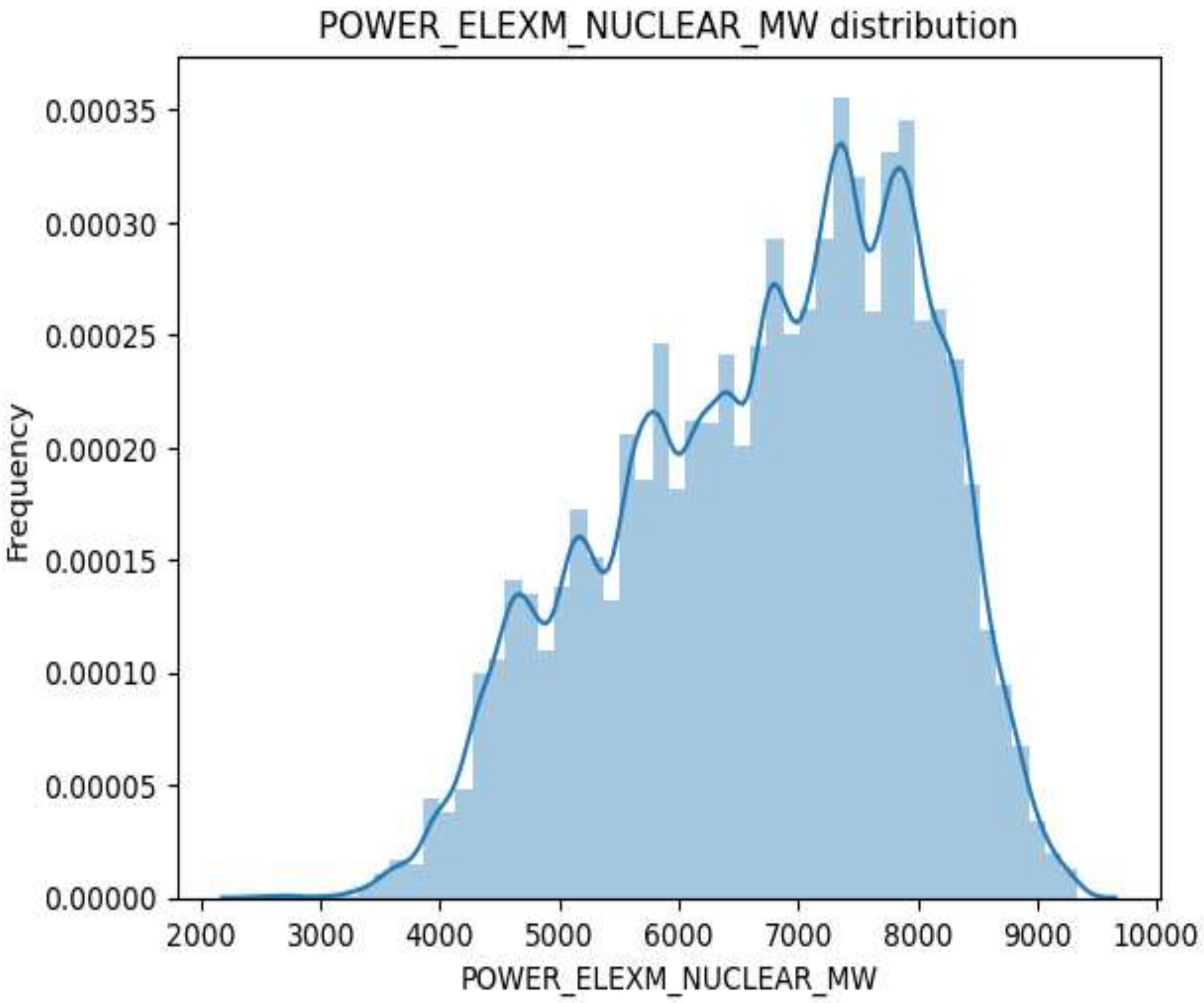
# Nuclear Energy AI-ML Case Study

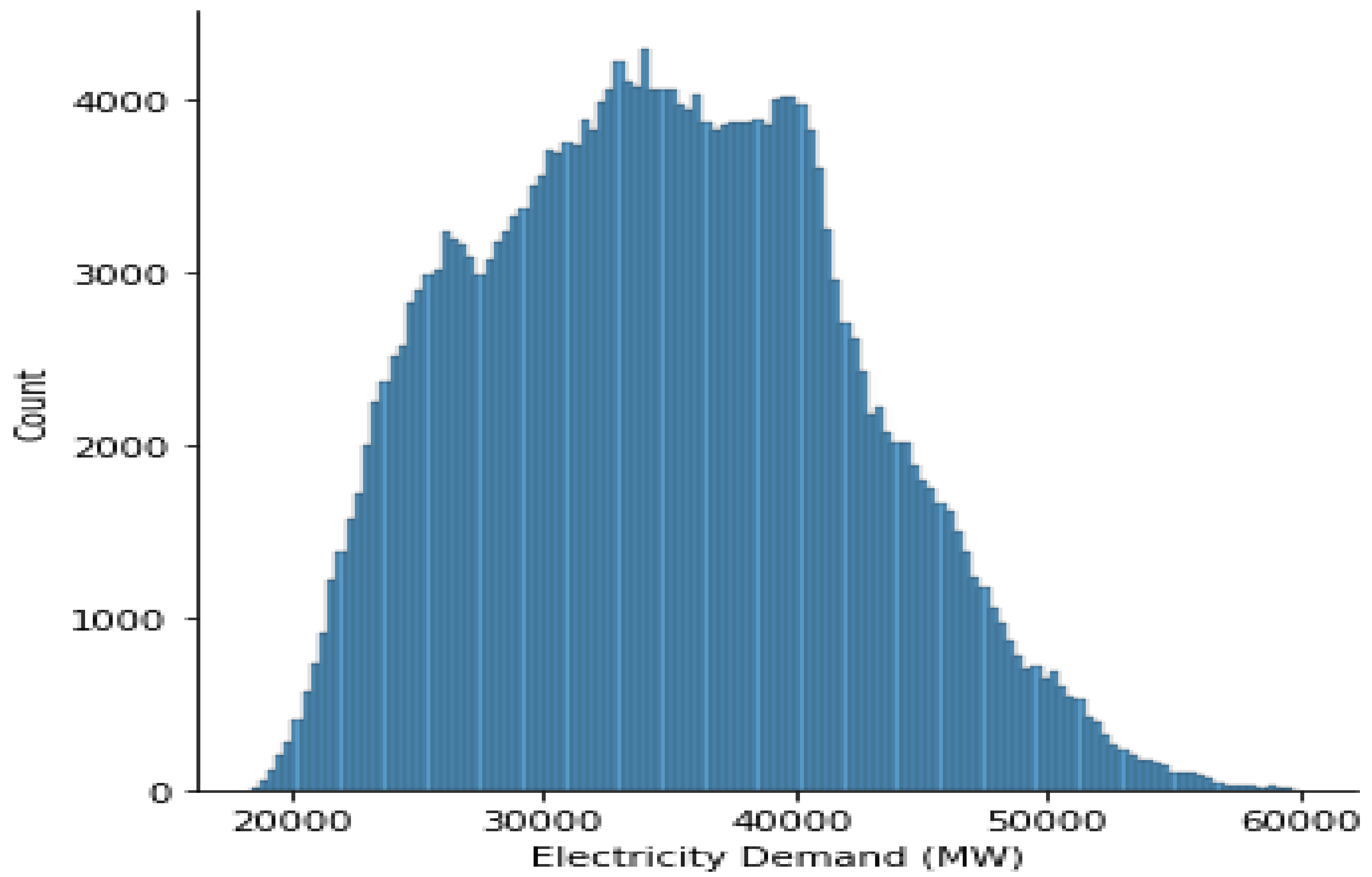
This dataset contains a wealth of information on the electrical half-hourly data for Great Britain from 2008 up until present day. This dataset is sourced from both the Elexon Portal and National Grid, providing you with an in-depth view into electricity supply and demand in the UK. It includes conventional generation, wind generation, nuclear generation, pumped storage and imports & exports. With columns such as ELEXM_SETTLEMENT_DATE, ELEXM_SETTLEMENT_PERIOD, ELEXM_UTC etc., this dataset is ideal for anyone looking to gain a truly comprehensive understanding of current energy situation in Britain.

The main aim of this project is to predict future energy consumption, production and pricing and identify the factors that influence energy consumption using AI/Auto-ML.

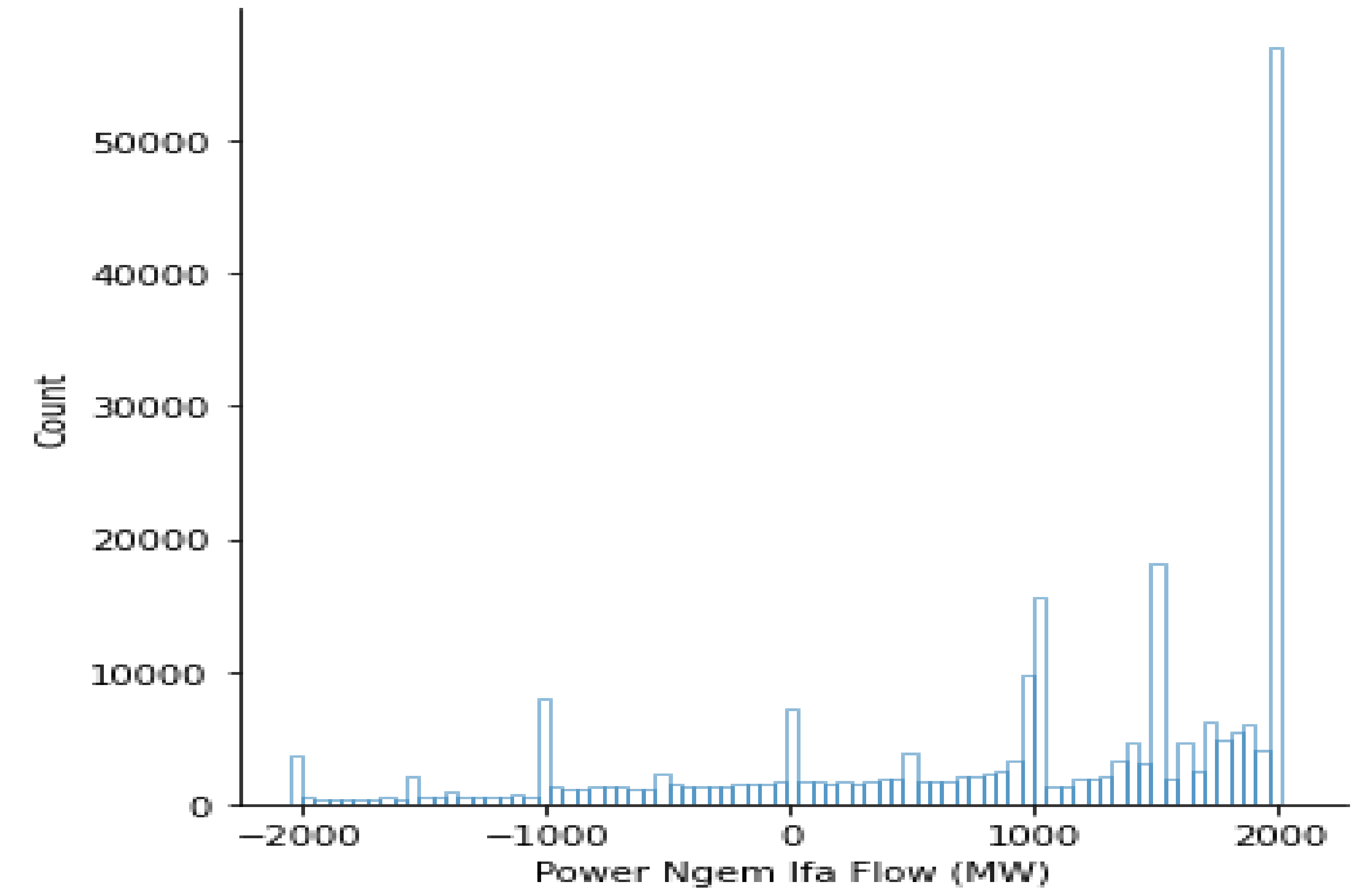



- The business feature used here is POWER_ELEXM_NUCLEAR_MW.
- It is the Nuclear electricity generation in MW unit
- The distribution plot shows nuclear electricity generation from 3000-9000.

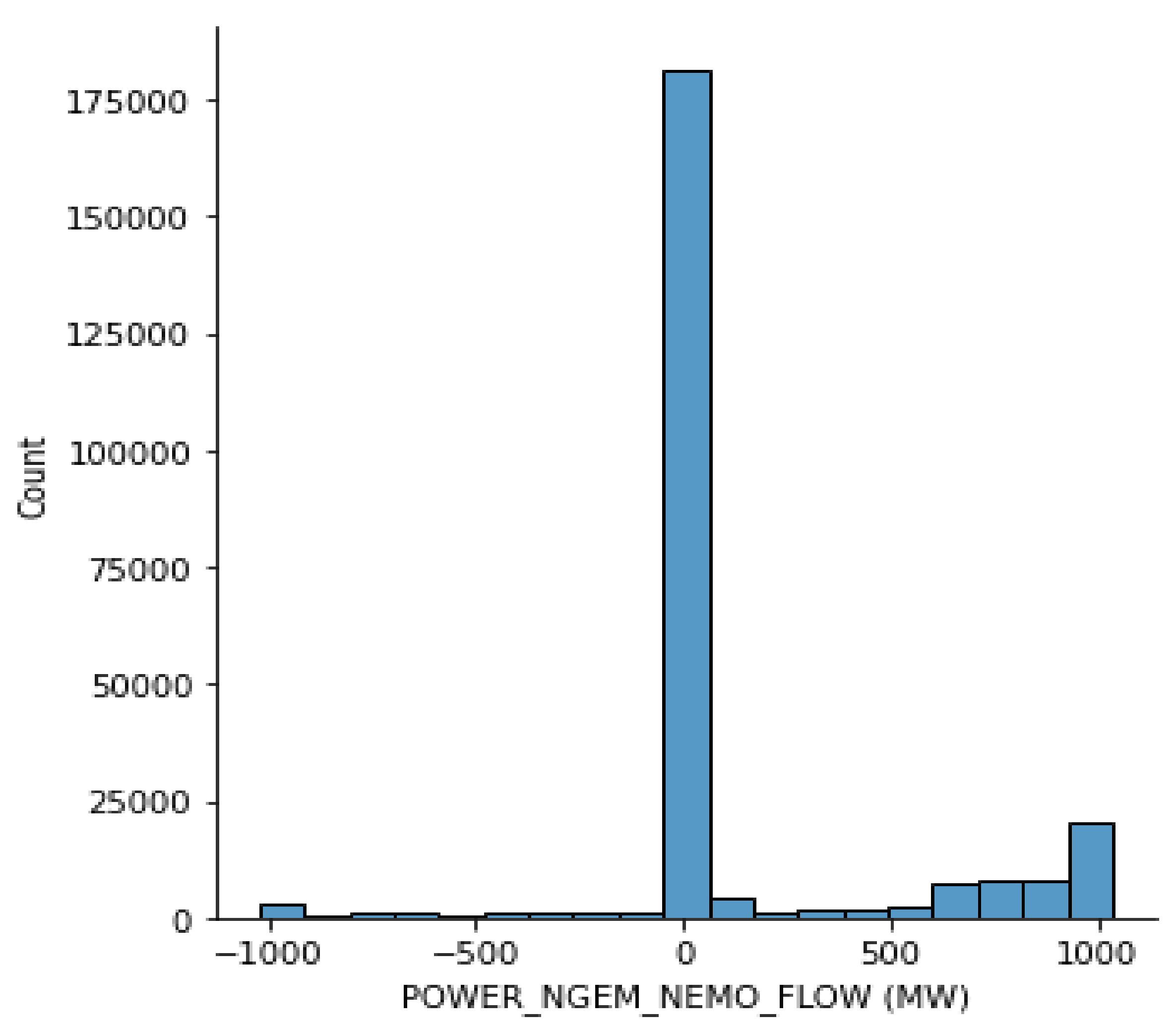
# Features Responsible



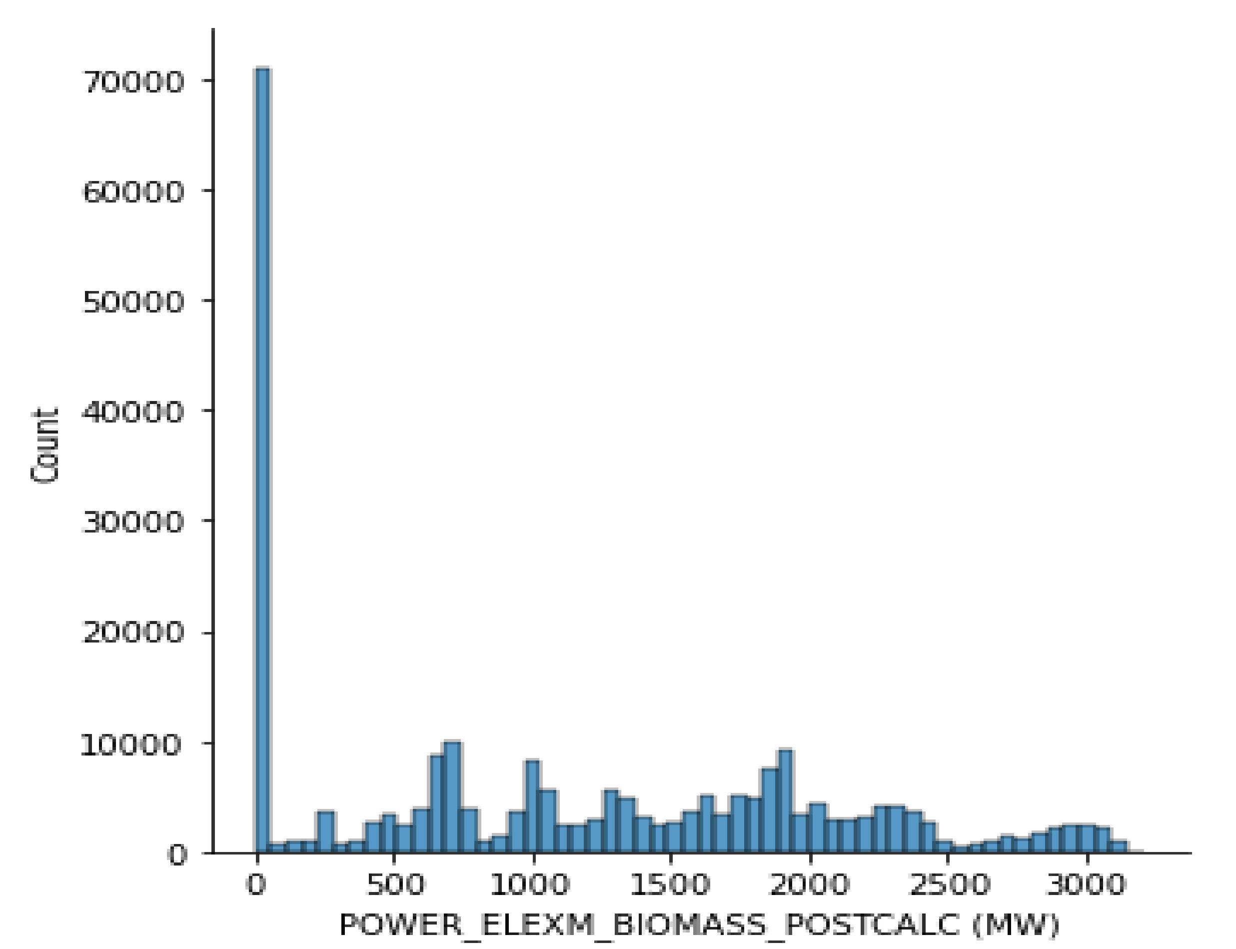

- **Electricity Demand =** For the electricity demand around 35,000 MW, the most nuclear energy has been generated.




- **Power Ngem Ifa Flow =** For Power-Ngem-Ifa-Flow value > 1395 the nuclear power generation is more.




- **POWER_NGEM_NEMO_FLOW (MW) =** The Power-Ngem-Nemo-Flow values closer to zero lead to more nuclear energy generation.




- **POWER_ELEXM_BIOMASS_POSTCALC =** The Power-Elexm-Biomass-Postcalc values less than 1990 units lead to the highest nuclear energy generation.

# Auto-ML Methodology Results

| Algorithms | Test Accuracy (25 percentile) | Test Accuracy (50 percentile) | Test Accuracy (75 percentile) | Test Accuracy (90 percentile) |
|---|---|---|---|---|
| **Lasso** | 22.4 | 34.5 | 55.5 | 40.2 |
| **Random Forest** | 78.2 | 99.8 | 89.8 | 91 |
| **XGBoost** | 69.5 | 76 | 85.8 | 88 |
| **MLP** | 42.2 | 61.9 | 76.7 | 99.1 |
| **RNN** | 42.1 | 48.3 | 45.9 | 40.5 |
| **Total Features** | 5 | 10 | 15 | 18 |
| **Avg. Accuracy** | 50.88 | 64.1 | 70.74 | 71.76 |

- **Based on our observation from the standard ML algorithms, 90th percentile has the best average accuracy.**
- **Random Forest was the best performing algorithm with 99.8%  accuracy in 50 percentile.**

# Conclusion

Auto-ML plays a crucial role in the nuclear energy industry. Data science algorithms can be used to monitor and analyze sensor data from nuclear power plants to detect anomalies and potential safety issues. This can help prevent accidents and ensure that the plants operate safely. The dataset has 246,677 records with all 22 features as Numerical.

For regression, models were created with algorithms using Auto-ML techniques like Decision Tree, Random forest, XGBoost, Multilayer Perceptron and Recurrent Neural Network. With these models, performance measurement values were obtained for feature sets of 5, 10, 15 and 18. The Auto-ML algorithms were able to predict the future energy consumption with an average accuracy between 55% – 67% and helped to identify factors that determine the future energy consuption. The major factors include Electrcity Demand, Power Ngem Ifa Flow, POWER_NGEM_NEMO_FLOW (MW) and POWER_ELEXM_BIOMASS_POSTCALC. The Random forest with 99.8 % accuracy in 50th percentile where tree showed a threshold of POWER_NGEM_NEMO_FLOW_MW <= -0.5 units and POWER_ELEXM_BIOMASS_POSTCALC_MW >= 2263.5 units which leads to highest future energy consumption.

Overall, Auto-ML plays a key role in simulating and modeling nuclear reactions and power generation processes. These simulations can help improve reactor design, optimize fuel usage, and predict the behavior of nuclear materials.