PREDICTING POWER OUTPUT BASED ON WEATHER CONDITION ON WIND TURBINES

Naresh babu Katabathuni[1] ,Sareen Gundabathina[2], DivyasreeGummadi[3], Naga Venkata Komal Kunchaparthi[4], Sri Lakshmi Kalapala[5]

[1,2,3,4] Department of CSE, Gudlavalleru Engineering College, India.

[5]Assistant Professor, Department of CSE, Gudlavalleru Engineering College, India

ABSTRACT:

The purpose of the project is to predict the electricity output based on the rotating speed of the wind turbines; here we get the output based on the previous data. In this we have two independent variables LV Active Power, Wind Speed and two dependent variable Theoretical Power and wind direction. For that we use one of the regression algorithm like polynomial regression and predict the output by using r2_score.In this we train the model with the previous data and test the data by giving the independent variables and predicting whether it is giving expected output or not. We use a node red for deploying the nodes, we used http request, timestamp, for sending the message we used message payload, for power detection we used theoretical power, inputs as Active Power and Wind Speed and to implement the user interface. There are two steps in the process of wind power prediction. In the first step raw data is collected by the power plant information system and is filtered. This prepares a valid data to be used for building a prediction model. In the second step we use all the regression algorithms to build a model to predict the wind power. It is to achieve a high accuracy with respect to the measured data.

INTRODUCTION:

Wind plant has lower cost of energy compared to other renewable energy source for large scale application. Due to the different geographical patterns, weather, and properties of the wind turbines, a wind turbine may have various performances given different situations. If the total output of a wind power plant can be predicted with high accuracy, more useful information can be provided to the power companies to help in scheduling the power generation. This information will allow a more flexible and intelligent control at a WPP (Wind power Plant). Methods for predicting wind power generation can be categorized into physical methods, statistical methods, methods based on neural networks, and hybrid methods. The physical methods rely heavily on numeric whether prediction, which is confined by the sensors and monitoring devices placed within the WPP. The quality of hardware chosen, the parameters settings, the computation time, the time delay, and the sampling rates influence the accuracy of data collected from the WPP. It is easier to predict a single wind turbines performance rather than a whole WPP power generation. Statistical and neural networks methods are based on the historical data and have a low prediction cost. The relationship between input data and output data based on the historical measured data is learned and then a nonlinear relationship model between them is built. But when new data not previously included in the training data set is used as input into this kind of model, the prediction error might be large, which is a disadvantage. Different prediction methods mentioned above can be combined as hybrid methods to achieve better prediction results. But this will increase the complexity of the model. In this paper, multiple linear regressions are applied to predict the power generation at wind power plant.

In this WPP, data of wind information, such as wind speed, wind direction, wind power generation, humidity and air pressure are collected by a Plant Information (PI) system.

• The raw data set will be screened by probabilistic neural network to prepare high quality data for building neural network models.

• The model's inputs do not rely on the data of wind speed and wind direction from all turbines; representative wind turbines can be found to compress the length of the input data.

• The inputs are expressed as complex-valued data (vector representation) which combine wind speed and wind direction.

• The complex-valued recurrent neural network model's time series inputs are generated based on the historical data values of the WPP rather than the predicted values by the model at the previous steps.

• The result to be predicted is the total power generation of the whole WPP rather than outputs of some single wind turbines.

The evaluation of wind potential in a region requires systematic data collection and analysis on wind speed and regime. Generally, a rigorous assessment requires specific surveys of the region where the wind farm will be placed. There are three major markets for the field of global wind power generation: Europe, USA and China. Wind energy penetration levels continue to rise, led by Denmark with a 40% use of this energy, followed by Uruguay, Portugal and Ireland with over 20%; Spain and Cyprus with about 20%; Germany with 16%; and the major markets of China, the US and Canada with 4%, 5.5% and 6% wind energy, respectively. The forecast of five years ahead is almost 60 GW of new wind power installations in 2017, rising to an annual market of 75 GW by 2021, and an accumulated installed capacity of more than 800GW by the end of 202. Wind energy is a clean and renewable alternative for the production of electric energy, presenting great social acceptance. In the social feature, wind power plants do not cause major environmental impacts such as in hydroelectric plants and allow the compatibility between the production of electricity from the wind and the use of land for livestock and agriculture. Wind generation occurs through the contact of the wind with the blades of the wind device. When rotating, these blades convert wind speed into mechanical energy that drives the rotor of the wind generator, which produces electricity. According to, tropical regions receive solar rays almost perpendicularly and are therefore warmer than the Polar Regions.

RELATED WORK:

Research objectives are prepared after correlating the various works done by contemporary researchers. Most of the researchers have developed methods for wind speed based forecasting. Apart from wind speed forecasting, many other parameters required to assess the wind energy potential are studied. Meteorological information coupled with topographical data has to be utilized for wind power estimation at a particular place. After this, wind turbine power curves have to be mapped against the wind parameters. [1] This is to establish the number of energy units that can be generated at wind plant level, for a month. In most of the wind plants in India, wind speed measurements taken from a single location for the whole plant which is an average indicator of wind power potential at plant level. Deviations from mean position of the wind measurement to the wind turbine are not considered. Hence, Average wind speed indicator continues to contribute for wind 38 power estimation at plant level. Wind speed and wind power generation along with changing values across time series in the present focus of work.

Wind power forecasts associated with wind resource assessment methodologies, data and experience through modeling will help the wind power plants to meet the changing grid requirements. [1] Forecasting serves two purposes – allows wind generators to take operational decisions which are aligned to market prices of energy, allows the system operator to better plan ancillary services required to balance generation from variable renewable energy sources.

Wind power estimation models are majorly classified as models based on physical or numerical weather prediction models, statistical models. [1] Physical models are based on the meteorological data and open boundary fluid flow equations. Other energy equations are solved as nonlinear solvers in predicting wind energy potential of selected space.

A number of different approaches have been applied to forecast wind speed and the power produced by wind farms. Potter and Negnevitsky [2] applied the adaptive neurons fuzzy inference approach to forecast short-term wind speed and direction.

Barbounis et al. [2] used the nonlinear recursive least-squares method to train a recurrent neural network (NN) based on the meteorological data. Their model has improved the accuracy of long-term wind speed and power forecasting. Damousiset al.[8] developed a fuzzy logic model and trained it with a genetic algorithm. The model was then used to forecast wind speed over horizons ranging from 0.5 to 2 h. Li et al. [2] compared regression and NN models for wind turbine power estimation, and reported that the NN model outperformed the regression model. Sfetsos [2] 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.

Developing prediction models for wind farms is a challenge, as the power is mainly determined by the wind speed that is difficult to forecast accurately. [3] The wind speed depends on parameters such as air pressure, temperature, terrain topography, etc. The stochastic nature of a wind farm environment calls for new modeling approaches to accurately predict the power to be produced in the future time periods.

Two different methodologies for power prediction have been employed. [3] The models are built using historical data collected by supervisory control and data acquisition (SCADA) systems installed at a wind farm. A short-term power prediction is important in dispatching power to meet customer needs. For long horizon predictions, meteorological data are usually used.

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 DataModeler. 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. [4] The model obtained for energy prediction gives a very reliable prediction of the energy output for newly supplied weather data.

Genetic programming has its main success stories in the field of symbolic regression. Given a set of input output vectors, the task is to find a function that maps the input to the output as best as possible, while avoiding over fitting. [4] The resulting function is later often used to predict the output for a newly given input. Syntax trees represent functions in this case, and the functions are changed by crossover and mutation to produce new functions. The quality of a syntax tree is determined by how well it maps the given set of inputs to their corresponding outputs.

Wind energy plays an increasing role in the supply of energy world wide. [5] The energy output of a wind farm is highly dependent on the weather conditions present at its site. [5] If the output can be predicted more accurately, energy suppliers can coordinate the collaborative production of different energy sources 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. [6] To deal with the interaction of the different parameters, we use symbolic regression based on the genetic programming tool DataModeller. Our studies are carried out on publicly available weather and energy data for a wind farm in Australia. [6] 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 newly supplied weather data.

PROPOSED METHODOLOGY:

Multiple linear regression attempts to model the relationship between two or more explanatory variables and a response variable by fitting a linear equation to observed data. Every value of the independent variable *x* is associated with a value of the dependent variable *y*. The population regression line for *p* explanatory variables $x_1, x_2, \ldots, x_p$ is defined to be $\mu_y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p$. This line describes how the mean response μ_y changes with the explanatory variables. The observed values for *y* vary about their means μ_y and are assumed to have the same standard deviation σ. The fitted values $b_0, b_1, \ldots, b_p$ estimate the parameters $\beta_0, \beta_1, \ldots, \beta_p$ of the population regression line.

Since the observed values for *y* vary about their means μ_y, the multiple regression model includes a term for this variation. In words, the model is expressed as DATA = FIT + RESIDUAL, where the "FIT" term represents the expression $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots \beta_p x_p$. The "RESIDUAL" term represents the deviations of the observed values *y* from their means μ_y, which are normally distributed with mean 0 and variance σ. The notation for the model deviations is ε.

Formally, the model for multiple linear regression, given *n* observations, is $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots \beta_p x_{ip} + \varepsilon_i$ for $i = 1,2, \ldots n$.

In the least-squares model, the best-fitting line for the observed data is calculated by minimizing the sum of the squares of the vertical deviations from each data point to the line (if a point lies on the fitted line exactly, then its vertical deviation is 0). Because the deviations are first squared, then summed, there are no cancellations between positive and negative values. The least-squares estimates b_0, b_1, ... b_p are usually computed by statistical software.

The values fit by the equation $b_0 + b_1x_{i1} + \ldots + b_px_{ip}$ are denoted $\hat{y}_i$, and the residuals e_i are equal to $y_i - \hat{y}_i$, the difference between the observed and fitted values. The sum of the residuals is equal to zero.

The variance σ^2 may be estimated by $s^2 = \dfrac{\sum e_i^2}{n-p-1}$, also known as the mean-squared error (or MSE).
The estimate of the standard error s is the square root of the MSE.

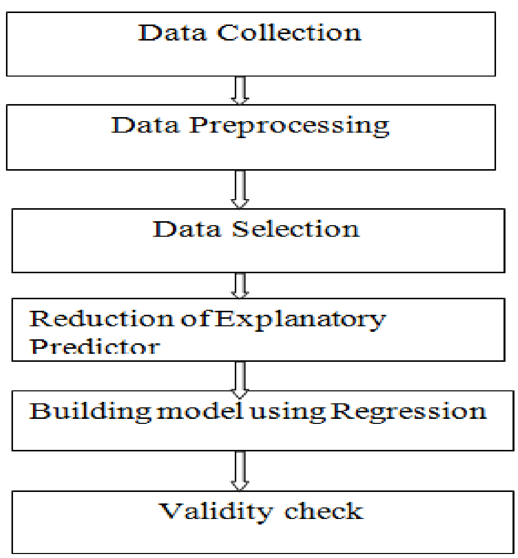
Multiple regression analysis has three main uses:

1. You can look at the strength of the effect of the independent variables on the dependent variable.

2. You can use it to ask how much the dependent variable will change if the independent variables are changed.

3. You can also use it to predict trends and future values.

Architecture:



3.3 DATA PREPARATION

As we are using machine learning algorithm we need a large datasets in order to get the test results. We have number of websites for downloading the data we used Kaggle website for the dataset. In that we have a number of datasets regarding the project we collected this dataset because it is having many inputs for efficient accuracy and it is a continuous data process running, we are also having many number of datasets regarding the output prediction but they are not much efficiently known by the user.

In this data we have input variables as LV Active Power, Wind Speed and output variable as Theoretical Power which are independent and dependent variables. We have Wind Speed data for every m/s. We have more accurate results.

Date/Time	LV ActivePower (kW)	Wind Speed (m/s)	Theoretical_Power_Curve (KWh)	Wind Direction (Ä°)
01 01 2018 00:00	380.0477905	5.31133604	416.3289078	259.9949036
01 01 2018 00:10	453.7691956	5.672166824	519.9175111	268.6411133
01 01 2018 00:20	306.3765869	5.216036797	390.9000158	272.5647888
01 01 2018 00:30	419.6459045	5.659674168	516.127569	271.2580872
01 01 2018 00:40	380.6506958	5.577940941	491.702972	265.6742859
01 01 2018 00:50	402.3919983	5.604052067	499.436385	264.5786133
01 01 2018 01:00	447.6057129	5.793007851	557.3723633	266.1636047
01 01 2018 01:10	387.2421875	5.306049824	414.8981788	257.9494934
01 01 2018 01:20	463.6512146	5.584629059	493.6776521	253.4806976
01 01 2018 01:30	439.725708	5.523228168	475.7067828	258.7237854
01 01 2018 01:40	498.1817017	5.724115849	535.841397	251.8509979
01 01 2018 01:50	526.8162231	5.934198856	603.0140765	265.5046997
01 01 2018 02:00	710.5872803	6.547413826	824.6625136	274.2329102
01 01 2018 02:10	655.1942749	6.199746132	693.4726411	266.7331848
01 01 2018 02:20	754.7625122	6.505383015	808.0981385	266.7604065
01 01 2018 02:30	790.1732788	6.634116173	859.4590208	270.4931946
01 01 2018 02:40	742.9852905	6.378912926	759.4345366	266.5932922
01 01 2018 02:50	748.2296143	6.446652889	785.2810099	265.5718079
01 01 2018 03:00	736.6478271	6.415082932	773.1728635	261.1586914
01 01 2018 03:10	787.2462158	6.437530994	781.7712157	257.5602112
01 01 2018 03:20	722.8640747	6.220024109	700.7646999	255.9264984
01 01 2018 03:30	935.0333862	6.89802599	970.7366269	250.0128937
01 01 2018 03:40	1220.609009	7.60971117	1315.048928	255.9857025
01 01 2018 03:50	1053.771973	7.288355827	1151.265744	255.4445953

Fig 1: Dataset of our project

LV Active Power(kW):

A very simple and effective means to eliminate leading or lagging power factor errors, reduce voltage fluctuations, enhance equipment operating life and improve system power capacity. The range offers consolidated features in one package without the risk of resonance, while stabilizing electrical networks by providing harmonic mitigation, power factor correction and load balancing.

In a simple alternating current (AC) circuit consisting of a source and a linear load, both the current and voltage are sinusoidal If the load is purely resistive, the two quantities reverse their polarity at the same time. At every instant the product of voltage and current is positive or zero, the result being that the direction of energy flow does not reverse. In this case, only active power is transferred.

Wind speed (m/s):

Wind speed is affected by a number of factors and situations, operating on varying scales (from micro to macro scales). These include the pressure gradient, Rossby waves and jet streams, and local weather conditions. There are also links to be found between wind speed and wind direction, notably with the pressure gradient and terrain conditions.

Wind speeds usually mean the movement of air in an outside environment, but the speed of movement of air inside is also important in many cases,including weatherforecasting, aircraft and maritime operations, construction and civil engineering. High wind speeds can cause unpleasant side effects, and strong winds often have special names,including gales, hurricanes, and typhoons.

Theoretical power curve(KWh):

The **power** of a binary hypothesis test is the probability that the test rejects the null hypothesis (H_0) when a specific alternative hypothesis (H_1) is true. The statistical power ranges from 0 to 1, and as statistical power increases, the probability of making a type II error (wrongly failing to reject the null hypothesis) decreases. For a type II error probability of β, the corresponding statistical power is 1-β. For example, if experiment 1 has a statistical power of 0.7, and experiment 2 has a statistical power of 0.95, then there is a stronger probability that experiment 1 had a type II error than experiment 2, and experiment 2 is more reliable than experiment 1 due to the reduction in probability of a type II error. It can be equivalently thought of as the probability of accepting the alternative hypothesis (H_1) when it is true that is, the ability of a test to detect a specific effect, if that specific effect actually exists.

Wind direction:

Wind direction is reported by the direction from which it originates. For example, a *northerly* wind blows from the north to the south.[1] Wind direction is usually reported in cardinal directions or in azimuth degrees. Wind direction is measured in degrees clockwise from due north. Consequently, a wind blowing from the north has a wind direction of 0° (360°); a wind blowing from the east has a wind direction of 90°; a wind blowing from the south has a wind direction of 180°; and a wind blowing from the west has a wind direction of 270°. In general, wind directions are measured in units from 0° to 360°, but can alternatively be expressed from -180° to 180°.

CONCLUSION AND FUTURE SCOPE:

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 where reported accuracy can be achieved, and provided clear trade-offs in prediction accuracy when decreasing the input space to the wind speed variable.

We are pleased that the presented framework is so simple that it can be used by literally everybody for predicting Theoretical power based on wind energy production —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.

Several forecasting models were discussed and a lot of researches on the models, which have their own characteristics, were presented. The major focus was on emphasizing the diversity of various forecasting methods available and also on providing a comparison of present mechanisms to determine the best available.

REFERENCES:

[1] O. Kramer and F. Gieseke. Short-term wind energy forecasting using support vector regression. In International Conference on Soft Computing Models in Industrial and Environmental Applications, pages 271–280. Springer, 2011.

[2] A. Kusiak, H. Zheng, and Z. Song. Short-term prediction of wind farm Power: A data mining approach. IEEE Transactions on Energy Conversion, 24(1):125 – 136, 2009.

[3] R. Poli, W. B. Langdon, and N. F. McPhee. A Field Guide to Genetic Programming. lulu.com, 2008. ISBN 978-1-4092-0073-4.

[4] M. Schmidt and H. Lipson. Age-fitness Pareto optimization. In Genetic Programming Theory and Practice VIII, Genetic and Evolutionary Computation, chapter 8, pages 129–146. Springer, 2010.

[5] I. Sanchez. Short- term prediction of wind energy production. International Journal of Forecasting, 22(1):43 – 56, 2006.

[6] M. Webb and S. Scuglia. Wind power: A favored climate change response. Global Economic Research: Fiscal Pulse (Scotia bank), 2007.

Github-link: https://github.com/katabathuni-naresh/Predicting-power-based-on-wind-turbines

