Available online at: https://ijact.in

Date of Submission	30/11/2018
Date of Acceptance	21/01/2019
Date of Publication	07/02/2019
Page numbers	2985-2989 (5 Pages)

ISSN:2320-0790

COMPUSOFT
An International Journal of Advanced Computer Technology

SHORT TERM PRICE FORECASTING USING TREE BASED METHODS

A K Srivastava[1], D Singh[2], A S Pandey[3], Sanjay Kumar[4]

[1] EED, IET, Dr. Rammanolar Lohiya Avadh University, Faizabad, (U.P.), India

[2, 4] EED, Indian Institute of Technology (BHU), Varanasi (U.P.), India
[3] EED, Kamla Nehru Institute of Technology, Sultnapur (U.P.), India

Abstract: In this paper, electricity price forecasting using J48, Random forest and Bagging are used to effectively forecast the electricity price. These models are obtained by recursively partitioning the data space and fitting a simple prediction model within each partition. The effectiveness of the proposed methods has been validated through comprehensive tests using real price data from Australian electricity market. The comparison of these methods shows that the bagging is having an edge as the accuracy is concerned.

Keywords: Price forecasting; Data mining toolbox; weka; J48; Random forest; Bagging

I. INTRODUCTION

As globally electricity market is being de-regulated, now Generators, Distributors and Consumers have to be ready for even small and immediate change in electricity load and supply. Now a day's generators as well as consumers are free to choose, to buy and sell the electricity as per their choice. In this scenario electricity price forecasting becomes topic of great interest. Every market participant needs to know the accurate electricity price for each load block to achieve maximum profitability. If the electricity market price can be predicted properly, the generating companies and large scale enterprises as main market participating and deciding entities can reduce their risks and maximize their outcomes further.

Electricity price forecasting has now become a need of de-regulated electricity power market. The accurate and efficient price forecasting at right time and in convenient way is desire of every market player. Electricity price forecasting, unlike load forecasting, is much more complex because of the unique characteristics and uncertainties in operation as well as bidding strategies. In the current deregulated scenario, the forecasting of electricity demand and price has emerged as one of the major research fields in electrical engineering [1]. Researchers and academicians are engaged in the activity of developing tools and algorithms for load and price forecasting.

Electricity price forecasting is very different from electricity load forecasting as it has characteristic of high volatility, very fast change of frequency of peak price, high dependency upon seasons and nature of non-storability. These characteristics arise from various reason which is due to electricity price dependency upon various factors like environment factors, market factors and past data of price and load itself. The various factors that affect the electricity price are shown in figure 1[2].

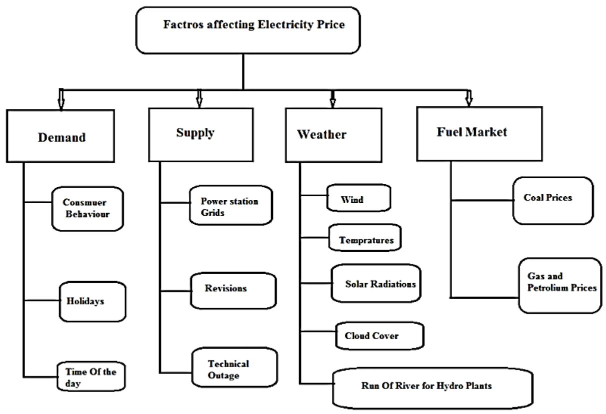

Fig 1: Factors affecting Electricity Price [2]

In the 1980s, statisticians Breiman et al (1984) [3] developed CART (Classification and Regression Trees), which is a sophisticated program for fitting trees to data. Since the original version, CART has been improved and given new features, and it is now produced, sold, and documented by Salford Systems. A real-time pricing type scenario is envisioned in reference [4] where energy prices could change on an hourly basis with the consumer having the ability to react to the price signal through shifting his electricity usage from expensive hours to other times when possible.

M. K. Kim et al [5] proposed a new forecasting method for short-term spot prices in the Nordic power market. It proposes a Cuckoo search Levenberg-Marquardt (CSLM)-trained, CSLM feed-forward neural network (CSLM-FFNN) for the solving process that combines the improved Levenberg Marquardt and Cuckoo search algorithms. The proposed model considers actual power generation and system load as input sets to facilitate the efficient use of both transmission and power generation resources by direct market participants. P. Sarikprueck et. al. [6] proposed a hybrid method for very short term market price forecasting to improve prediction accuracy on both nonspike and spike wholesale market prices. First, support vector classification is carried out to predict spike price occurrence, and support vector regression is used to forecast the magnitude for both non-spike and spike market prices. Additionally, three clustering techniques including classification and regression trees, K-means, and stratification methods are introduced to mitigate high error spike magnitude estimation. E. E. Elattar et. al. [7] presented a new approach to short-term electricity price forecasting. The proposed method is derived by integrating the kernel principal component analysis (KPCA) method with the local informative vector machine (IVM), which can be derived by combining the IVM with the local regression method.

C. González et al (2015) [8] proposed, a regression tree-based models like Bagging and Random Forests to identify the variables dominating the marginal price of the commodity as well as for short-term (one hour and day ahead) electricity price forecasting for the Spanish-Iberian market. They highlighted the effectiveness of the proposed ensemble of tree-based models which emerge as an alternative and promising tool, competitive with other existing methods.

The present work aims to compare the different models of tree for price forecasting. The effectiveness of J48, Random Forest and Bagging methods have been compared. It is observed that bagging has a slightly edge over J48 and random forest based forecasting.

II. ELECTRICITY PRICE FORECASTING WITH TREES BASED METHODS

The electric power price forecasting problem is not easy to handle due to nonlinear and random-like behaviors of system loads, weather conditions, and variations of social and economic environments, etc. A good amount of research has already gone in this area. However, linear regression is a good model, where there is a single predictive formula holding over the entire data-space. When the data has lots of features which interact in complicated, nonlinear ways, assembling a single global model can be very difficult and hopelessly confusing when you do succeed.

An alternative approach to nonlinear regression is to sub-divide or partition the space into smaller regions, where the interactions are more manageable. We then partition the sub-divisions again and this is called recursive partitioning. Until finally we get to chunks of the space which are so tame that we can fit simple models to them.

The advantage of tree methods is that they are able to explore and highlight complex or hidden relationships in the data, for this reason they are used for getting well-fitted models that represent very accurately the data behavior and are particularly useful for prediction.

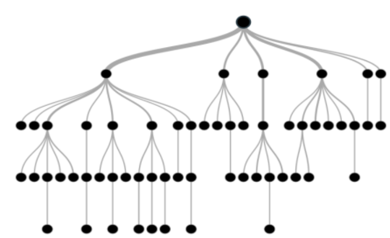

Fig 2: A model of decision tree

[A] J48 Method: The C4.5 algorithms for building decision trees is implemented in Weka as a classifier called J48. Classifiers, like filters, are organized in a hierarchy: J48 has the full name weka.classifiers.trees.J48. 48 is an extension of ID3. J48 are the improved versions of C4.5 algorithms or can be called as optimized implementation of the C4.5. The output of J48 is the Decision tree. If we have a list of dependent (target) and independent variables (predictors), then by applying a decision tree like J48 we can predict the target variable of a new dataset record. The attribute which is to be predicted is known as dependent variable and the other attributes which help in predicating it are known as independent variables in the dataset. Its objective is to progressively generalize a decision tree until it gains equilibrium of flexibility and accuracy.

The additional features of J48 are accounting for missing values, decision trees pruning, continuous attribute value ranges, derivation of rules, etc. In the WEKA data mining tool, J48 is an open source Java implementation of the C4.5 algorithms.

[B] Random Forests: RF is more recent than the other techniques. It was developed by Breiman [9] as a way of obtaining more accurate predictions without overfitting the data. A random forest is a collection or ensemble of decision trees that is built using the whole dataset considering all features, but in random forests a fraction of the number of rows is selected at random and a particular number of features are selected at random to train on and a decision tree is built on this subset. In Random Forests, a different subset of training data is selected, with replacement to train each tree. Class assignment is made by the number of votes from all the trees and for regression the average of the results is used.
The basic difference is that Random Forest (RF) is a collection or ensemble model of numerous Decision Trees (DT).
The reasons for selecting Random Forest over decision tree are listed below:

1. The single DT would lead to over-fit model if the dataset is huge, the same way like a single person might have its own perspective on the complete population.
2. However, if we implement the voting system and ask different individuals to interpret the data then we would be able to cover the patterns in a much meticulous way.

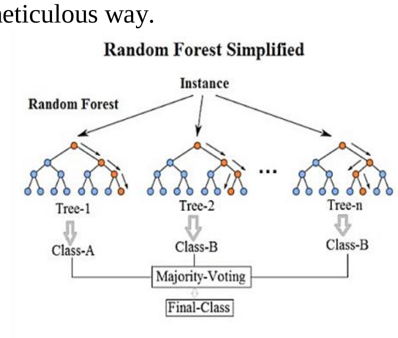

Fig 3: Random forest Functioning

The Functioning of Random Forest is as follows:
1. In Random Forest, N number of tree is grown if the number of cases in the training set is N and the sample N case is at random.
2. This Sample set will work as training set for growing of the tree. If there are M input variables in the Sample set, then m random variables out of the M will be chosen such that m<<M at any node.
3. The value of m is held constant during the forest growing and each tree grows to the largest extent possible.
4. Each tree is grown to the largest extent possible and there is no pruning.
5. Predict new data by aggregating the predictions of the n tree trees (i.e., majority votes for classification, average for regression.

[C] Bagging: Bagging is the abbreviation of "Bootstrap Aggregation"; it was introduced by Bradley Efron in 1979. Bagging is well known methods for estimating standard errors, bias, and constructing confidence intervals for parameters. It consists on making use of several single trees, each of them constructed with a different, randomly chosen sample from the original group, using the rest of the data to validate and predict. Each sample contains the same number of data as the initial set but differs from it in that repetitive data may exist. This is the reason why this technique is known a bootstrapping: "resampling with replacement".
Working principle of bagging: -
1. It takes original data set D with N training example.
2. Then it creates M subset form the N training example where M<N.
3. Each Data set D_M is generated from D by sampling with replacement but each data set D_M has the same number of examples as in data set D.
4. These data sets are reasonably different from each other (since only about 63% of the original examples appear in any of these data sets)
5. Train models $C_1, C_2, \ldots, C_M$ using $D_1, D_2, \ldots, D_M$.
6. The predictions of all the classifiers are combined using a mean, median or mode value depending on the problem at hand.
7. Useful for models with high variance and noisy data

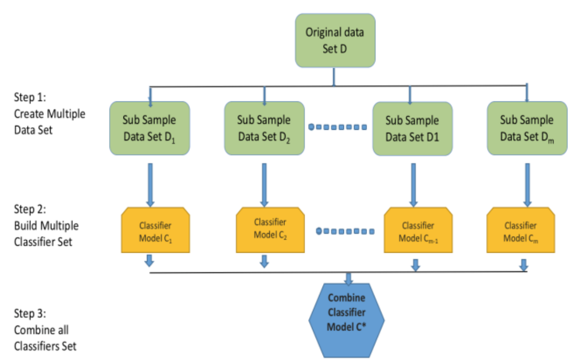

Fig 4: Steps involved in Bagging

Bagging is a technique used to reduce the variance of our predictions by combining the result of multiple classifiers modelled on different sub-samples of the same data set. The most notable benefits of bagging is:

1. It has surprisingly competitive performance & rarely over fits.
2. It is capable of reducing variance of constituent models

III. RESULT AND DISCUSSION

For short term price forecasting data is collected from Australian Energy Market Operator for New South Wales, Australia. The data consists of half hourly load and price of all seasons from January 2014 to June 2016. The weather data of Sydney City is taken from www.weatherzone.com/au. Weather data considered in the present study are half hourly wind speed, temperature, and humidity. All the data were quantized at 40 levels, each level consisting of 2.5 percent of the range. Thus, for a particular week, all the data have been classified to have only 40 discrete values. Table 1 show the list of features which are assumed to affect half-hourly electricity prices.

Table-1 List of input variables for electricity price forecasting

Variable	Variable Timing	Feature Name
Load (L)	$L_{(T-23:00)}$	L1
	$L_{(T-23:30)}$	L2
	$L_{(T-24:00)}$	L3
	$L_{(T-01:30)}$	L4
	$L_{(T-01:00)}$	L5
	$L_{(T-00.30)}$	L6
Price (P)	$P_{(T-23:00)}$	P1
	$P_{(T-23:30)}$	P2
	$P_{(T-24:00)}$	P3
	$P_{(T-01:30)}$	P4
	$P_{(T-01:00)}$	P5
	$P_{(T-00.30)}$	P6
Wind Speed (W)	$W_{(T-23:00)}$	W1
	$W_{(T-23:30)}$	W2
	$W_{(T-24:00)}$	W3
	$W_{(T-01:30)}$	W4
	$W_{(T-01:00)}$	W5
	$W_{(T-00.30)}$	W6
Temperature (T)	$T_{(T-23:00)}$	T1
	$T_{(T-23:30)}$	T2
	$T_{(T-24:00)}$	T3
	$T_{(T-01:30)}$	T4
	$T_{T-01:00}$	T5
	$T_{T-00.30}$	T6
Humidity (H)	$H_{T-23:00}$	H1
	$H_{T-23:30}$	H2
	$H_{T-24:00}$	H3
	$H_{T-01:30}$	H4
	$H_{T-01:00}$	H5
	$H_{T-00.30}$	H6
Day Timing (Ho)	Ho_T	Ho

The set of input feature thus consisted of 31 features. The training set consisted of 2016 data sets. The training set was taken on the concept of analogous weeks. The data set corresponded to the five analogous weeks of the months of the previous year and the preceding week the same year. For example, if the forecasting is to be performed for the week of 1-7 Aug 2015, the training set would include the data corresponding to 24-30 July 2015, 1-7 Aug 2014, 24-30 July 2014, 17-23 July 2014, 8-14 Aug 2014, 15-21 Aug 2014. The classifier uses stratified 10 fold cross-validation classification accuracy methodologies. Thus, whole of the data is tested in this method at least once.

Forecast accuracy after creating training set and testing set data, the daily forecast using J48 and Random Forest (RF) and Bagging is computed. MAPE (Mean Absolute Percentage Error) is calculated for the whole Week. The mean absolute percentage errors for three seasons are calculated using J48, RF and Bagging and compare with each other are shown in Table 2.

When MAPE is compared season wise it is observed that the Bagging provides lesser MAPE 10.1 than the J48 12.17 and RF 10.74 respectively for Spring Season. It is found for all three seasons.

Fig 5 shows the forecasting result of First week of July 2015 for J48, Random Forest and Bagging it observed from the result that bagging provide lesser MAPE 7.39 than RF 8.51 and J48 9.06 respectively.

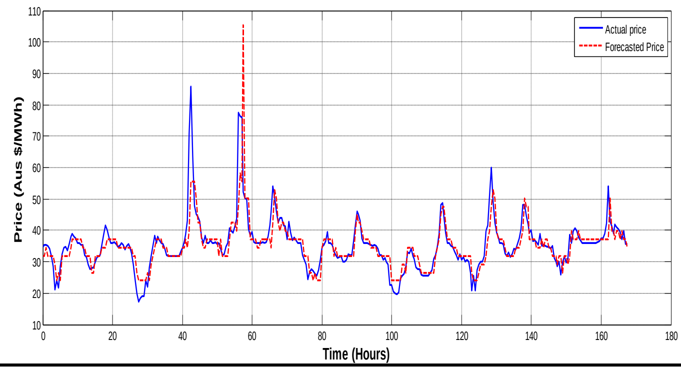

Fig. 5 July 01-07, 2015 Forecasting with Bagging in Spring Season

Fig 6 shows the forecasting result of Third week of October 2015 for J48, Random Forest and Bagging. It observed from the result that bagging provide lesser MAPE 7.95 than RF 8.46 and J48 9.76 respectively.

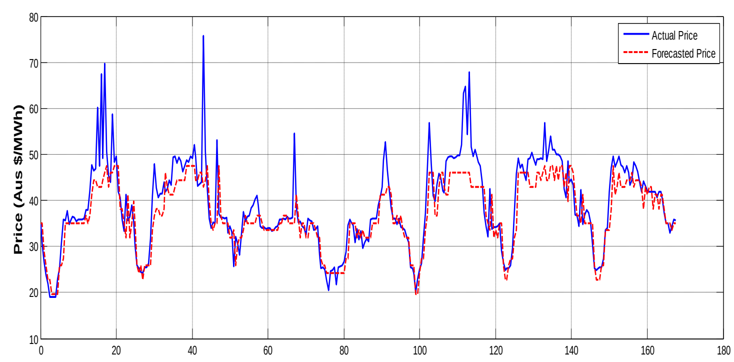

Fig. 6 October 15-21, 2015 Forecasting with Bagging in Winter Season

Table: 2 Mean Absolute Percentage Error (MAPE) Season Wise

Season	Month	First (1-7)			Second (8-14)			Third (15-21)			Forth (22-28)					
		J48	RF	Bagging	J48	RF	Bagging	J48	RF	Bagging	J48	RF	Bagging	J48	RF	Bagging
Winter	Jul-15	9.06	8.51	7.39	11.72	10.35	9.4	15.64	14.15	13.81	12.27	9.96	9.81	12.17	10.74	10.10
Spring	Oct-15	11.64	9.78	8.31	8.91	11.28	8.76	9.76	8.46	7.95	9.3	9.48	7.67	9.90	9.75	8.17
Summer	Jan-16	7.9	7.35	6.92	17.49	15.28	17.59	13.95	11.89	11.97	10.42	8.53	8.43	12.44	10.76	11.22

Fig 7 shows the forecasting result of Forth week of January 2016 for J48, Random Forest and Bagging. It observed from the result that bagging provide lower MAPE 8.43 than RF 8.53 and J48 10.42 respectively.

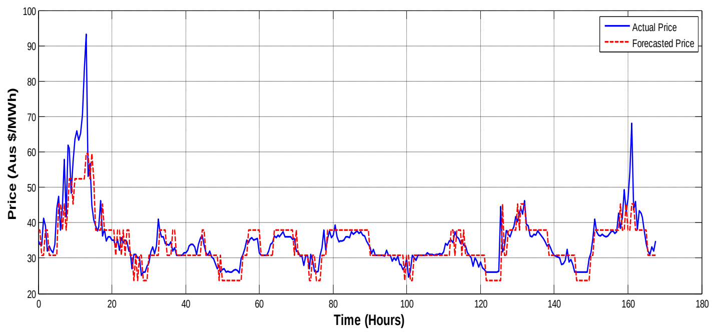

Fig. 7 January 22-28, 2016 Forecasting with Bagging in Summer Season

IV. Conclusion

In this paper, tree based methods J48, Random Forest and Bagging are proposed for predicting the electricity prices. The Mean Absolute Percentage Error (MAPE) has been calculated day season wise. Forecasts for a week in three seasons were carried out to test the accuracy of the three methods. Based on the results, it is observed from the result that Bagging method provides better forecast accuracy than J48 and Random Forest.

V. References

[1] D. W. Bunn, “Forecasting loads and prices in competitive power markets,” in Proceedings of the IEEE, vol. 88, no. 2, pp. 163-169, Feb. 2000.

[2] Singh, Nitin & R. Mohanty, S. (2015). A Review of Price Forecasting Problem and Techniques in Deregulated Electricity Markets. Journal of Power and Energy Engineering, Vol. 03, pp 1-19, 2015

[3] Breimann, L., Friedman, J.H., Olshem, J.H., Stone, C.J.: ‘Classification and regression trees’ (Chapman and Hall/CRC, 1984)

[4] A. Khotanzad, E. Zhou and H. Elragal, “A Neuro-Fuzzy Approach to Short-Term Load Forecasting in a Price-Sensitive Environment”, IEEE Transactions on Power Systems, Vol. 17, No. 4, pp. 1273-1282, November 2002,

[5] M. K. Kim, "Short-term price forecasting of Nordic power market by combination Levenberg–Marquardt and Cuckoo search algorithms," in *IET Generation, Transmission & Distribution*, vol. 9, no. 13, pp. 1553-1563, 1 10 2015.

[6] P. Sarikprueck, W. Lee, A. Kulvanitchaiyanunt, V. C. P. Chen and J. Rosenberger, "Novel Hybrid Market Price Forecasting Method With Data Clustering Techniques for EV Charging Station Application," in *IEEE Transactions on Industry Applications*, vol. 51, no. 3, pp. 1987-1996, May-June 2015.

[7] E. E. Elattar, "Day-ahead price forecasting of electricity markets based on local informative vector machine," in *IET Generation, Transmission & Distribution*, vol. 7, no. 10, pp. 1063-1071, October 2013.

[8] C. González, J. Mira-McWilliams and I. Juárez, "Important variable assessment and electricity price forecasting based on regression tree models: classification and regression trees, Bagging and Random Forests," in IET Generation, Transmission & Distribution, vol. 9, no. 11, pp. 1120-1128, 6 8 2015.

[9] Hong-Tzer Yang and Chao-Ming Huang, “A New Short-Term Load Forecasting Approach Using Self-organizing Fuzzy ARMAX Models”, IEEE Transactions on Power Systems, Vol. 13, No. 1, February 1998, pp. 217-225.