Predicting the energy of Wind Turbine based on Weather Condition using IBM Cloud

1. INTRODUCTION

a. Overview:

Wind power is one of the fastest-growing renewable energy technologies. Usage is on the rise worldwide, in part because costs are falling. Global installed wind-generation capacity onshore and offshore has increased by a factor of almost 75 in the past two decades, jumping from 7.5 gig watts (GW) in 1997 to some 564 GW by 2018, according to IRENA's latest data. Production of wind electricity doubled between 2009 and 2013, and in 2016 wind energy accounted for 16% of the electricity generated by renewables. Many parts of the world have strong wind speeds, but the best locations for generating wind power are sometimes remote ones. Offshore wind power offers tremendous potential. Wind turbines first emerged more than a century ago. Following the invention of the electric generator in the 1830s, engineers started attempting to harness wind energy to produce electricity. Wind power generation took place in the United Kingdom and the United States in 1887 and 1888, but modern wind power is considered to have been first developed in Denmark, where horizontal-axis wind turbines were built in 1891 and a 22.8-metre wind turbine began operation in 1897.

b. Purpose:

The purpose of this project is to predict the energy output of a wind turbine given certain environmental factors such as Wind Velocity, Wind Direction, Air Pressure and Air Temperature. With a compiled dataset, training linear regression over the dataset accuracy of 90.5% was achieved. This prediction is more accurate and using these data future estimates about energy output can be made which is necessary for government policies and for statistics.

2.   LITERATURE SURVEY

a. Existing problem:

Stancu[1] says that end-to-end approach based on the CRISP-DM framework offers some guidelines for future projects. Well-tuned ANNs can deliver accurate predictions for forecasting wind energy production but carry costs associated with the re-sources, time, and computational power required to find the optimum hyperparameter combination. Tree-based models provide transparency that is lacking in black box algorithms such as ANN. SVR may have been the best solution pro-posed if a single metric considering both performance and training time had been proposed. Wind power prediction remains an open topic despite the promising results. For the analyzed time frame, no special event or natural hazard occurred; the wind energy production as a dependent variable was consistently associated with the correct independent variables. A key feature of this re-search is the insights it offers into how data pre-processing and the model's optimization were performed.

b. Proposed solution:

A proposal to use linear regression over the dataset containing various environmental factors is made. The resulting model has an accuracy of 90.5%. Further accurate estimation and assumptions can be made after getting the prediction from the resulting model.

3.   THEORETICAL ANALYSIS

a. Block diagram:
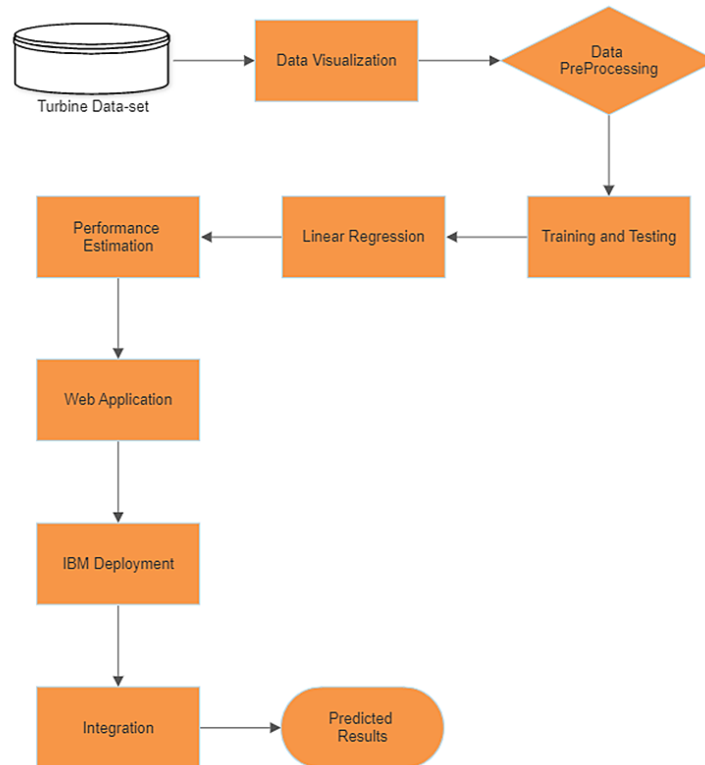
b. **Hardware / Software designing:**

Jupyter Notebook is used as main software.

The required libraries to be imported to Python script are:

Numpy: NumPy is the fundamental package for scientific computing in Python. It is a Python library that provides a multidimensional array object, various derived objects (such as masked arrays and matrices), and an assortment of routines for fast operations on arrays, including mathematical, logical, shape manipulation, sorting, selecting, I/O, discrete Fourier transforms, basic linear algebra, basic statistical operations, random simulation and much more..

Pandas: It is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language.

Matplotlib: Visualisation with python. It is a comprehensive library for creating static, animated, and interactive visualizations in Python.

Seaborn: Seaborn is a library for making statistical graphics in Python. Seaborn helps you explore and understand your data. Its plotting functions operate on dataframes and arrays containing whole datasets and internally perform the necessary semantic mapping and statistical aggregation to produce informative plots.

Pickle: The pickle module implements serialization protocol, which provides an ability to save and later load Python objects using special binary format.

Scikit-learn: Scikit-learn is probably the most useful library for machine learning in Python. The sklearn library contains a lot of efficient tools for machine learning and statistical modeling including classification, regression, clustering and dimensionality reduction. We can also call sklearn as short.

Sklearn.model_selection: train_test_split is a function in Sklearn model selection for splitting data arrays into two subsets: for training data and for testing data. With this function, you don't need to divide the dataset manually. By default, Sklearn train_test_split will make random partitions for the two subsets.

Test_size(): This parameter decides the size of the data that has to be split as the test dataset. This is given as a fraction. For example, if you pass 0.5 as the value, the dataset will be split 50% as the test dataset. If you're specifying this parameter, you can ignore the next parameter.

Spyder: Spyder software is used to build the flask Application which enabled us to connect the model to front end.
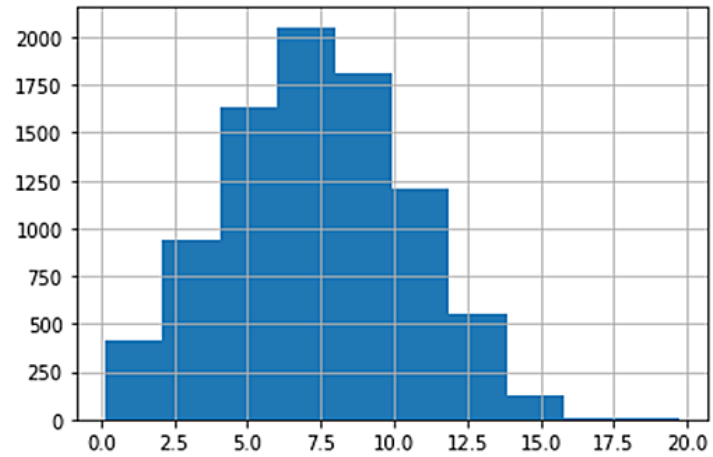
HTML and CSS to create the front-end part of the web page. The application was run by python flask

4.　　EXPERIMENTAL INVESTIGATIONS:

a. **Plotting the Wind Speed to visualize the average Speed:**
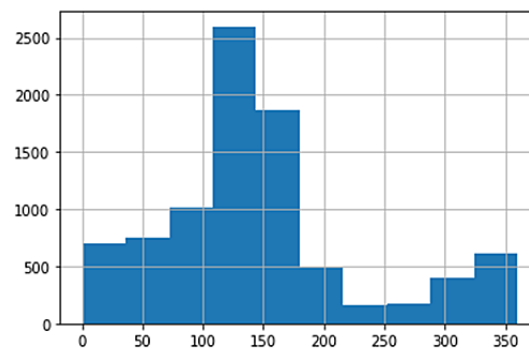
```
In [4]: df['WindSpeed'].hist()
Out[4]: <AxesSubplot:>
```
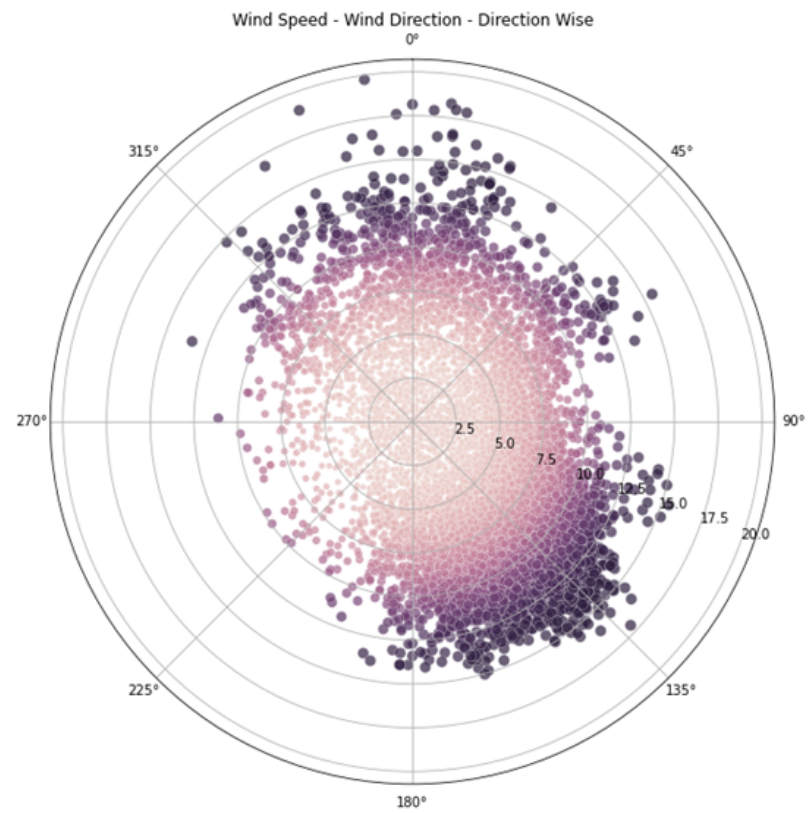

b. **Plotting the Wind Direction (in degrees) to see which direction wind flows the most:**
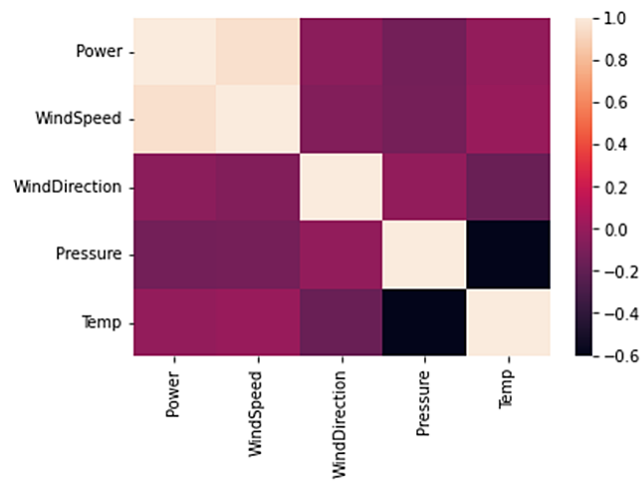
```
In [5]: df['WindDirection'].hist()
Out[5]: <AxesSubplot:>
```
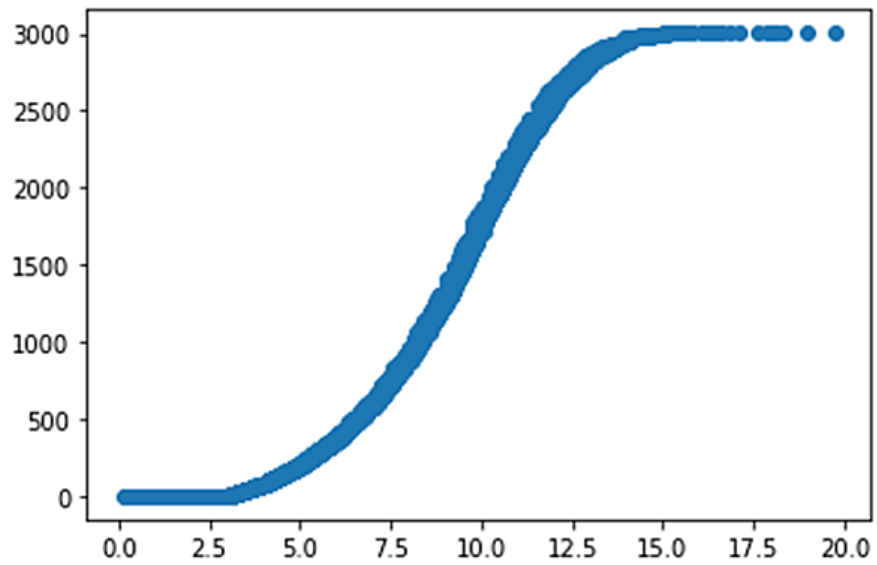

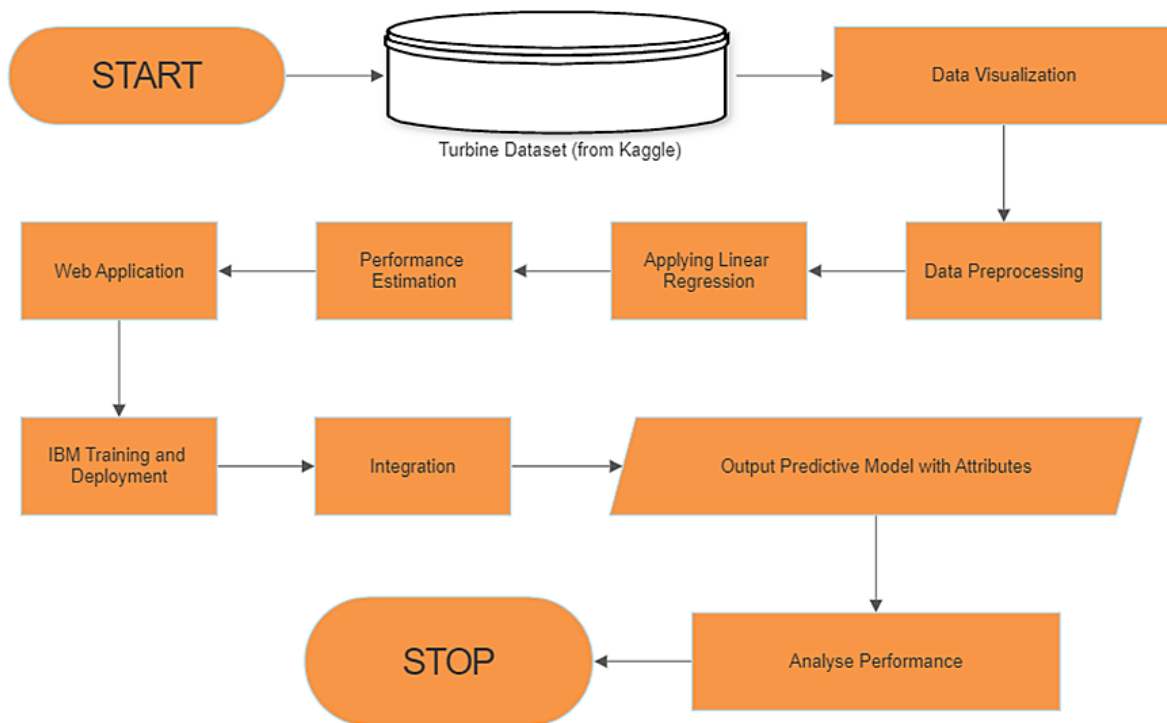c. **Plotting Wind Speed, Direction wise on a radian scale to visualize Wind flow:**

Wind Speed - Wind Direction - Direction Wise

d. Heat map of the correlation factor b/w each column:

b. **Plotting the Power Curve between Wind Speed and Power.**

5. FLOWCHART

6. RESULT:

LINEAR REGRESSION:

According to the training, Linear Regression performed the best with R2 Score of 90.50%

7. ADVANTAGES & DISADVANTAGES

List of advantages and disadvantages of the proposed solution:

A. **Advantages:**

i. It can be used accurately to predict wind energy output of a turbine.

ii. Same output can be used for other scientific data related works or researches (since prediction is very accurate)

B. **Disadvantages:**

i. Features such as rainfall, mechanical factors inside a turbine isn't considered and that can have influence over the prediction

ii. Risk of error associated with even the highest accuracy machine learning algorithms

8. APPLICATIONS

The areas where this solution can be applied:

i) Energy Sector

ii) Further analysis on the design of turbines to make them more efficient

iii) Setting up or discovering geographical locations where a wind turbine can operate effectively.

9. CONCLUSION

The project was aimed at predicting the output energy of wind turbines under environmental factors. The model using Linear Regression had a high accuracy. This machine learning model would be a particular help for the Energy Sector. Since we all are making a change to renewable resources we have to make sure to use the resources very carefully. This model can predict that setting up a wind turbines to produce wind energy on a certain location based on weather data can be profitable or not. Thus reducing costs and optimizing the solution.

10. FUTURE SCOPE

1. In the upcoming future, we will convert our web application to an android application with more enhanced and innovative features.

2. We'll also try to introduce new features like mechanical aspects of Wind Turbine, rainfall and other factors which can affect the prediction thus making our model

11. BIBLIOGRAPHY

References:

[1] Rashid,haider, Batunlu, Forecasting of Wind Turbine output using ANN

[2] Farhad Elyasichamazkoti, Abolhasan Khajehpoor, Application of machine learning for wind energy from design to energy-Water nexus: A Survey, Energy Nexus,

[3] https://www.geeksforgeeks.org/machine-learning/

[4] https://www.geeksforgeeks.org/types-of-regression-techniques/

APPENDIX:

GitHub Link to access the project files and source code:
https://github.com/smartinternz02/SI-GuidedProject-46892-1651812493