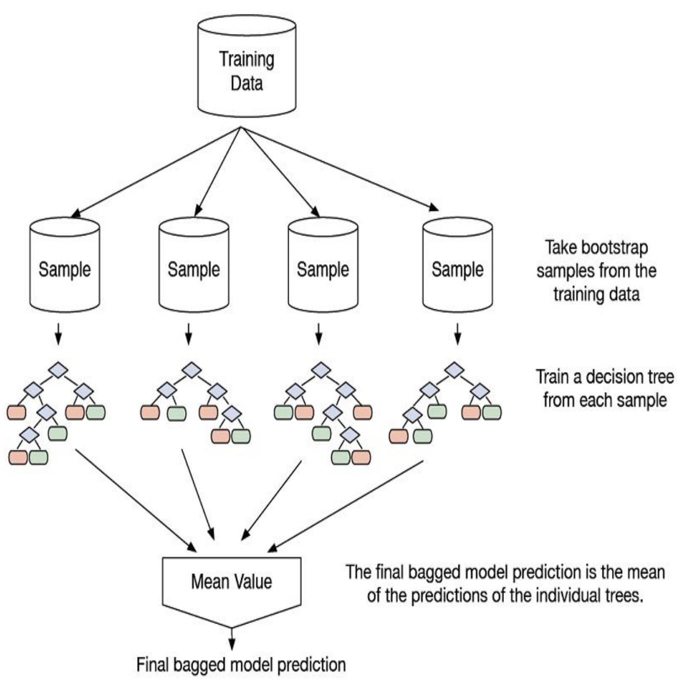

Figure 2: Bagging Model

2.2.3 Bagging Algorithm:

Input: - The historical data is containing feature values along with the value corresponding load. Let the data D containing kc classes.

1) Take a bootstrapped replica Dm by randomly drawing from.
2) Call Learn Decision Tree with Dm and receive a single decision tree DTm.
3) Add DTm to the ensemble, E.

2.3 M5P

M5P has been assessed on a few learning undertakings for which consequences of different techniques are additionally accessible. When looking at approaches, a typical proportion of execution is relative blunder, the proportion of the fluctuation of the residuals to the difference of the objective qualities themselves. Another helpful measurement is the connection among real and anticipated qualities (a connection coefficient of 1 shows just that there is a direct connection among genuine and anticipated qualities). A few creators report rate deviation, the normal over the instances of the proportion of the leftover to the objective worth.

In many preliminaries, M5P's exhibition was surveyed in a 10-way cross-approval in which the accessible data was separated into ten equivalent estimated squares. For each square thus, a model was built utilizing just cases in the staying nine squares, at that point tried on cases in the \hold-out" square. Each case was in this way tried once utilizing a model built without reference to the case. All m5p results revealed beneath were acquired utilizing similar default estimations of parameters. To show the impacts of shaping models at the leaves and of smoothing, results are additionally given with these highlights debilitated. At the point when we construct engineering with the goal that the M5P calculation remains mindful to online produce the assessment work, we make an operator fit for adjusting to dynamic and out of reach encompassing. now, conceivable engineering exploits in examination with conventional techniques like the hereditary calculation, taking into account that in spite of the fact that the last doesn't require information got from the issue, it requires a pre-characterized assessment strategy for the outcome. Neural networks (NNs) are not as straightforward as semi-experimental regression-based models. In contrast to neural networks, the tree acquired are aftereffect of reasonable standards. Likewise the model, calculation needn't bother with a few experimental parameters tree to be introduced, for example, number of concealed layers, number of neurons and other learning parameters.

3. RESULT AND DISCUSSION

The historical load data of New South Wales, Australia (taken from AEMO) and weather data of Sydney City (www.weatherzone.com/au.) has been taken half hourly from January 2014 to June 2016 for the forecasting. Wind speed, temperature, and humidity are considered under weather data in the present study. Table-1 show the list of input variables affecting the half-hourly predicted load.

A set of input variables consisting of 25 features along with training set containing 1965 dataset has been used. The concept of similar weeks is taken into account for the training set. The data set corresponded to the five similar weeks of the months of the previous year and the preceding week the same year i.e. if forecasting is to be done for 8-14 Nov 2014, training set will include data corresponds to 1-7 Nov 2014, 25-31 Oct 2013, 1-7 Nov 2013, 8-14 Nov 2013, 15-21 Nov 2013 and 22-28 Nov 2013. Stratified 10 FCV classification methodology has been used for the classifier, making it the whole of the data tested at least once. The forecast error for Regression Tree i.e. Random Forest (RF), Bagging and M5P is computed for whole week in terms of MAPE (Mean Absolute Percentage Error). The MAPE for three seasons are calculated using Regression Trees are compared with each other and are shown in Table 2.

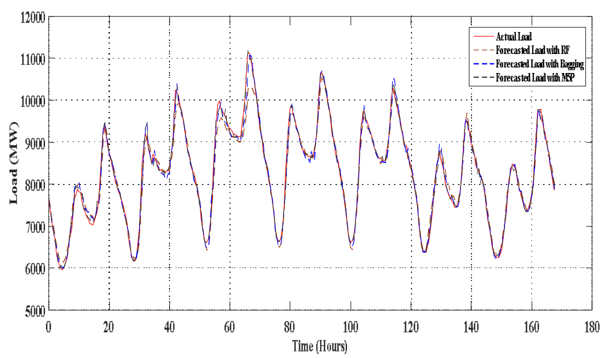

Figure 3: June 08-14, 2015 Forecasting with Regression Trees Method in Winter Season

Fig. 3 shows the forecasting result of Second week of June 2015 for RF, Bagging & M5P. It observed from the result that bagging provide lesser MAPE 1.04 than RF 1.64 and Bagging 1.19 respectively.

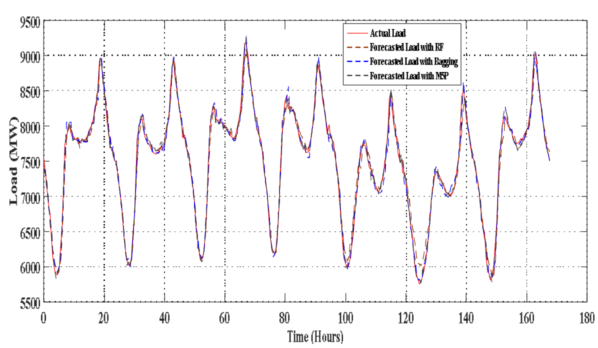

Figure 4: Sep 15-21, 2015 Forecasting with Regression Trees Method in Spring Season

Fig 4 shows the forecasting result of Third week of Sep 2015 for RF, Bagging & M5P. It observed from the result that bagging provide lesser MAPE 0.65 than RF 1.13 and Bagging 0.86 respectively.

Fig. 5 shows the forecasting result of Forth week of January 2016 for RF, Bagging & M5P. It observed from the result that bagging provide lesser MAPE 0.75 than RF 1.50 and Bagging 0.84 respectively.

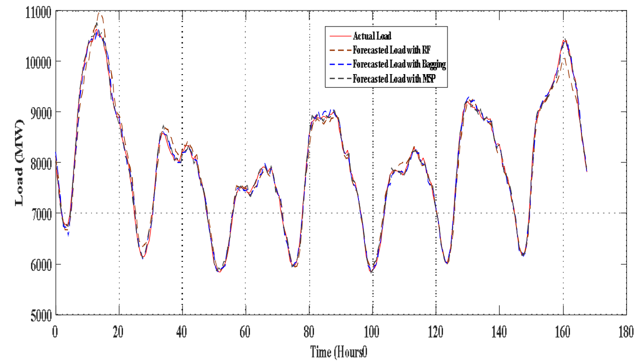

Figure 5: Jan 22-28, 2016 Forecasting with Regression Trees Method in Summer Season

Table 1: Variables Affecting the Half-Hourly Predicted Load

Variable	Variable Timing	Variable Name
Load (Lo)	$Lo_{(t-24:00-01:00)}$	Lo_6
	$Lo_{(t-24:00-00:30)}$	Lo_5
	$Lo_{(t-24:00-00:00)}$	Lo_4
	$Lo_{(t-01:30)}$	Lo_3
	$Lo_{(t-01:00)}$	Lo_2
	$Lo_{(t-00.30)}$	Lo_1
Wind Speed (Wi)	$Wi_{(t-24:00-01:00)}$	Wi_6
	$Wi_{(t-24:00-00:30)}$	Wi_5
	$Wi_{(t-24:00-00:00)}$	Wi_4
	$Wi_{(t-01:30)}$	Wi_3
	$Wi_{(t-01:00)}$	Wi_2
	$Wi_{(t-00.30)}$	Wi_1
Temp (Te)	$Te_{(t-24:00-01:00)}$	Te_6
	$Te_{(t-24:00-00:30)}$	Te_5
	$Te_{(t-24:00-00:00)}$	Te_4
	$Te_{(t-01:30)}$	Te_3
	$Te_{(t-01:00)}$	Te_2
	$Te_{(t-00.30)}$	Te_1
Humidity (Hu)	$Hu_{(t-24:00-01:00)}$	Hu_6
	$Hu_{(t-24:00-00:30)}$	Hu_5
	$Hu_{(t-24:00-00:00)}$	Hu_4
	$Hu_{(t-01:30)}$	Hu_3
	$Hu_{(t-01:00)}$	Hu_2
	$Hu_{(t-00.30)}$	Hu_1
Half Hourly Timing (Ho)	$Ho_{(t-00.00)}$	H_o

4. CONCLUSION

In this paper, regression tree based methods (Random Forest, Bagging and M5P) are presented for short term load forecasting (STLF). The (MAPE) has been calculated for all the seasons. Forecasts for a week in all seasons were carried out to test the accuracy of these three methods. Based on the experimental results, it is observed that M5P method provides better forecast accuracy and outperforms other regression tree methods.

Table 2: Mean Absolute Percentage Error (MAPE) for all Season

Sr. No.	Season	Month	First			Second			Third			Forth		
			RF	Bagging	M5P	RF	Bagging	M5P	RF	Bagging	M5P	RF	Bagging	M5P
1	Winter	June	3.12	2.1	1.36	1.64	1.19	1.04	1.62	1.19	0.90	1.26	0.91	0.82
2	Spring	Sep	1.45	0.98	0.90	1.38	1.03	0.95	1.13	0.86	0.65	2.05	1.49	1.10
3	Summer	Jan	1.13	0.87	0.77	2.83	1.64	0.83	2.28	1.46	1.05	1.50	0.84	0.75

REFERENCES

1. G. Gross, F. D. Galiana. **Short-term load forecasting**, *Proceeding of the IEEE*, Vol.75, No.12, pp.1558, 1573, Dec. 1987. https://doi.org/10.1109/PROC.1987.13927
2. J. W. Taylor, and R. Buizza. **Neural network load forecasting with weather ensemble predictions**, *IEEE Transactions on Power Systems*, Vol. 17, pp. 626-632, 2002.
3. C. Cecati, J. Kolbusz, P. Różycki, P. Siano and B. M. Wilamowski. **A Novel RBF Training Algorithm for Short-Term Electric Load Forecasting and Comparative Studies**, *IEEE Transactions on Industrial Electronics*, Vol. 62, No. 10, pp. 6519-6529, Oct. 2015. https://doi.org/10.1109/TIE.2015.2424399
4. I. Mogram and S. Rahman. **Analysis and evaluation of five short-term load forecast techniques**, *IEEE Trans. On Power Systems*. Vol.4, No.4, pp 1484-1491, 1989. https://doi.org/10.1109/59.41700
5. J. Y. Fan, Member, J. D. McDonald, Senior Member. **A real-time implementation of short-term load forecasting for distributed power systems**, *IEEE Transactions on Power Systems*, Vol. 9, No.2, March 1994.
6. Q. Mu, Y. Wu, X. Pan, L. Huang and X. Li. **Short-term Load Forecasting Using Improved Similar Days Method**, *Asia-Pacific Power and Energy Engineering Conference, Chengdu*, 2010, pp. 1-4.
7. A. D. Papalexopoulos, T.C Hesterberg. **A regression-based approach to short-term system load forecasting**, *IEEE Transactions on Power Systems*, Vol. 5, No.4, pp.1535,1547, Nov 1990. https://doi.org/10.1109/59.99410
8. Q. W. Luthuli and K. A. Folly. **Short term load forecasting using artificial intelligence**, *2016 IEEE PES PowerAfrica, Livingstone*, 2016, pp. 129-133.
9. M. Chaouch. **Clustering-Based Improvement of Nonparametric Functional Time Series Forecasting: Application to Intra-Day Household-Level Load Curves**, *IEEE Transactions on Smart Grid*, Vol. 5, No. 1, pp. 411-419, Jan. 2014.
10. Manish Kumar Singla, Jyoti Gupta, Parag Nijhawan, Amandeep Singh Oberoi. **Electrical Load Forecasting Using Machine Learning,** *International Journal of Advanced Trends in Computer Science and Engineering*, Vol. 8, No. 3, pp. 615-619, 2019. https://doi.org/10.30534/ijatcse/2019/45832019
11. Lorwin Felimar Torrizo, Aaron Don Africa. **Next-Hour Electrical Load Forecasting using an Artificial Neural Network: Applicability in the Philippines,** *International Journal of Advanced Trends in Computer Science and Engineering*, Vol. 8, No. 3, pp. 831-835, 2019. https://doi.org/10.35940/ijitee.I3174.0789S319
12. K. Nose-Filho, A. D. P. Lotufo and C. R. Minussi. **Short-Term Multinodal Load Forecasting Using a Modified General Regression Neural Network**, *IEEE Transactions on Power Delivery*, vol. 26, no. 4, pp. 2862-2869, Oct. 2011.
13. J.W. Taylor, L.M. de Menezes, and P.E. McSharry. **A comparison of univariate methods for forecasting electricity demand up to a day ahead**, *International Journal of Forecasting*, vol. 22, pp. 1-16, 2006. https://doi.org/10.1016/j.ijforecast.2005.06.006
14. M. R. Haq and Z. Ni. **A New Hybrid Model for Short-Term Electricity Load Forecasting**, *IEEE Access*, Vol. 7, pp. 125413-125423, 2019. https://doi.org/10.1109/ACCESS.2019.2937222