I J C T A, 9(37) 2016, pp. 263-267

Classification of Optimal and Safe Operating Periods of Wind Turbine Using Decision Tree

B. Priyadharshini[*] and Velappa Ganapathy[**]

Abstract: The electric power production in the world is predominantly fossil fuel based. Fossil fuels are non-renewable, that is, they draw on finite resources. In addition, they contribute to the production of greenhouse gases and particulates. In contrast, renewable energy resources, such as wind, solar, ocean, biomass, hydro, etc., can be replenished at a generally predictable rate and have no direct greenhouse gas or particulate emissions. This study aims in identifying suitable time to operate the wind turbines at various districts of Tamil Nadu state. The historical data is fed as input to the WEKA tool. The tool uses an algorithm named J48 and a decision tree is generated. The output generated is a multiclass classification. The output categorizes the stochastic nature of wind and gives information to the beneficiaries about the favorable periods of operating wind turbines.

Keywords: Wind parameters, Classification, Decision tree, Weka tool.

1. INTRODUCTION

In wind energy forecasting, it is essential to know the speed at which the turbine starts rotating and produces electricity. Too little wind cannot deliver sufficient power. Too much wind is susceptible to damage. Therefore it is essential to predict the safe operating periods of wind turbine generators so that precautionary measures can be taken to avoid any damage to the rotor and the blades. Also the operation of wind turbines at low speeds can be prevented. Where load shedding is common in states like Tamil Nadu, it is essential in determining the optimal operating periods of wind power generators.

Classification is a data mining task of predicting the value of a categorical variable by building a model based on one or more categorical variables. Decision tree builds classification model in the form of a tree structure. C4.5 is an algorithm used to generate a decision tree developed by Ross Quinlan. C4.5 is an extension of Quinlan's earlier ID3 algorithm. J48 is an open source Java implementation of the C4.5 algorithm in the Weka data mining tool. Weka (Waikato Environment for Knowledge Analysis) is a popular suite of machine learning software written in Java, developed at the University of Waikato, New Zealand. Weka contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization.

A confusion matrix contains information about actual and predicted classifications done by a classification system. Performance of such systems is commonly evaluated using the data in the matrix. The following table shows the confusion matrix for a two class classifier. The entries in the confusion matrix have the following meaning in the context of our study:

A is the number of correct predictions that an instance is negative, b is the number of incorrect predictions that an instance is positive, c is the number of incorrect of predictions that an instance is negative, and d is the number of correct predictions that an instance is positive.

[*] School of Computing, SRM University, Chennai, India. *Email: priyadharshini.b@ktr.srmuniv.ac.in*
[**] School of Computing, SRM University, Chennai, India. *Email: ganapathy.v@ktr.srmuniv.ac.in*

		Predicted	
		Negative	*Positive*
Actual	*Negative*	a	B
	Positive	c	D

All correct guesses are located in the diagonal of the table, so it's easy to visually inspect the table for errors, as they will be represented by any non-zero values outside the diagonal.

2. RELATED WORKS

In [1] the authors compare various forecasting approaches using time series analysis. Apart from the traditional model like ARMA, feed forward and recurrent neural networks, other approaches like ANFIS and Neural logic networks are also included. The authors [2] have trained the network in autoregressive manner using back propagation and cascade correlation algorithm. In [3], the model is trained using a genetic algorithm based learning scheme. In [4] the authors prove that the recurrent models outperform the static ones while they exhibit significant improvement over the persistent method. In [5] a self-organized map is trained to classify the forecasted local wind speed provided by the meteorological services. In [6] the authors have used Grey model, Grey predictor model GM (1,1) and a modified version of GM (1,1) known as adaptive alpha GM (1,1) for wind speed and wind power prediction. This paper [7] presents a statistical approach based on k-means clustering technique to manage environmental sampled data to evaluate and forecast of the energy deliverables by different renewable sources in a given site. In order to do wind speed forecasting hybrid models consisting of ARIMA and ANN models were developed [8].This paper [9] presents a robust two step methodology for accurate wind speed forecasting based on Bayesian combination algorithm and three neural network models such as ADALINE, BP network and RBF network. This article [10] presents adaptive Bayesian learning and Gaussian process approximation for wind power prediction. The Multiple Architecture System [11] is implemented by associating the predictions obtained from the different regression algorithms (MLR, MLP, RBF and SVM) making up the ensemble by three fusion strategies (simple, weighted and non-linear).In[12], a modified EMD-FNN model (empirical mode decomposition (EMD) based feed-forward neural network (FNN) ensemble learning paradigm) is proposed for wind speed forecasting. The forecasting engine [13] includes a new enhanced particle swarm optimization component and a hybrid neural network. A statistical-based wind power forecasting without using numerical weather prediction (NWP) inputs is carried out in this work [14] using adaptive wavelet neural network (AWNN) and a feed-forward neural network (FFNN).This paper[15] presents a method to improve the short term wind power prediction at a given turbine using information from numerical weather prediction and from multiple observation points.

3. PROPOSED WORK

The data was downloaded from Agricultural Meteorology Division, Government of India. The following parameters were given as input for a particular month such as Name of the District, Day, Rainfall, Maximum temperature, Minimum temperature, Total cloud cover, Maximum relative humidity, Minimum relative humidity, Wind speed and Wind direction.

The output has three classes such as 'very less', 'less' and 'sufficient'. 'Very less' (<8 kmph) tends to be the class where turbines cannot operate. 'Less' (8-12 kmph) tends to be the class where small wind turbines starts to rotate. 'Sufficient' (>13 kmph) tends to be the class where a turbine starts generating power. The tool takes data from the training set and a decision tree as given in Figure 1 is generated. This tree is converted to a set of rules by mapping from the root node to the leaf nodes one by one.

R1: IF wind speed <= 7 THEN class='very less'

R2: IF wind speed <= 12 and min temp > 25 THEN class='less'

R3: IF wind speed > 12 OR IF wind speed <12 AND min temp <=25 THEN class='sufficient'

Five districts input are given and the results are compared.

4. EXPERIMENTAL RESULTS

Selected districts such as Ariyalur, Kanyakumari, Pudukottai, Ramnathapuram and Tuticorin are taken for comparison for the month of May. Table I compares the results of wind speed in different districts.

```
=== Confusion Matrix ===(Ariyalur)

   a  b  c      <-- classified as
  13  0  0  |   a = very less
   0 11  0  |   b = sufficient
   0  0  7  |   c = less

=== Confusion Matrix ===(Kanyakumari)

   a  b  c      <-- classified as
   7  0  1  |   a = less
   0 10  0  |   b = very less
   0  0 13  |   c = sufficient

=== Confusion Matrix ===(Pudukottai)

   a  b  c      <-- classified as
  13  0  0  |   a = very less
   0 11  0  |   b = sufficient
   0  0  7  |   c = less
=== Confusion Matrix ===(Ramanathapuram)
   a  b  c      <-- classified as
  17  0  0  |   a = very less
   0  3  0  |   b = less
   0  0 11  |   c = sufficient

=== Confusion Matrix ===(Tuticorin)

   a  b  c      <-- classified as
  11  0  2  |   a = sufficient
   0 10  0  |   b = very less
   1  0  7  |   c = less
```

Table 1
Result Comparison of wind speed in 5 districts

Districts	*Accuracy*	*Mean absolute error*	*Root mean squared error*
Ariyalur	100	0	0
Kanyakumari	96.77	0.0403	0.1516
Pudukottai	100	0	0
Ramnathapuram	100	0	0
Tuticorin	90.32	0.0597	0.2356

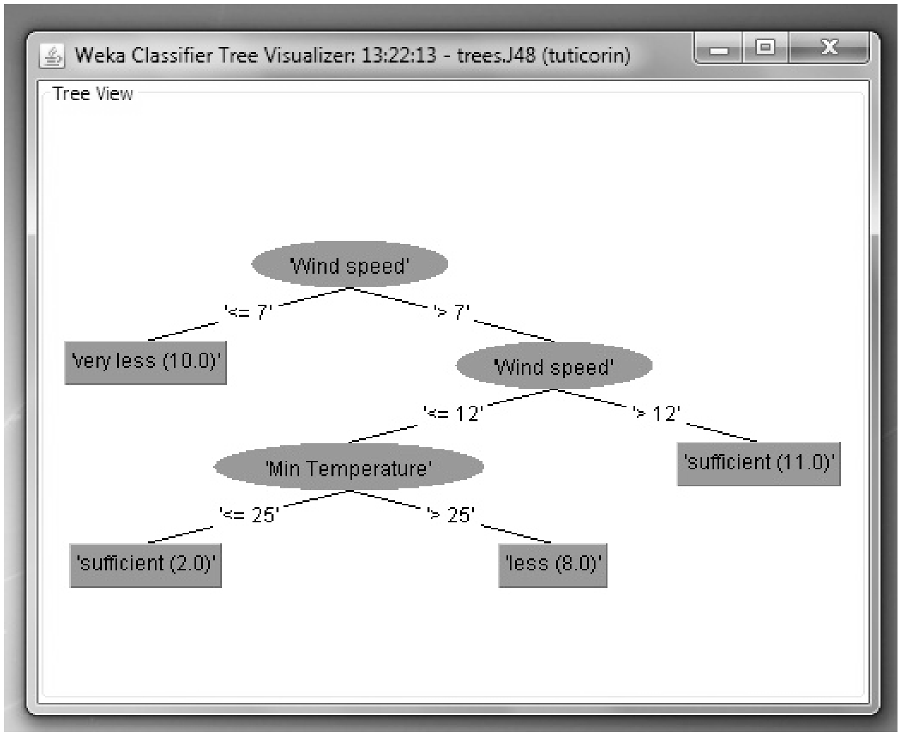

Figure 1: Decision tree

5. CONCLUSION

Meteorological data of various districts collected through were fed to the WEKA tool. The tool used the classification algorithm to classify the speed of wind turbine and gives the end user the appropriate time to operate the wind turbine. This finding forms a basis for forecasting wind power generation and thereby helps in planning and scheduling of power supply.

References

1. A. Sfetsos, "*A comparison of various forecasting techniques applied to mean hourly wind speed time series*", Vol. 21, pp. 23–35, Sep 2000.
2. Anurag More, M.C. Deo, "*Forecasting wind with neural networks*", Vol. 16, pp. 35-49, Feb 2003.
3. Damousis, I.G., Alexiadis, M.C.; Theocharis, J.B.; Dokopoulos, P.S., "*A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation*", Vol. 19, pp. 352-361, Jun 2004.
4. Thanasis G. Barbounis, John B. Theocharis, Minas C. Alexiadis, Petros S. Dokopoulos, "*Long-Term Wind Speed and Power Forecasting Using Local Recurrent Neural Network Models*", Vol. 21, pp. 273-284, Mar 2006.
5. George Sideratos and Nikos D. Hatziargyriou, "*An Advanced Statistical Method for Wind Power Forecasting*", Vol. 22, pp. 258-265, Feb. 2007.
6. T.H.M. El-Fouly, E.F. El-Saadany and M.M.A. Salama, "*Improved Grey predictor rolling models for wind power prediction*", Vol. 1, pp. 928-937, Nov 2007.
7. Annalisa Di Piazza, Maria Carmela Di Piazza, Antonella Ragusa, Gianpaolo Vitale, "*Environmental data processing by clustering methods for energy forecast and planning*", Vol. 36, pp. 1063-1074, Mar 2011.
8. Erasmo Cadenas, Wilfrido Rivera, "*Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA–ANN model*", Vol. 35, pp.2732-2738, Dec 2010.
9. Gong Li, Jing Shi, Junyi Zhou, "*Bayesian adaptive combination of short-term wind speed forecasts from neural network models*", Vol. 36, pp. 352-359, Jan 2011.
10. Ruddy Blonbou, "*Very short-term wind power forecasting with neural networks and adaptive Bayesian learning*", Vol. 36, pp. 1118-1124, Mar 2011.

11. Hassen Bouzgou, Nabil Benoudjit, "*Multiple architecture system for wind speed prediction*", Vol. 88, pp. 2463-2471, Jul 2011.
12. Zhenhai Guo, Weigang Zhao, Haiyan Lu, Jianzhou Wang, "*Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model*", Vol. 37, pp. 241-249, Jan 2012.
13. Nima Amjady, Farshid Keynia, and Hamidreza Zareipour, "*Wind Power Prediction by a New Forecast Engine Composed of Modified Hybrid Neural Network and Enhanced Particle Swarm Optimization*", Vol. 2, pp. 265-276, Jul 2011.
14. Kanna Bhaskar, S. N. Singh, "*AWNN-Assisted Wind Power Forecasting Using Feed-Forward Neural Network*", Vol. 3, pp. 306-315, Apr 2012.
15. Muhammad Khalid, Andrey V. Savkin, "*A Method for Short-Term Wind Power Prediction With Multiple Observation Points*", Vol. 27, pp. 579-586, May 2012.
16. Paolo Giudici, Applied Data Mining-Statistical Methods for business and Industry, Wiley Student Edition, 2005.
17. Margaret H. Dunham, Data Mining-Introductory and Advanced Topics, Pearson Education, 2003.
18. Han J. and Kamber M., Data mining concepts and techniques, Morgan Kauffmann publications, 2006.