Energy Efficient Medium-Term Wind Speed Prediction System using Machine Learning Models

Sasirekha Sankar *, S. Amudha, P. Madhavan, Dev Krishna Lamba
SRM Institute of Science and Technology, Kattankulathur, Chennai
Email: sasireks@srmist.edu.in

Abstract-With the day by day exhausting nonrenewable resources, it becomes crucial to focus on the renewable sources of energy and to get the maximum output from them. Wind energy is a major source of energy in many parts of India. Other such renewable sources of energy are solar energy, biomass energy etc. Our major objective during this project was to predict the wind speed for medium-term so as to help wind farms in channelizing the wind energy in an efficient manner and to get the maximum throughput from the wind farms.

Keywords- Wind speed prediction, neural networks, SVR, SVM, Deep learning, Medium term.

1. Introduction

Depletion of Natural Fossil Energy and growing global electricity needs is driving countries to incline towards renewable energy and its efficient utilization.

The pioneer type of renewable energy was solar and tidal which is now followed by wind power in last few decades. This source of wind power is completely free, unlimited and boundless.

The natural winds wavering nature is a major challenge encountered by the wind farms as it’s a production is neither constant nor stable.

This initiatives the need for wind speed prediction as it’s a driving force for accurate and efficient utilization of wind power.

The wind speed prediction is not just exploited for knowing the production of the wind farms but also use to judge multiple factors like deciding the location of the wind farm, planning the expected production, based on the terrain and weather choosing the size of the windmill.

The wind speed prediction is usually done using the traditional statistical methods like auto regression modelling with moving averages, but with invent of Machine Learning and AI
algorithms are increasing the efficiency of the prediction. In this paper Support Vector Regression (SVR) method is used to determine the wind speed from the dataset and predict accurately for medium term.

Support Vector Regression (SVR) is a regression algorithm similar to Support Vector Machine (SVM) where the working principles is same in both the concepts. The main purpose of this technique is to predict the single real time values rather than group of values. SVR ensures suitable prediction model and it can clearly predict the non-linearity in the real time data values.

Briefly, SVR can be thought of as an adapted form of SVM where the dependent variable is numerical rather than categorical. The main benefit of this SVR is usage of a non-parametric technique. The dependent variable and the in-dependent variables does not create an impact on the distribution output model as it’s reflected in the SLR model whether the outcome depends on the Gauss Markov assumptions. Instead, it depends on the kernel functions. The resultant variable have greater interpretation in SVR as it allows the building of a non-linear model without altering the explanatory variables.

The main principle in back of SVR is that it is not only caring about real time prediction but also used to predict long time errors, which is less than error \( (\epsilon) \) is less than certain value. This is known as the principle of maximal margin. SVR is a useful technique that provides the user with high flexibility in terms of distribution of underlying variables, relationship between independent and dependent variables and the control on the penalty term.

Basically, we have to decide a decision boundary at ‘e’ distance from the original hyper plane such that data points closest to the hyper plane or the support vectors are within that boundary line.

![Graphical model representation.](image)

**Fig. 1. Graphical model representation.**

Basically, we have to decide the value ‘e’ as distance between the original hyper plane and the boundary. The model needs to identify the closest data points closest either to the support vector or the hyper plane.

### 2. Overall Framework

#### A. Input data

Our input data includes information on wind speed and other related meteorological parameters that can affect the wind speed predictions such as the temperature, surface
pressure, direction, time etc. The data was collected for the year 2014 from a wind farm in the state of Tamil Nadu, India. We have a total of approximately 1608 x 7 data in .csv format consisting of wind speed and related meteorological factor’s information in time series.

B. Data Pre-processing

Before any of the actual prediction can be made, the raw data needs to be pre-processed. This process of data preprocessing includes verifying the following: i. The dataset is in appropriate format as required by the prediction model ii. There are no null values in the dataset. iii. Any abnormal data is removed or corrected. iv. Any redundant data/column is removed. v. All the rows that are not required for our prediction are removed in order to improve the accuracy.

C. Prediction Model

Once data pre-processing is done, it can be used as input to the SVR prediction model. The first step during the prediction involves dividing the entire dataset into training and test sets. The training set is used to train the system and to study the previous patterns of wind speed. It also helps to identify a relationship among wing speed and its dependency on other related meteorological factors. The training is achieved through the data set allocated for training and to predict the future wind speeds and to measure the accuracy of the predictions. The whole data set is proportioned as 20% for testing purpose and 80% of the remaining dataset for the training purpose. Once the testing and training sets are identified, we feed the training dataset to the model to train itself and realize the various inter-dependencies among the wind speed and other meteorological factors such as direction, temperature, surface pressure, time etc that can directly affect the wind speed of an area. Once these dependencies are established and the model trains itself, the test dataset is applied to serve the outcome of the predictions.

D. Outcome

The wind speed outcome from our prediction is in the same format as our input i.e. tabular format and consists of a single column with the predicted wind speed values.

E. Prediction Score

The accuracy for the predictions is calculated using r2_score functionality from scikit learn package. It takes two variables as input for the accuracy calculation which are the test set and the prediction outcome that we get. The results of the prediction are plotted in a graph format for easy visualization. Figures 2-4 show the results.

F. Outcome Graph
3. Result and Conclusion

This paper presents a wind forecasting method using Support Vector Regression (SVR). The predictions were based on the wind speed time series data that was collected for an year from a wind farm and also takes into account the several meteorological factors as input for accurate predictions.
The model is built based on the data segregated as training data set and the model is validated based on the test data set. The test data sets benchmark for the accuracy of the refined model. Using this SVR methodology, the final accuracy of the predictions that we got was 60.78%.

References

[1] https://medium.com/coinmonks/supportvectorregression-or-svr-8eb3aecf6d0ff
[2] https://www.kdnuggets.com/2017/03/building-regression-models-support-vector-regression.html
[3] Vikram Bali, Ajay Kumar - Deep Learning based wind speed forecasting.
[4] Huai Nana, Dong Lei, Hao Ying – Short-term wind speed prediction using CNN-GRU model.
[5] Yiwei FU, Wei HU, Maolin Tang, Rui YU - Multi-step Ahead Wind Power Forecasting Based on Recurring Neural Networks (RNN).
[6] Yu-Xi-Wu, Jia-Qi Zhu, Qing Biao - Datadriven wind speed forecasting using deep feature extraction and LSTM
[7] Qiaomu Zhu, Jinfu Chen, Xiang Bai – Learning temporal and spatial correlations jointly- A unified framework for wind speed prediction.