PREDICTION OF WIND SPEED AND WIND DIRECTION USING MACHINE LEARNING ALGORITHM

Esther Pushpam V.S, Asst.Prof,
Department of CSE,
Sri Ramakrishna Institute of Technology,
Coimbatore - 641010

Karnan M
Computer Science and Engineering
Sri Ramakrishna Institute of Technology
Coimbatore-10, India

Madhumitha R
Computer Science and Engineering
Sri Ramakrishna Institute of Technology
Coimbatore-10, India

Mohilesh M
Computer Science and Engineering
Sri Ramakrishna Institute of Technology
Coimbatore-10, India

Abstract

A group of wind turbines in the same location used to produce electricity . Turbines can work efficiently with optimum wind speed , heavy wind cause damage to the turbine, In this project , the wind speed and direction for wind farm can be predicted to the efficient work of wind turbines. So, the output of the wind turbines will be having greater efficiency. Big data and Machine learning is outlined because the massive assortment of datasets that is advanced to method. Predicting wind speed and direction is one of the most important and critic tasks in a wind farm. Machine learning techniques are often used to predict the time series non-linear wind evolution. In this context, this project proposes a short term wind data prediction model based on Least Square Support Vector Machines (LSSVM) and Random forest algorithm in their regression mode, which have the advantage of being simple, fast and well adapted for the short term. This project also tries to prove how wind direction may influence power generation, and why it is important to predict it. A real time series data set contains the historical values of parameters like wind speed, wind direction which are used to predict the wind speed and direction . The predicted model will be evaluated with Mean Absolute , Mean Square Error value and both models will be compared for their performance.

Keywords—Wind speed, Wind direction, Machine learning

I. Introduction

A. Background History

A powerhouse, also called a wind generation station or alternative energy plant, could be a group of wind turbines within the same location accustomed produce electricity. Wind farms vary in size from atiny low number of turbines to many hundred wind turbines covering an intensive area. Wind farms are often either onshore or offshore. Many of the most important operational onshore wind farms are located in China, India, and also the u.

s.. Because they require no fuel, wind farms have less impact on the environment than many other sorts of power generation. Wind farms have, however, been criticized for his or her visual impact and impact on the landscape. The faster the common wind speed, the more electricity the turbine will generate, so faster winds are generally economically better for power station developments. The balancing factor is that strong gusts and high turbulence require stronger dearer turbines, otherwise they risk damage.

B. Problem Statement

Wind is a free energy source however it is highly unpredictable which is significant problem for prediction of wind speed and wind direction . Wind data need to be collected is a time series manner to process the data with its accuracy. Data need to cleaned and explored to identify its insights because the wind characteristic varies based on humidity , temperature , pressure in time to time. To make a prediction its is necessary to identifying appropriate machine learning algorithm which can be obtained based on literature survey. With a collected data of wind farm , the model need to build and trained well to achieve greater accuracy.

C. Wind Turbine

The amount of energy generated by a wind turbine depends on wind speed (main factor the area swept by the blades air density. Wind turbines requires as follows

1) A minimum wind speed (generally 12-14 km/h) to begin turning and generate electricity
2) Strong winds (50-60 km/h) to generate at full capacity
3) Winds of less than 90 km/h; beyond that speed, the turbines must be stopped to avoid damage.

D. Machine Learning

Machine learning is very important because it gives enterprises a view of trends in customer behavior and business operational patterns, in addition as supports the event of latest products. Many of today's leading companies, like Facebook, Google and Uber, make machine learning a central a part of their operations. Machine learning has become a major competitive differentiable for several companies. Classical machine learning is usually categorized by how an algorithm learns to become more accurate in its predictions. There are four basic approaches: supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning. the kind of algorithm data scientists like better to use depends on what kind of data they need to predict.

- Development of new machine learning algorithms that learn more accurately, utilize data from dramatically more diverse data sources available over the Internet and intranets, and incorporate more human input as they work;
- Integration of these algorithms into standard database management systems; and
- An increasing awareness of data mining technology within many organizations and an attendant increase in efforts to capture, warehouse, and utilize historical data to support evidence-based decision making.

In this kind of machine learning, data scientists supply algorithms with labeled training data and define the variables they need the algorithm to assess for correlations. Both the input and therefore the output of the algorithm is specified.

Unsupervised learning: this sort of machine learning involves algorithms that train on unlabeled data. The algorithm scans through data sets trying to find any meaningful connection. the information that algorithms train on in addition because the predictions or recommendations they output are predetermined.

Semi-supervised learning: This approach to machine learning involves a combination of the 2 preceding types. Data scientists may feed an algorithm mostly labeled training data, but the model is liberal to explore the information on its own and develop its own understanding of the info set.

Reinforcement learning: Data scientists typically use reinforcement learning to show a machine to finish a multi-step process that there are clearly defined rules. Data scientists program an algorithm to finish a task and provides it positive or negative cues because it works out the way to complete a task. except for the foremost part, the algorithm decides on its own what steps to require along the way.

E. Applications

Prediction of Wind speed and Wind direction are often utilized in foretelling, Aviation and Maritime operations, Construction projects, Growth and Metabolism rate of the many plant species, and has countless other implications

Weather forecasting:

Making a weather outlook involves three steps: observation and analysis, extrapolation to seek out the long run state of the atmosphere, and prediction of particular variables. one qualitative extrapolation technique is to assume that weather features will still move as they need been moving.

Aviation and Maritime operations:

AMO - aviation and maritime operations is to shield the people and nation's critical infrastructure through the coordinated use of air and marine assets to detect, interdict and stop acts of terrorism and also the unlawful movement of individuals, illegal drugs, and other contraband toward or cross the borders of the u. s..

F. Scope

The scope of the project is to predict the wind speed and wind direction for wind forms to run the turbines efficiently to produce Wind power. Wind power describes the process by which wind is used to generate electricity. Wind turbines convert the kinetic energy in the wind into mechanical power. A generator can convert mechanical power into electricity. Mechanical power can also be utilized directly for specific tasks such as pumping water and the extent of wind resource may help determine what type of a turbine you want to install, and many landowners install wind turbines simply because they believe strongly in using non-polluting, inexhaustible forms of energy.

II. LITERATURE SURVEY

1) A comparison of permanent magnet and wound rotor synchronous machines for portable power generation

Michelle Bash; Steve Pekarek; Scott Sudhoff; Jennifer Whitmore; Michelle Frantzenet al. In the year of 2019, Permanent magnet and wound rotor synchronous machines (PMSMs and WRSMs) are often used in diesel engine-based portable power generation systems. In these applications, there is a growing desire to improve machine efficiency in order to reduce fossil fuel requirements. In addition, there is a desire to reduce mass to improve mobility. To attempt to address these competing performance objectives, a system analyst is confronted with numerous choices, including machine type (PM or WR), converter architecture (active/passive), and control. Herein, to assist the analyst, design tools capable of performing automated multi-objective optimization of PMSMs and WRSMs connected to both active and passive rectifiers are described. The tools are then used to derive tradeoffs between mass and efficiency for a 3 kW application

2) An energy-efficent routing protocol for WSN-based intelligent mining system

Yanyan Mao; Dapeng Cheng et al. Comparing In recent years, Wireless sensor network (WSNs) has become the main technology to construct underground network for intelligent mining system. Due to the complex roadway topology, there needs an optimized energy conservation routing protocol. We proposed EERP, an energy-efficient routing protocol for for WSN-based intelligent mining system, which can prolonged lifetime of a network and reduce energy consumption through construct a dominator chain by region partition. Simulation results show that EERP has 8.2 times stable working time than LEACH. When region length is 100m, EERP still can maintain 99% energy after 3000rd transmission.

3) Energy efficient power management

M. Magno; L. Benini; L. Gaggero; J. P. La Torre Aro; E. Popovici et al. (2017), Technology advancements of sensors, low power mixed-signal/RF circuits, and Wireless Sensor Networks (WSNs) made it possible to construct compact and low cost solutions for a wide range of applications such as healthcare, surveillance, building monitoring, sports and fitness, etc. A new emerging generation of WSN is the Body Sensor Network which is suitable to monitor of the human body, mainly for health care applications but also for sports and fitness. Due to the growing interest in this area, novel sensors for monitoring bio-electric signals like ECG, EMG, EOG and EEG are being developed. In this paper we present a low-power designed body sensor networks (BSN) platform for on-body physiological measurements and wireless data communications. The BSN node is capable of hosting high sensitivity electric potential dry surface sensors that can be used in either contact or non-contact mode to measure ECG and EMG signals. The wireless connectivity is provided from a Bluetooth module to be connected to a PC or Smartphone and an 802.15.4 transceiver. Moreover as the platform is designed with low power in mind it incorporates adaptive power modes and an ultra-low power wake up radio to switch off/on the sensors and the radio transceivers in order to save energy. Experimental measurements show the acquisition of the novel sensors and the low power management consumption achievable with the node in different modalities. The power consumption in deep sleep is only 1.8uW. Thus, due to the presence of the novel sensors, the low power consumption and the wireless connectivity, the BSN platform will greatly facilitate the research and development activities for pervasive healthcare, medicine, wearable medical devices and other emerging biomedical engineering fields.

4) Energy-efficient time synchronization algorithm

Qianping Wang; Wei Wang; Yu-e Su; Yuan An et al. In the year of 2020, energy efficiency is an important standard for evaluating the effectiveness of the time synchronization protocols. Hence, energy should be conserved as much as possible provided that algorithm can achieve a certain accuracy. In this paper, we propose an energy efficient time synchronization algorithm in which each layer of the network can be synchronized merely by receiving broadcast messages rather than relying on two-way message exchange. Simulation results show that our improved strategy reduced the energy consumption and prolonged the lifetime of network compared with TPSN.

5)Time-series prediction of wind speed using machine learning algorithms: A case study Osorio wind farm, Brazil

A.Khosravi, L. Machado, R. O. Nunes, Machine learning algorithms (MLAs) are applied to predict wind speed data for Osorio wind farm that is located in the south of Brazil, near the Osorio city. Forecasting wind speed in wind farm regions is valuable in order to obtain an intelligent management of the generated power and to promote the utilization of wind energy in grid-connected and isolated power systems. In this study, multilayer feed-forward neural network (MLFFNN), support vector regression (SVR), fuzzy inference system (FIS), adaptive neuro-fuzzy inference system (ANFIS), group method of data handling (GMDH) type neural network, ANFIS optimized with particle swarm optimization algorithm (ANFIS-PSO) and ANFIS optimized with genetic algorithm (ANFIS-GA) are developed to predict the time-series wind speed data. The Time-series prediction describes a model that predicts the future values of the system only using the past values. Past data is entered as input and future data to be used for represents MLA output. The developed models are examined on 5-min, 10-min, 15-min and 30-min intervals of wind speed data. The results demonstrated that the GMDH model for all time intervals can successfully predict the time-series wind speed data with a high accuracy. Also, the combination of ANFIS models with PSO and GA algorithms can increase the prediction accuracy of the ANFIS model for all time intervals.

6) Prediction of wind speed and wind direction using artificial neural network, support vector regression and adaptive n euro-fuzzy inference system

A.Khosravi, R. N. N. Koury, L.Machado, J.J.G.Pabon In this study, three models of machine learning algorithms are implemented to predict wind speed, wind direction and output power of a wind turbine. The first model is multi-layer feed- forward neural network (MLFFNN) that is trained with different data training algorithms. The second model is support vector regression with a radial basis function (SVR-RBF). The third model is adaptive n euro-fuzzy inference system (ANFIS) that is optimized with a partial swarm optimization algorithm (ANFIS-PSO). Temperature, pressure, relative humidity and local time are considered as input variables of the models. A large set of wind speed and wind direction data measured at 5-min, 10-min, 30-min and 1-h intervals are utilized to accurately predict wind speed and its direction for Bushehr.

7) An intelligent method for wind power forecasting based on integrated power slope events prediction and wind speed forecasting

Fudong Li, Huan-yu Liao In this paper, we study an intelligent wind power prediction method by taking the prediction time horizons and prediction accuracy into account. The wind power slope events are defined, and multiple support vector machines are applied to the classification of slope down/up events for multi step-ahead scenarios. The wind speed series are decomposed by using the maximum overlap discrete wavelet transform (MODWT), and each decomposed signal is forecast using an adaptive wavelet neural network (AWNN) individually. The network is trained for wind speed prediction up to 24 h ahead. Based on slope events forecasting and wind speed forecasting, an improved radial basis function neural network (RBFNN) is proposed to predict wind power up to 24 h ahead. The proposed model is tested by using wind power data collected from a real wind farm. The analysis results validate that both the prediction time horizons and the prediction accuracy are guaranteed, and the proposed method can be applied to the optimal scheduling of wind farms 1 day in

8)Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extreme optimization

JieChena, Guo-QiangZengb, WunengZhoua, WeiDua, Kang-DiLua As an essential issue in wind energy industry, wind speed forecasting plays a vital role in optimal scheduling and control of wind energy generation and conversion. In this paper, a novel method called Ensemble LSTM is proposed by using nonlinear-learning ensemble of deep learning time series prediction based on LSTMs (Long Short Term Memory neural networks), SVRM (support vector regression machine) and EO (extremal optimization algorithm). First, in order to avert the drawback of weak generalization capability and robustness of a single deep learning approach when facing diverse form data, a cluster of LSTMs with diverse hidden layers and neurons are employed to explore and exploit the implicit information of wind speed time series. Then predictions of LSTMs are aggregated into a nonlinear-learning regression top- layer composed of SVRM and the EO is introduced to optimize the parameters of the top-layer. Lastly, the final ensemble prediction for wind speed is given by the fine- turning top-layer. The proposed Ensemble LSTM is applied on two case studies data collected from a wind farm in Inner Mongolia, China, to perform ten-minute ahead utmost short term wind speed forecasting and one-hour ahead short term wind speed forecasting. Statistical tests of experimental results compared with other popular prediction models demonstrated the proposed Ensemble LSTM can achieve a better forecasting performance.

III. EXISTING AND PROPOSED SYSTEM

A. Existing System

In existing system, they developed a Binary Electronic Sequence Calculator (BESK) for real time forecasts. But the duration of forecast was less because the computers didn't have the capacity to store longer forecast data. Later, they also use ANN (Artificial Neural Network), SVM (Support Vector Machine) for prediction purpose and they use only one algorithm at a time to predict the wind speed and wind direction accuracy, by finding the Mean Absolute Error for the above algorithms. And they are also comparing different types of artificial neural network to obtain the greater efficiency

B. Proposed System

In Proposed systems consist of input data which is obtained from dataset. it consists of parameters like sea level pressure, temperature, relative humidity, wind directions etc.. The proposed system consists of six blocks. Data collection, pre-processing, model building are all different steps whereas splitting the data into training and testing is considered to be a single step. the next step is to do the prediction which is the output block. The first block is the Meteorological Data. This is raw data which needs to be pre-processed before using it. So, this step includes removal of missing values if any, scaling it to a standard form. The next step is to build a model. The two models that are used are: SVM (Support Vector Machines), Random Forest. The data set has been split into training and testing in the next step. Predictions are obtained which are then matched against the actual values to calculate the error.

IV. SYSTEM SPECIFICATION

A. . Introduction

A technology that enables a machine to stimulate human behavior to help in solving complex problems is known as Artificial Intelligence. Machine learning is a subset of AI and allows machines to learn from past data and provide an accurate output. In our proposed system , the prediction of wind speed and direction has been made on Time series data .The data will be obtained from wind farm. In Prepossessing the data need to be cleaned and explore to understand the insights, some cases normalization also required to minimize the computation .Then the required parameters are filtered using feature extraction techniques . The time series cleaned data can be trained with Least Square Support Vector Machine and Random forest build model will be tested for its performance , it can be measured with Mean Absolute Error and Mean Square Error.

B. Block Diagram

In this block diagram, we are going to discuss about Data Collection, Data Processing, Feature Extraction, Model building and Training which is shown in *figure 1.1*

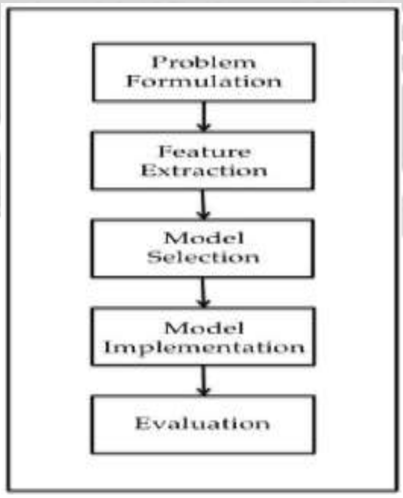


Figure 1.1

C. Flow Chart

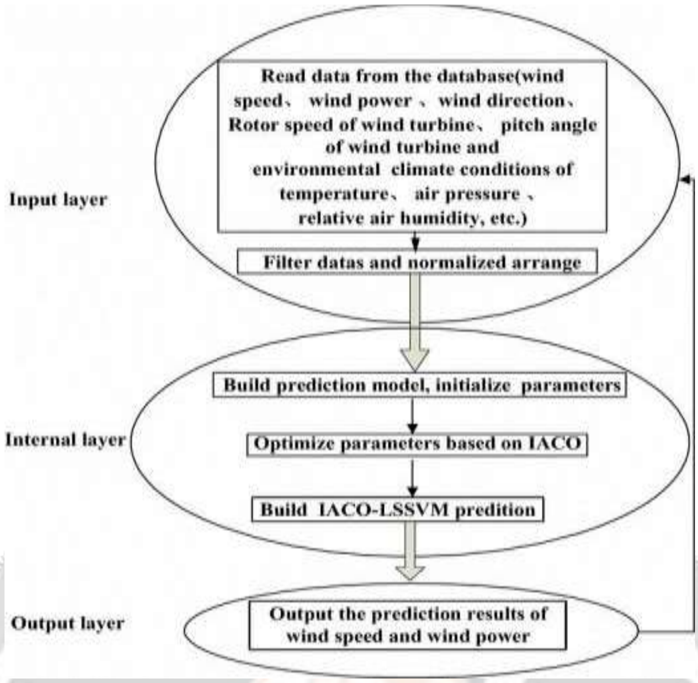


Figure 1.2

D. Performance Evaluation

Whereas most learning algorithms perform acceptably on data sets with tens of thousands of training examples, many important data sets are significantly larger. For example, large retail customer databases and Hubble telescope data can easily involve a terabyte or more. To provide reasonably efficient data mining methods for such large data sets requires additional research. Research during the past few years has already produced more efficient algorithms for such problems as learning association rules and efficient visualization of large data sets. Further research in this direction might well lead to even closer integration of machine learning algorithms into database management systems.

MAE - Mean Absolute Error

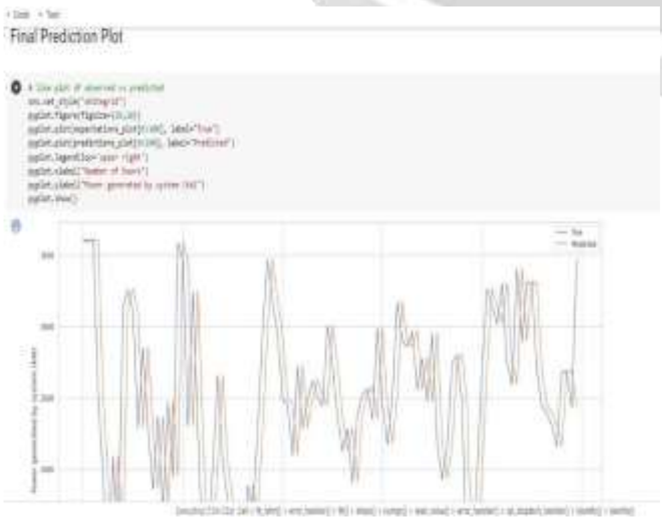
$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

MSE - Mean Squared Error

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

V. RESULT AND CONCLUSION

A. Result



B. *Conclusion*

Wind prediction plays lot of role in wind farm , weather forecating , aviation maritime operations .. In this project wind prediction will be made for wind farm , to improve power generation and minimize damage in turbine. The literature survey has been made based wind power generation and time series data prediction . From literature survey , it is identified the time series data components trends, seasonality and residual has be analyzed and then it it also learned that wind speed plays important role in turbine function to generate energy. short term wind prediction using machine learning algorithms LSSVM , RF was suggested Predictions are quite satisfying if model is trained on just one day. if one day ahead is being predicted. For long term predictions, larger data set would have to be used for training which would include data for all four seasons. This can be used for planning of production and usage of wind turbines, which would significantly decrease problems which occur due to variability of wind. the detection and calculation of down time periods and losses is somehow uncomplicated.

REFERENCES

1. Agrawal, R., Imielinski, T., and A. Swami. Database mining: A performance perspective. IEEE Trans. on Knowl. Data Eng. 5, 6 (2021), 914 925.
2. Chauvin, Y. and Rumelhart, D. Backpropagation: Theory, Architectures, and Applications. Lawrence Erlbaum Associates, Hillsdale, N.J., 2021.
3. Clark, P. and Niblett, R. The CN2 induction algorithm. Mach. Learn. 3, 4 (Mar.2019), 261 284.
4. Gray, J., Bosworth A., Layman A., and Pirahesh, H. Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals. Microsoft Tech. Rep. MSR-TR-95-22, Redmond, Wash., 2019.
5. Heckerman, D., Geiger, D., and Chickering, D. Learning Bayesian networks: The combination of knowledge and statistical data. Mach. Learn. 20, 3 (Sept. 2019), 197 243.
6. Faloutsos C. and Lin, K. FastMap: A fast algorithm for indexing, data mining, and visualization. ACM SIGMOD (2019), 163 174.
7. Mitchell, T. Machine Learning. McGraw-Hill, New York, .
8. Muggleton, S. Foundations of Inductive Logic Programming. Prentice Hall, Englewood Cliffs, N.J..
9. Quinlan J. C4.5 Programs for Machine Learning. Morgan Kaufmann, San Mateo, Calif
10. de Vries EN, Prins HA, Crolla RM, den Outer AJ, van Andel G, van Helden SH, et al. Effect of a comprehensive surgical safety system on patient outcomes. N Engl J Med. 2010;363:1928–1937. [PubMed] [Google Scholar]
11. Reed M, Huang J, Brand R, Graetz I, Neugebauer R, Fireman B, et al. Implementation of an outpatient electronic health record and emergencydepartment visits, hospitalizations, and office visits among patients with diabetes. JAMA. 2013;310:1060–1065. [PMC free article] [PubMed] [Google Scholar]
12. Jaffe MG, Lee GA, Young JD, Sidney S, Go AS. Improved blood pressure control associated with a large-scale hypertension program. JAMA. 2013;310:699–705. [PMC free article] [PubMed] [Google Scholar]
13. Shnorhavorian M, Bittner R, Wright JL, Schwartz SM. Maternal risk factors for congenital urinary anomalies: results of a population-based case-control study. Urology. 2011;78:1156–1161. [PubMed] [Google Scholar]..