PROJECT SCOPE DOCUMENTATION

PROJECT NAME :
SBSPS-Challenge-1207-Predicting-the-energy-output-of-wind-turbine-based-on-weather-condition.

PROJECT MANAGER: SHEVYA SOLANKI **DATE** :20/06/2020

1.PROJECT SUMMERY :

- 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 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. 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. The model obtained for energy prediction gives a very reliable prediction of the energy output for newly supplied weather data.

2.PURPOSE OF PROJECT :

- Wind energy plays an increasing role in the supply of energy worldwide. The energy output of a wind farm is highly dependent on the wind conditions present at its site. 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.
- Developing a time series model to Predict the power output of wind farm based on the weather condition in the site (1Hr prediction to 72Hrs. prediction) Build an application to recommend the Power Grid to suggest the best time to utilize the energy from wind farm.

3.PROJECT REQUIREMENT :

3.1 FUNCTIONAL REQUIREMENT :

- Predicting The Energy Output Of Wind Turbine Based On Weather Condition.

3.2 TECHNICAL REQUIREMENT :

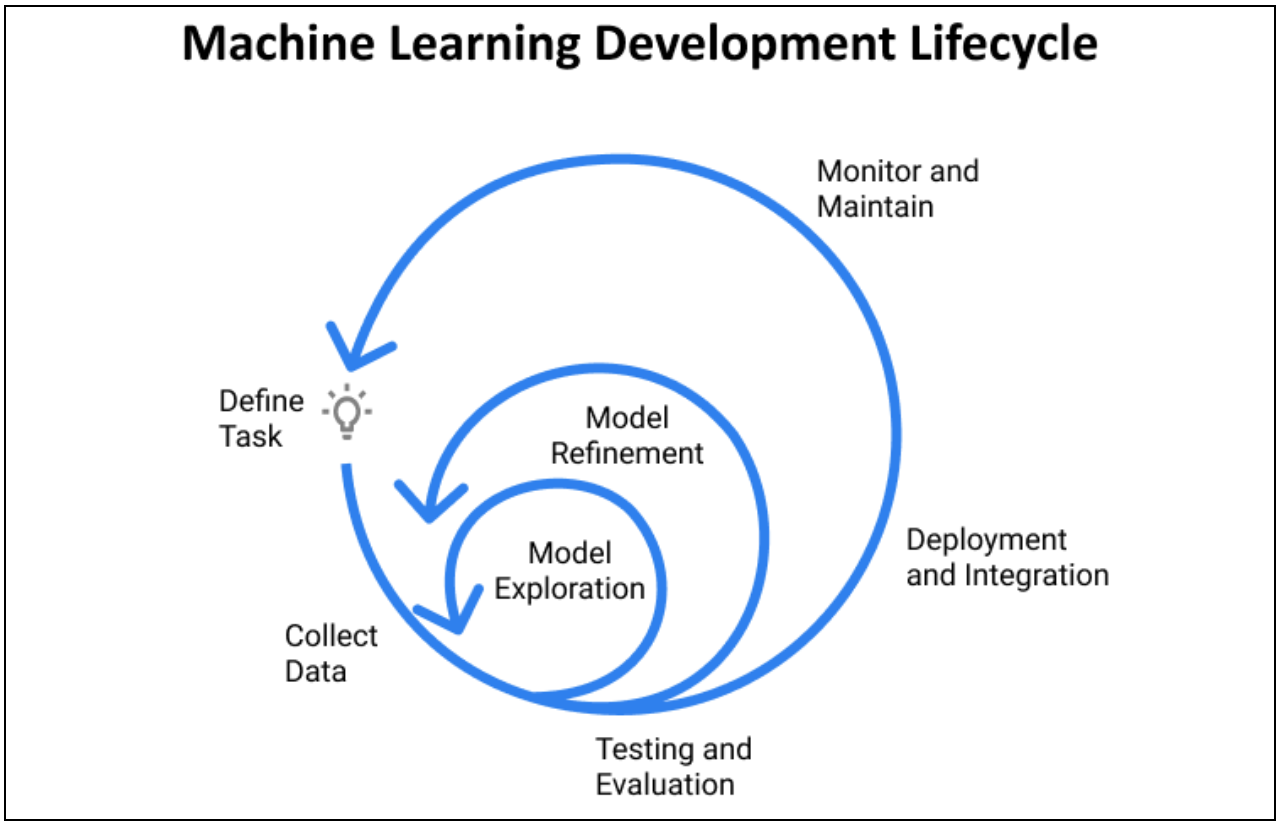
- Python, IBM Cloud, IBM Watson, Github , Node Red in IBM Watson, Jupyter,Notebook Watson,Python 2 or 3, IBM Watson Studio, IBM Cloud for Deployment, Android / any Web frameworks Use Weather Company Data from IBM Cloud.

3.3 HARDWARE REQUIREMENT :

Processor - i5 7th Gen
Speed - 2GHz or more
Hard Disc - 10 Gb or more

4. MAJOR DELIVERABLES :

- The flow chart shown eblow gives the information reharding how the model is built and the descriptiov of the stages it involves. It includes 7 stages as described :



1. **Planning and project setup**

a. Define the task and scope out requirements.
b. Determine project feasibility.
c. Setting up project codebase.

2. **Data collection and labeling :**

https://www.kaggle.com/berkerisen/wind-turbine-scada-dataset
a. Define ground truth (create labeling documentation)
b. Build data model
c. Validate quality of data

TOOLS - Excel SpreadSheets , CSV fifiles.

3. **Model exploration**

a. Establish baselines for model performance.
b. Overfifit simple model to training data.
c. Using various Libraries to explore data.

TOOLS- Matplotlib, Sklearn, Seaborn, Keras, Scikit-Learn.

4. **Model refifinement**

a. Perform model-specifific optimizations (hyper parameter tuning)
b. Iteratively debug model as complexity is added

5. **Testing and evaluation**

a. Evaluate model on test distribution; understand differences between train and test set distributions (how is “data in the wild” different than what you trained on)
b. Revisit model evaluation metric; ensure that this metric drives.

6. **Model deployment**

a. Expose model via a REST API , IBM Watson.
b. Deploy new model to small subset of users to ensure everything goes smoothly, then roll out to all users
c. Monitor live data and model prediction distributions

7. **Ongoing model maintenance**

a. Understand that changes can affect the system in unexpected ways .
b. Periodically retrain model to prevent model staleness
c. If there is a transfer in model ownership, educate the new team

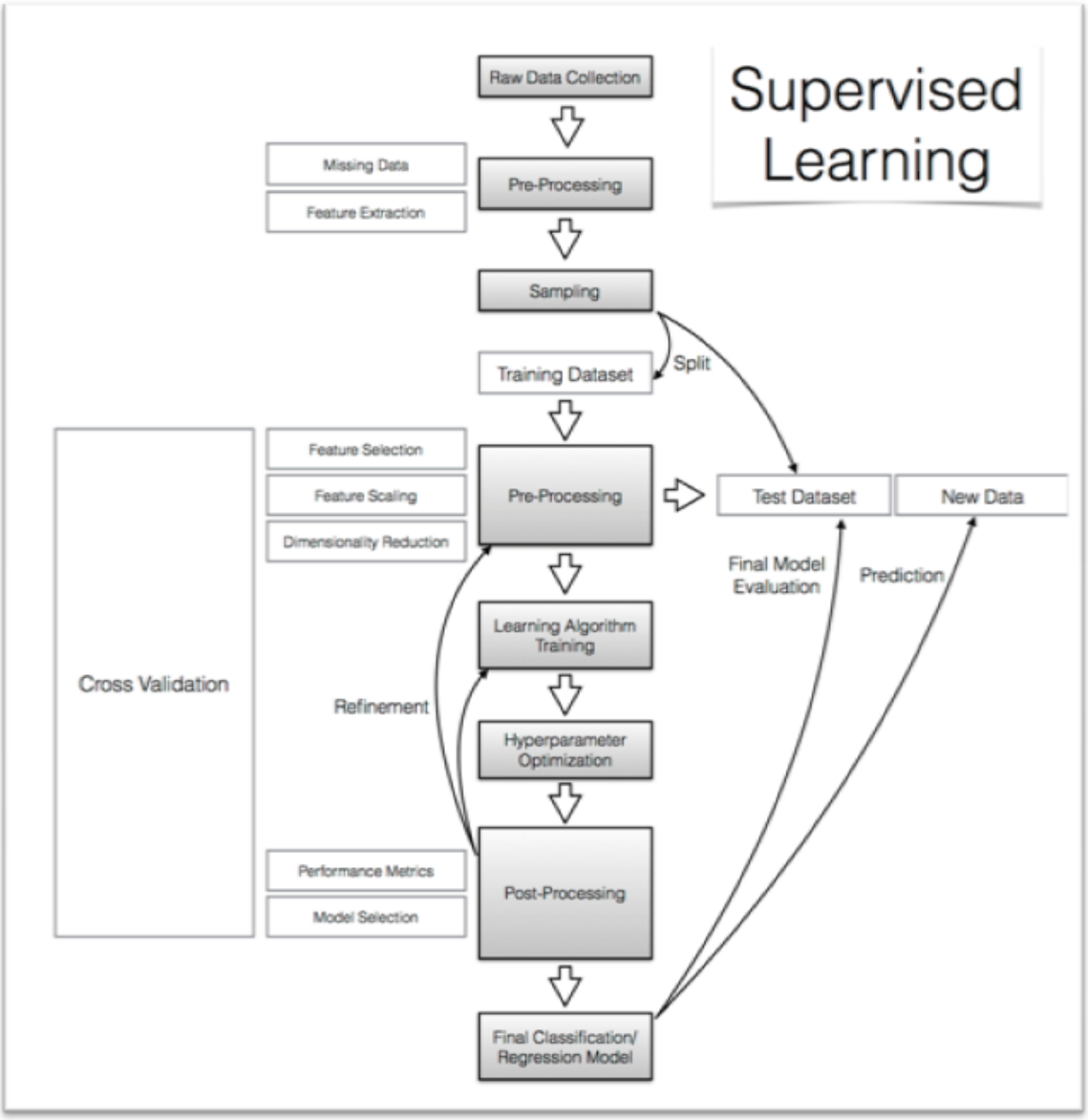
5. ADVANTAGE OF THE PROJECT :

- Renewable energy, such as wind and solar energy, plays an increasing role in the supply of energy world wide. This trend will continue because global energy demand is increasing, and the use of nuclear power and traditional sources of energy such as coal and oil is either considered unsafe or leads to a large amount of CO2 emission

- Wind energy is a key player in the field of renewable energy. The capacity of wind energy production has been substantially increased during the last years. In Europe, for example, the capacity of wind energy production has doubled from 2005 to 2007 [13]. However, levels of production of wind energy are hard to predict as they rely on potentially unstable weather conditions present at the wind farm. In particular, wind speed is crucial for energy production based on wind, and it may vary drastically over time. Energy suppliers are interested in accurate predictions, as they can avoid overproduction by coordinating the collaborative production of traditional power plants and weather-dependent energy sources.

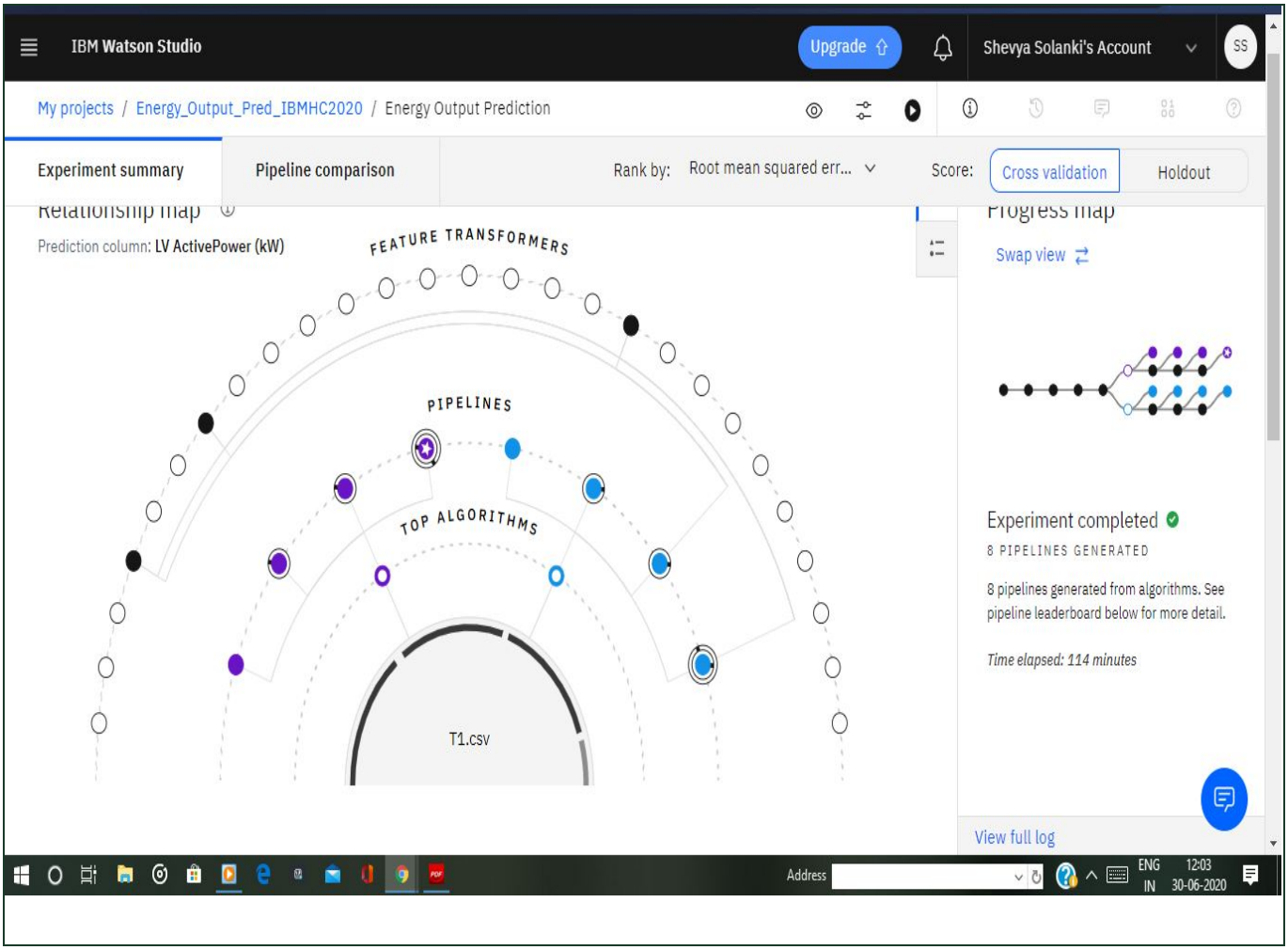
6. FLOWCHART :

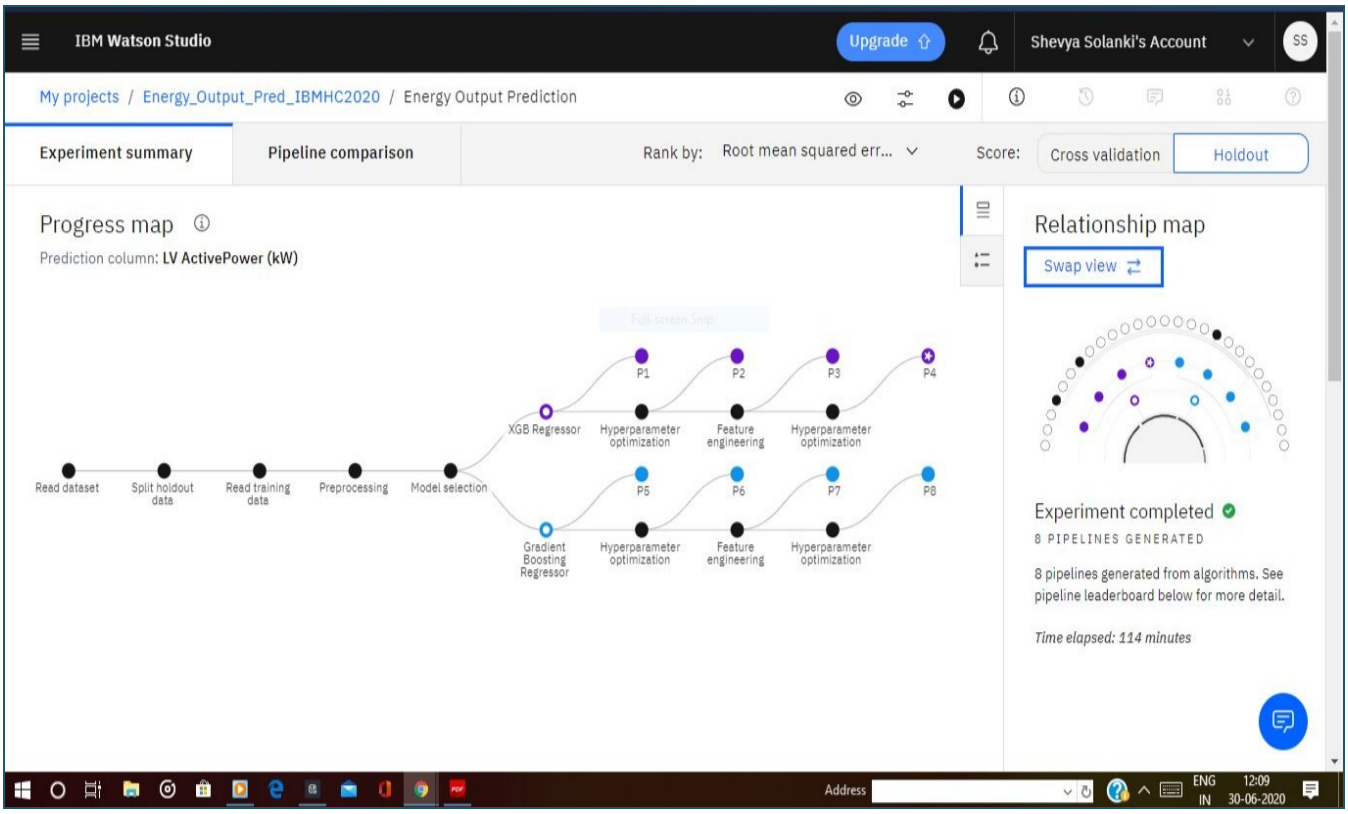
- How does the model actually works is shown by the flowchart below. We will use Regression algorithms to solve or to find solution to the problem statement .Flow chart explains the whole process from start to the end.


Supervised Learning
Raw Data Collection
Missing Data
Pre-Processing
Feature Extraction
Sampling
Split
Training Dataset
Feature Selection
Feature Scaling
Dimensionality Reduction
Pre-Processing
Test Dataset
New Data
Final Model Evaluation
Prediction
Learning Algorithm Training
Cross Validation
Refinement
Hyperparameter Optimization
Performance Metrics
Model Selection
Post-Processing
Final Classification/ Regression Model

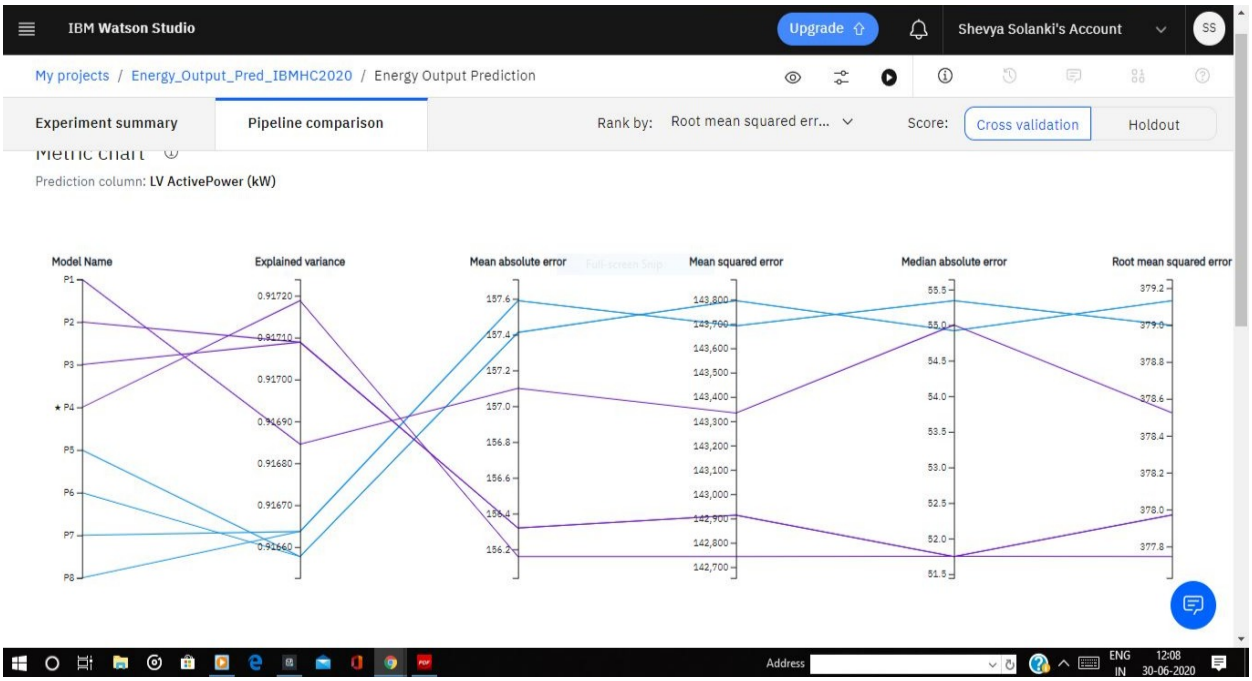
7. RESULT :

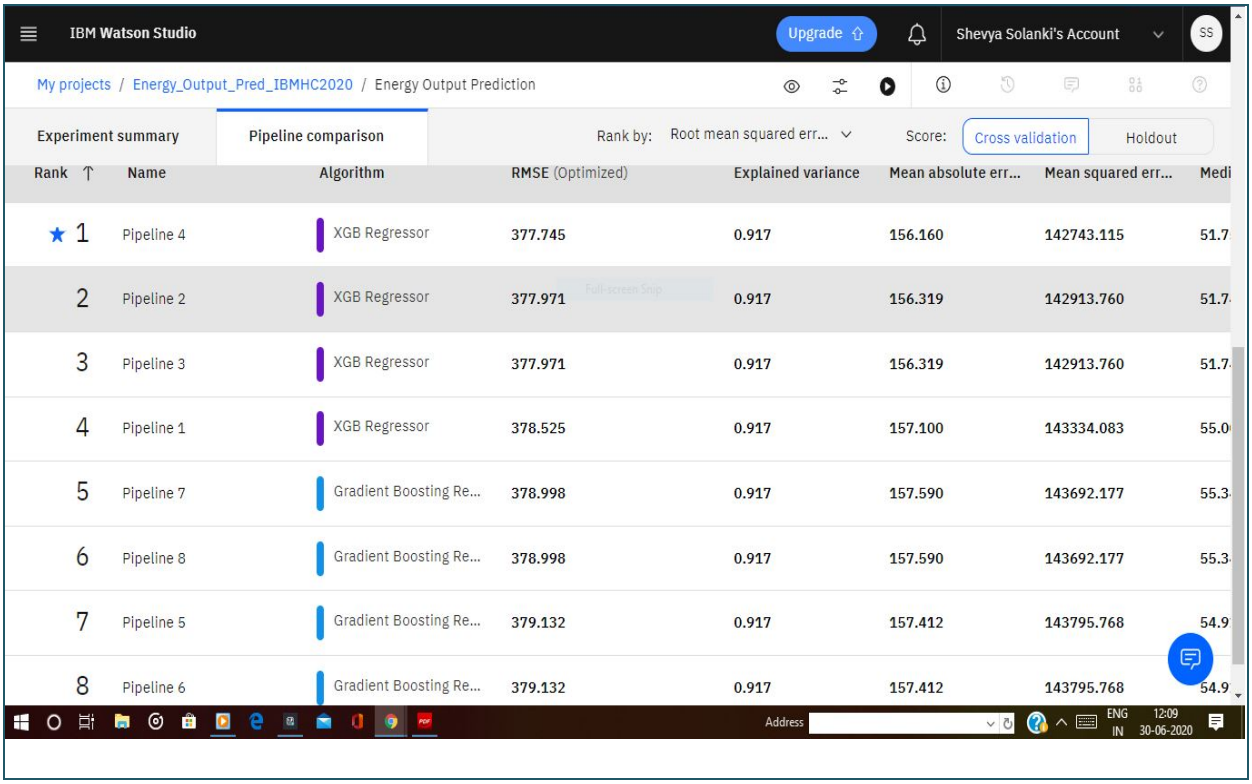
- The project is completed by performing IBM Machine Learning & IBM Watson Studio using Auto-AI experiment which consists of 8 algorithms to be performed out of which the best one is selected for deployment.
- Node-Red starter kit is used to build the UI(User Interface) so that user/customer can enter their details and get the predictions.
- Charts are also used to show the Energy output predictions throug various graphs such as Liner Graph Analysis & Bar graph to show the user the predicted values.



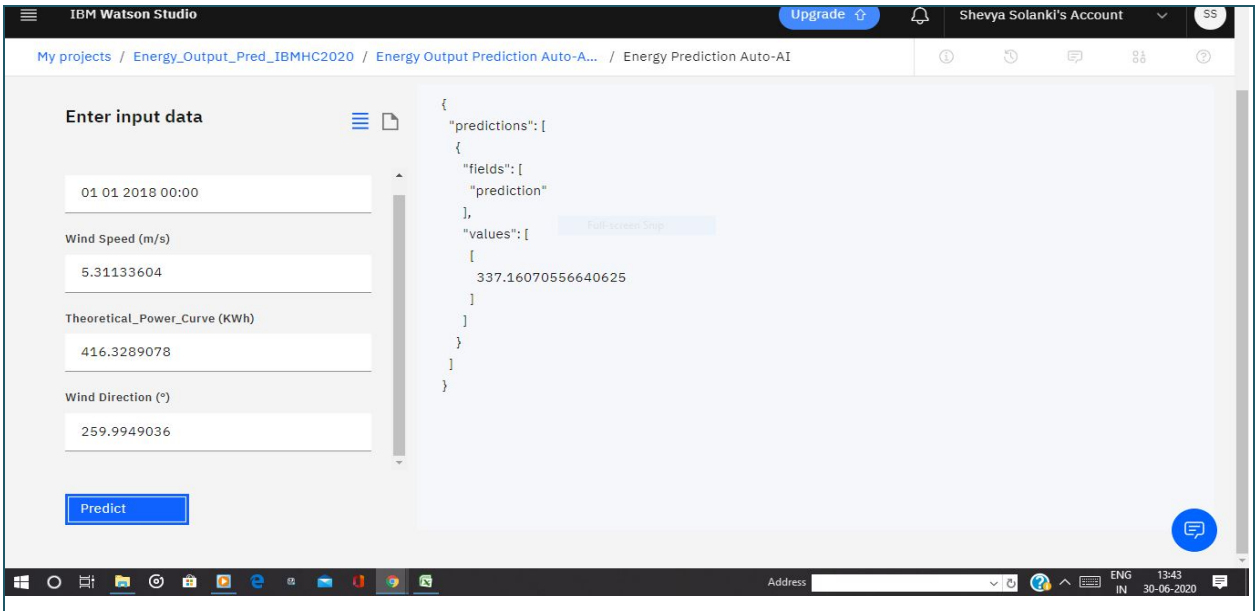


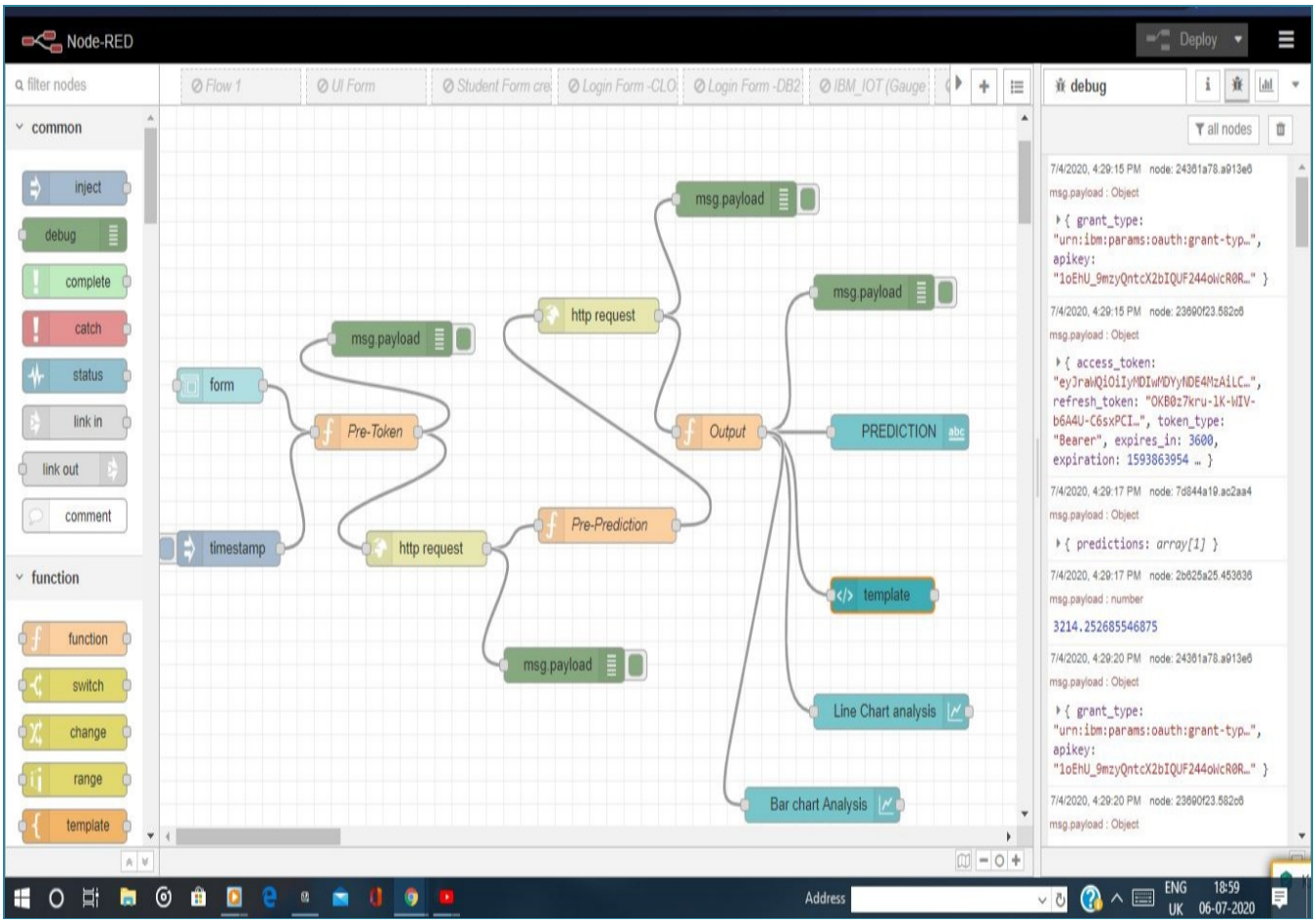
- Above image shown is the Progress Map for the various algorithms performed.



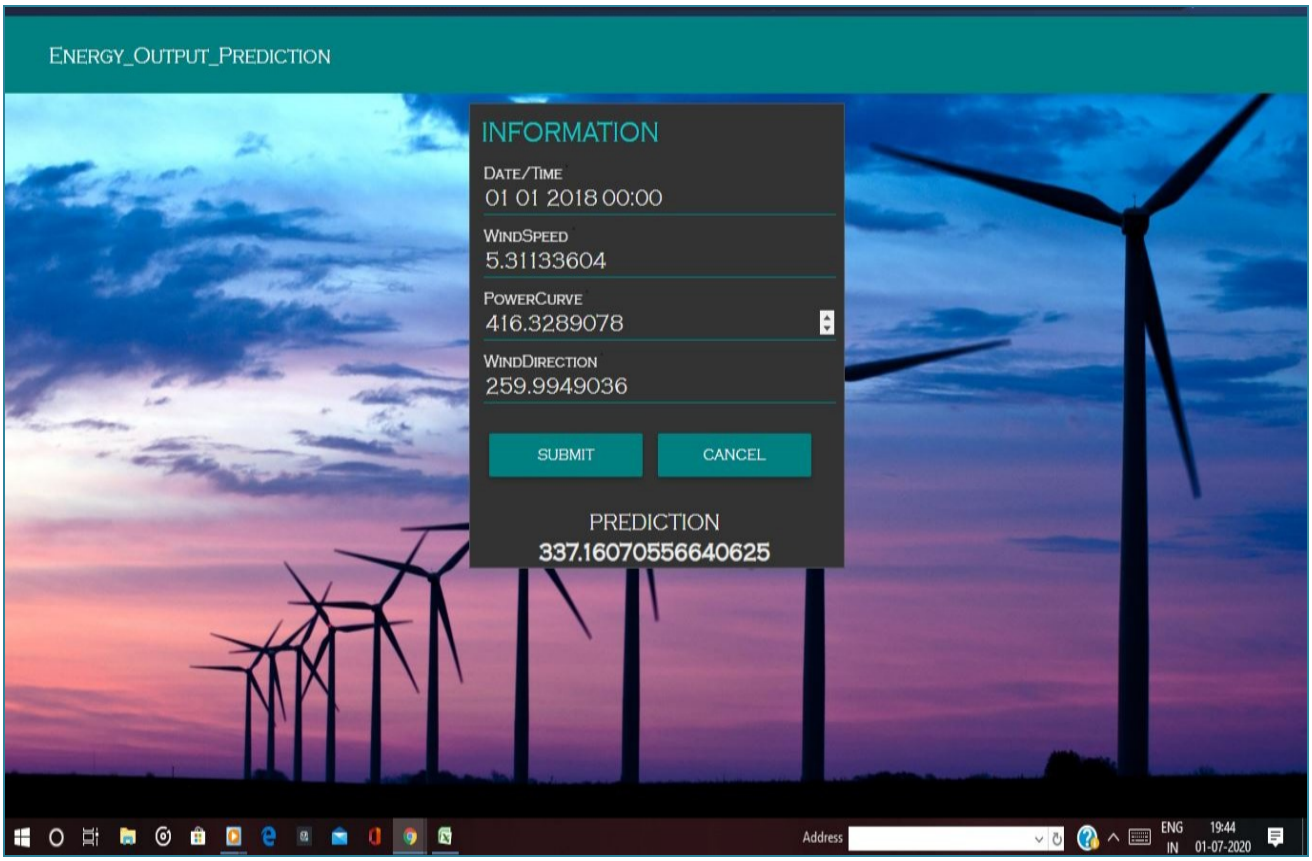


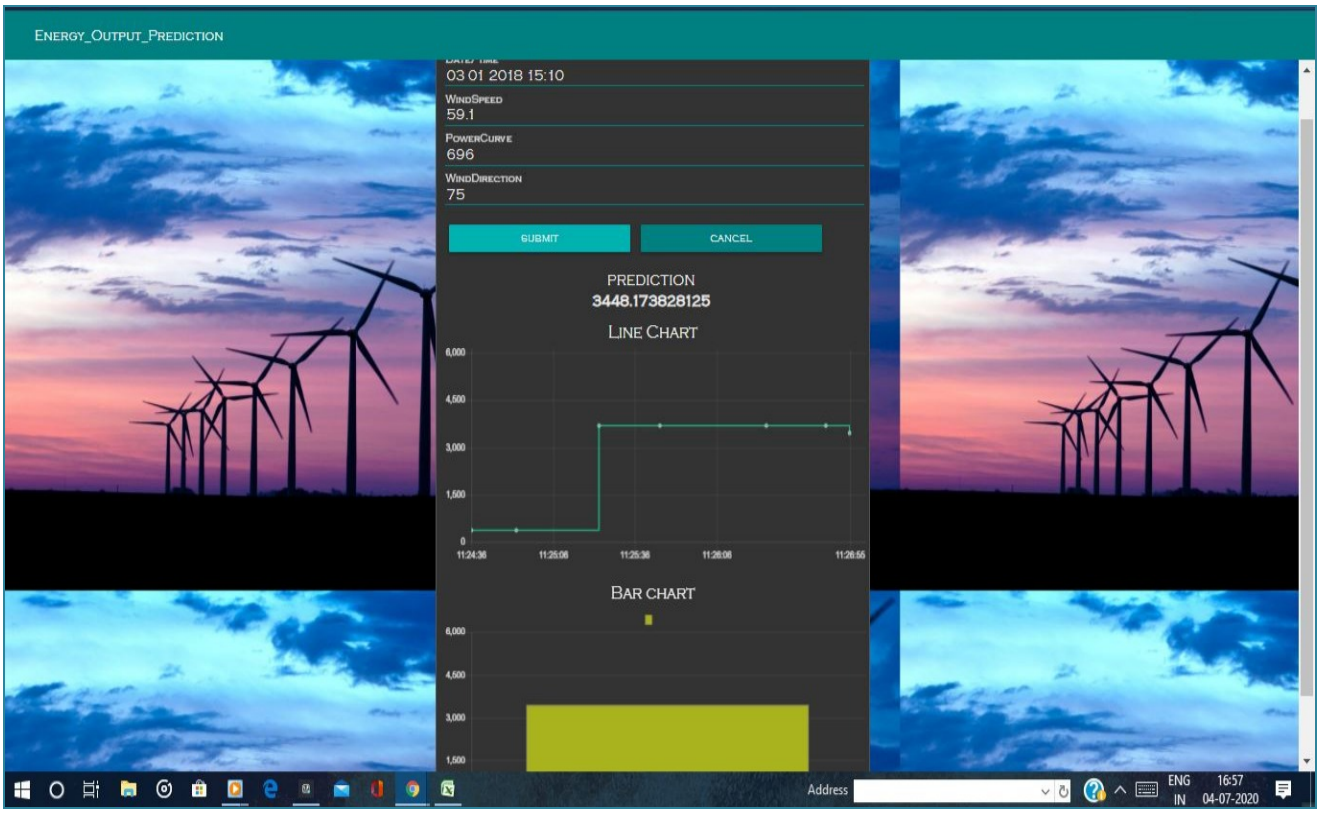
- **Above shown** are the ranking of 8 pipelines performed by IBM Watson Studio Auto-AI, RMSE(Root Mean Square Error) of best pipeline is 377.746 algorithm used is XGB Regressor.





- Above shown is the Node-Red flow to create UI (User Interface) various nodes are connected together to create a flow then deployed to get the predicted value.


ENERGY_OUTPUT_PREDICTION
INFORMATION
DATE/TIME
01 01 2018 00:00
WINDSPEED
5.31133604
POWERCURVE
416.3289078
WINDDIRECTION
259.9949036
SUBMIT
CANCEL
PREDICTION
337.16070556640625
Address
ENG
IN
19:44
01-07-2020


ENERGY_OUTPUT_PREDICTION
03 01 2018 15:10
WINDSPEED
59.1
POWERCURVE
696
WINDDIRECTION
75
SUBMIT
CANCEL
PREDICTION
3448.173828125
LINE CHART
6,000
4,500
3,000
1,500
0
11:24:36
11:25:06
11:25:36
11:26:06
11:26:55
BAR CHART
6,000
4,500
3,000
1,500
Address
ENG
IN
16:57
04-07-2020

Above shown is the UI where user will submit their details and get the predicted value as output , as well as graphs are used to show the predicted value through the graph for better conclusion.