# LITERATURE SURVEY


## ABSTRACT

Wind energy plays an increasing role in the supply of energy world wide. The energy output of a wind farm is highly dependent on the weather conditions present at its site. If the output can be predicted more accurately, energy suppliers can coordinate the collaborative production of different energy source more efficiently to avoid costly overproduction. In this paper, we take a computer science perspective on energy prediction based on weather data and analyze the important parameters as well as their correlation on the energy output. To deal with the interaction of the different parameters, we use symbolic regression based on the genetic programming tool Data Modeler. Our studies are carried out on publicly available weather and energy data for a wind farm in Australia. We report on the correlation of the different variables for the energy output. The model obtained for energy prediction gives a very reliable prediction of the energy output for a newly supplied weather data.


## INTRODUCTION

Wind speed/power has received increasing attention around the earth due to its renewable nature as well as environmental friendliness. With the global installed wind power capacity rapidly increasing, the wind industry is growing into a large-scale business. Reliable short-term wind speed forecasts play a practical and crucial role in wind energy conversion systems, such as the dynamic control of wind turbines and power system scheduling. A precise forecast needs to overcome problems of variable energy production caused by fluctuating weather conditions. Power generated by wind is highly dependent on the wind speed. Though it is highly non-linear, wind speed follows a certain pattern over a certain period of time. We exploit this time series pattern to gain useful information and use it for power prediction.

## MACHINE LEARNING

Machine learning is a branch of Artificial Intelligence, which help the computer to think like human and can take their own decision without human intervention. Due to rapidly development in Artificial Intelligent, Machine learning has lots of advancement in diagnosis of difference types of disease. Moreover, Machine learning algorithm gives us more accurate prediction and performance. Machine learning has been broadly divided into different types are

### a) SUPERVISED LEARNING

In easy word, supervised learning is types of learning method with the help of supervisor, teacher or instructor. It consists of training set of pattern associated with label data and makes it easy for algorithm from input to output and also easy to learn and predict. Some of supervised learning are classification such as KNN, SVM, Naïve Bayes, Neural network regression as linear and polynomial, Decision tree and Random forest. Developed prediction based on both input and output data

### a) UNSUPERVISED LEARNING

Unsupervised learning is also known as clustering. In unsupervised learning there is no training data set, no label and unknown output data. This type of learning method is like self-guide learning method. Some of the supervised learning methods are clustering such as K-Means clustering, SVD and PCA.

### b) SEMI SUPERVISED LEARNING

Semi supervised learning is types of learning method in Machine learning, these learning is in between training data with label (SL) and training data with no label (USL). These algorithms are performing better large amount of unlabeled data and less amount of label data

### c) REINFORCEMENT LEARNING

This is a type of machine learning based on agent, action, state, reward and environment. The software agent and machine to automatically define behavior with specific context based on their reward feedback.

## LITERATURE REVIEW

- Li, G., Shi, J., Zhou, J.: Bayesian adaptive combination of short-term wind speed forecasts from neural network models
- E. Frank, Y. Wang, S. Ingis, G. Holmes and I. H. Witten, "Using model trees for classification", Machine learning They include: the support vector machine regression (SVMreg) algorithm , multilayer perceptron (MLP) algorithm M5P tree algorithm Reduced Error Pruning (REP) tree (decision or regression tree) , and the bagging tree.
- B. G. Brown, R. W. Katz and A. H. Murphy, "Time series models to simulate and forecast wind speed and wind power", J. Appl. Meteorol.,Time series prediction focuses on determining future events based on known events, measured typically at successive times and spaced at (often uniform) time intervals.
- I H. Witten and E. Frank, Data Mining: Practical Machine Learning Tools and Techniques, CA, San Francisco:Morgan Kaufmann, 2005, 525 pp.For example, the boosting tree algorithm can be used to select the best predictors, as well as the wrapper approach using the genetic or the best-first search algorithms .
- Sfetsos presented a novel method to forecast the mean hourly wind speed using a time series analysis, and showed that the developed model outperformed the conventional forecasting models.

- T. G. Barbounis, J. B. Theocharis, M. C. Alexiadis and P. S. Dokopoulos, "Long-term wind speed and power forecasting using local recurrent neural network models", Trans. Energy Convers., vol. 21, no. 1, pp. 273-284, Mar. 2006.they T. G. Barbounis, J. B. Theocharis, M. C. Alexiadis and P. S. Dokopoulos, "Long-term wind speed and power forecasting using local recurrent neural network models", Trans. Energy Convers., vol. 21, no. 1, pp. 273-284, Mar. 2006.

## CONCLUSION

In this study, we showed that wind energy output can be predicted from publicly available weather data with accuracy up to 80% R2 on the training range and up to 85, 5% on the unseen test data. We identified the smallest space of input variables (windGust2 and dewPoint) where reported accuracy can be achieved, and provided clear trade-offs in prediction accuracy when decreasing the input space to the windGust2 variable. We demonstrated that an off-the-shelf data modeling and variable selection tool can be used with mostly default settings to run the symbolic regression experiments as well as variable importance, variable contribution analysis, ensemble selection, and validation. We are pleased that the presented framework is so simple that it can be used by literally everybody for predicting wind energy production on a smaller scale—for individual wind turbines on private farms or urban buildings, or for small wind farms. For future work, we are planning further study of the possibilities for longer-term wind energy forecasting.