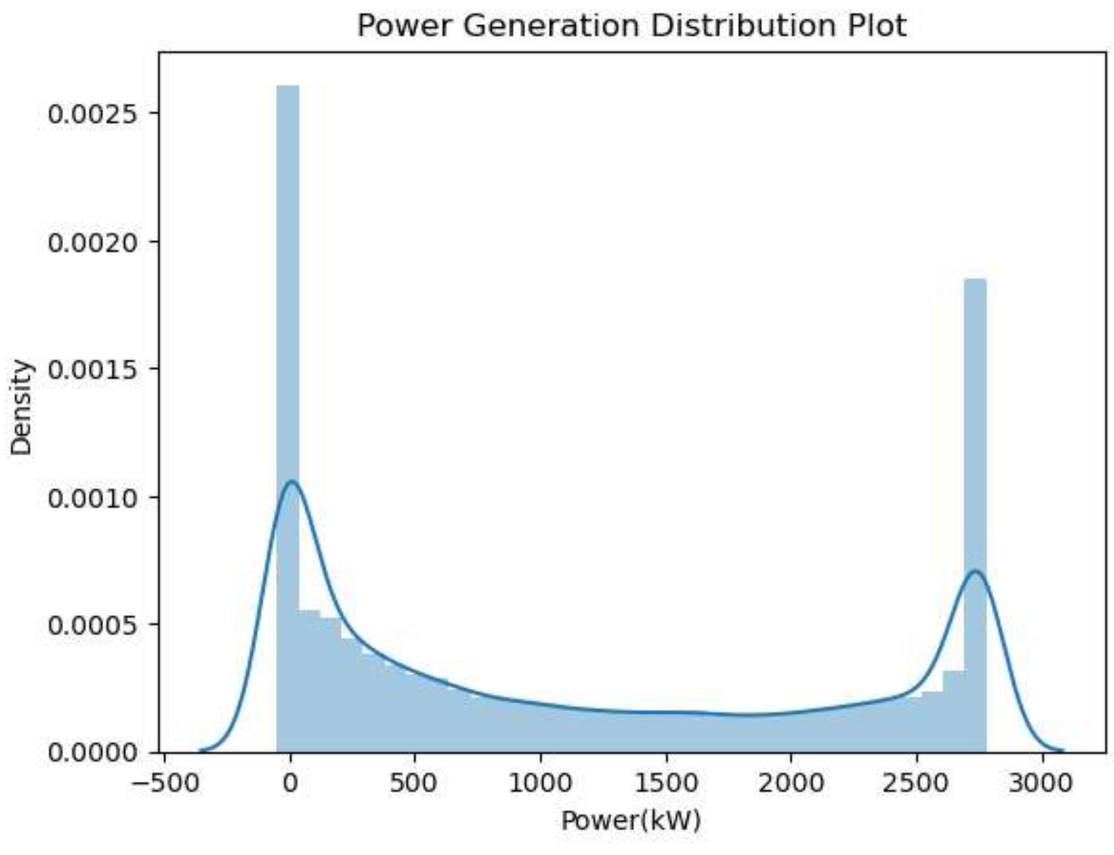
Renewable Energy AI-ML Case Study

Auto-ML is playing an increasingly important role in the renewable energy industry. With the growth of renewable energy sources such as solar and wind power, there is a need for advanced data analytics and predictive modeling to optimize their performance and reduce costs. Auto-ML algorithms can analyze large volumes of data from renewable energy systems to predict power output, detect faults, and optimize performance.

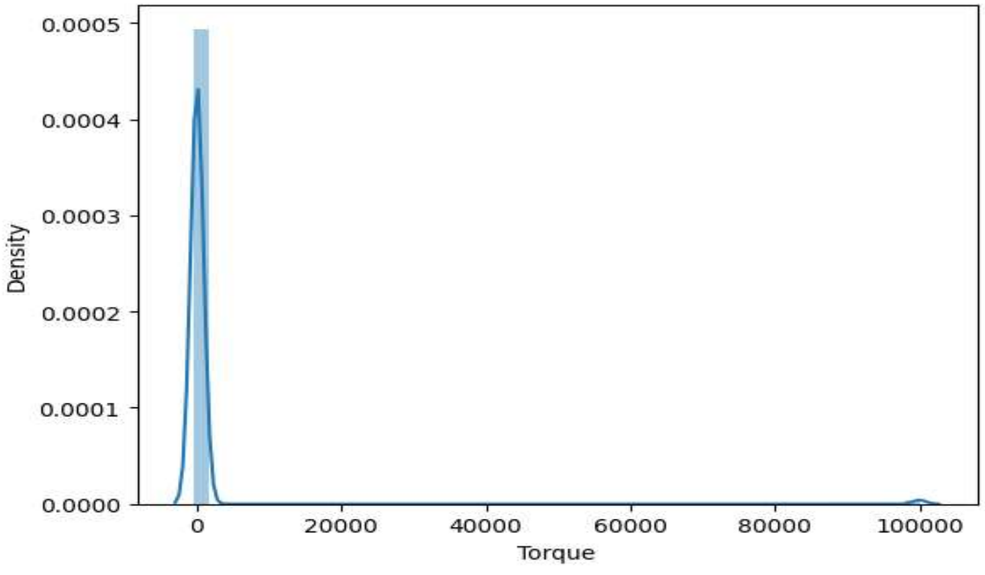
The data set consists of real-time SCADA data. Each data value belongs only to the relevant time period and the input variables transmitted in the data set for the time period to be predicted are prepared to be used to predict the power generation result in the same time period. In the shared dataset, the real-time power generation amount (Power(kW)) of a wind turbine between 01.01.2019 and 14.08.2021 is given on a 10-minute basis.

The aim of this project is to predict the Power Generation and identify factors that optimize Power Generation using Auto-ML.

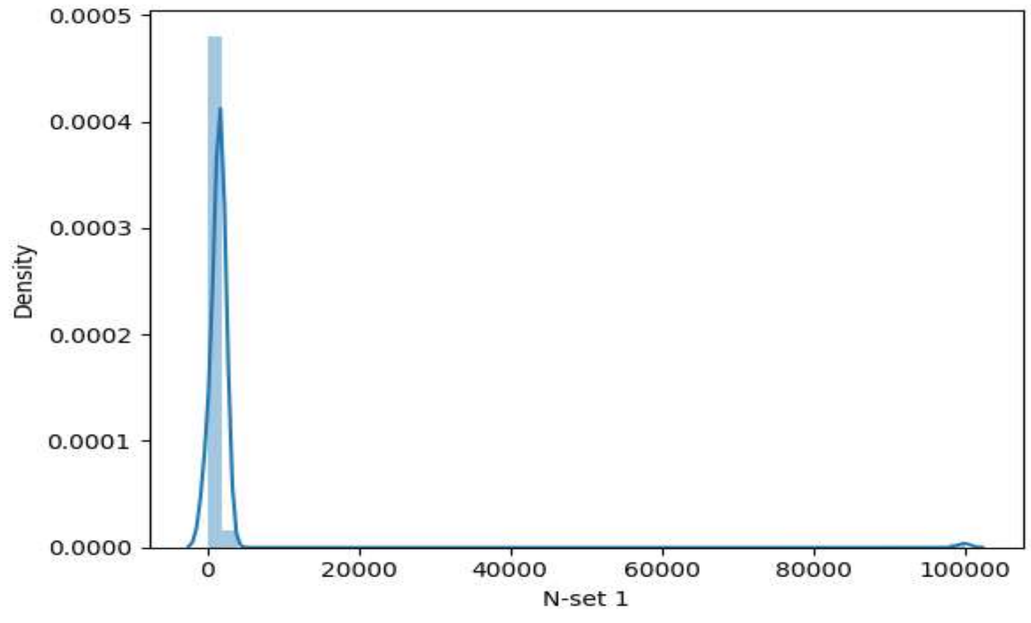


- The business feature used here is Power (KW).
- It is the real time power generation amount of a wind turbine.
- The distribution plot shows that most of the Power(kW) is from -25 to 2750.

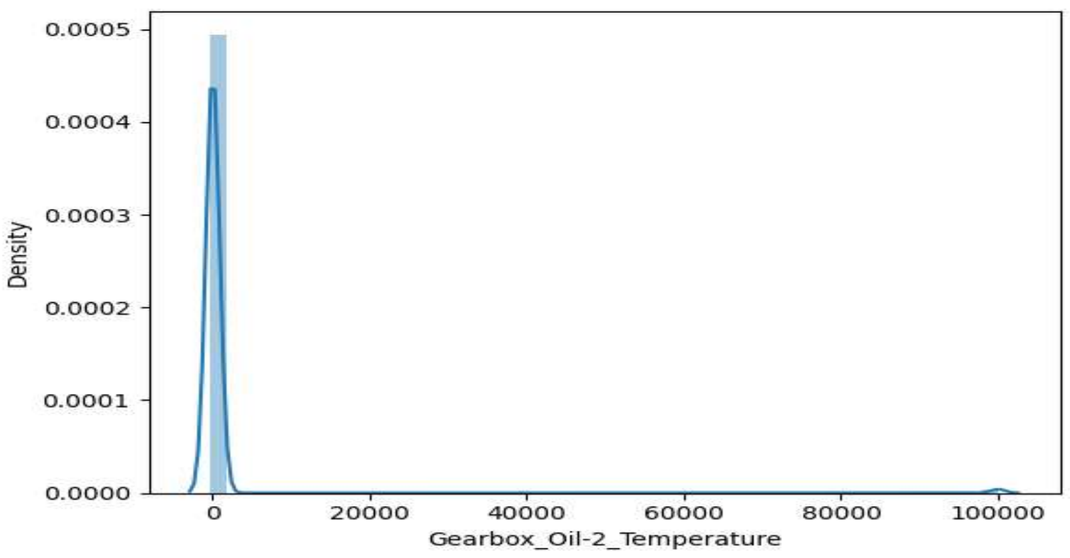
Features Responsible



- **Torque :** The feature is right skewed, 0 to 100 range being the most frequent value.
- Torque is the twisting force that is exerted on the turbine rotor by the wind. It is directly proportional to the force exerted by the wind on the rotor blades and the distance from the center of the rotor to the point where the force is applied.



- **N-set 1 :** The feature is right skewed, 0 to 100 range being the most frequent value.
- The N-set motor, or the generator motor, is responsible for converting the rotational energy of the wind turbine rotor into electrical energy.



- **Gearbox_Oil-2_Temperatue :** The feature is right skewed, 0 to 100 range being the most frequent value.
- The temperature of the gearbox oil is critical in maintaining proper lubrication and preventing wear and tear of the gearbox components. If the gearbox oil temperature is too low, the oil may become too viscous, leading to inadequate lubrication and increased wear on the gearbox components.

Auto-ML Methodology Results

Algorithms	Test Accuracy (25 percentile)	Test Accuracy (50 percentile)	Test Accuracy (75 percentile)	Test Accuracy (90 percentile)
Lasso	72.3	72.7	72.8	72.8
Random Forest	99.3	99.5	99.5	99.5
XGBoost	99.8	99.9	99.9	99.9
MLP	71.3	93.5	98.31	65.8
RNN	79.73	75.3	71.8	74.1
Total Features	20	40	60	72
Avg. Accuracy	84.48	88.18	84.46	82.42

- **Based on our observation from the standard ML algorithms, 50th percentile has the best average accuracy.**
- **XGBoost was the best performing algorithm with 99.9% accuracy in 50, 75 and 90 percentile.**

Conclusion

Auto-ML are solving the Renewable Energy industry in several ways. One of the main challenges of renewable energy sources is their intermittency. Auto-ML algorithms can be used to analyze historical data and weather patterns to predict the amount of energy that will be generated from renewable sources, allowing grid operators to better plan and manage energy supply. Auto-ML algorithms can be used to optimize the performance of renewable energy systems, such as wind turbines or solar panels. The dataset has 136,730 records with 1 Categorical Features and 77 Numerical Features.

For regression, models were created with algorithms using Auto-ML techniques like Lasso, Random forest, XGBoost, Multilayer Perceptron and Recurrent Neural Network. With these models, performance measurement values were obtained for feature sets of 20, 40, 60 and 72. The Auto-ML algorithms were able to predict the power generation with an average accuracy between 82% – 89% and helped to identify factors that determine the power generation. The major factors include Torque, N-set 1 and Gearbox_Oil-2_Temperatue. The Random forest with 99.5 % accuracy in 90th percentile where tree showed a threshold of Torque > 51.39 units, Pitch Demand Baseline_Degree > 0.9 units and N-set 1 > 1558.88 rpm which leads to highest Power Generation.

In summary, data science and Auto-ML are transforming the Renewable energy industry by enabling proactive maintenance, reducing downtime, improving efficiency, and reducing costs.