**Capstone Project**

On

Appliance Energy Prediction



Submitted by

Sabitendu Mahapatra

Email id- [sabitendumahapatra@gmail.com](mailto:sabitendumahapatra@gmail.com)

GitHub linkhttps://github.com/sabitendu/Capstone-Project-On-Appliance\_Energy\_Prediction.git

Project Summary

Energy is critical to economic and social growth. It enables nations to achieve their aim of higher living conditions.

So, these days, planning and operating energy production and consumption is a must. Understanding how energy is used is necessary for better management. This presented an opportunity to create a supervised machine learning method to forecast Appliance Energy use.

The initial phase, based on our dataset, was data prepressing. I understood the data and discovered that there are no null values in the dataset, and I received a detailed description of the characteristics involved.

Following that, Exploratory Data Analysis and Data Visualisation produced a concise knowledge of the link between features and label, i.e., the dependent variable it has also given an idea of the features to be chosen for the next step. The heatmap was utilised to understand the association between independent variables, based on which important features were chosen. Choosing the right elements to improve accuracy was difficult.

Standardisation was a critical step before fitting the model since it ensured that the feature values in the data had a zero mean and unit variance. When we construct any Machine Learning method, there is a chance that the objective function will not perform properly unless it is normalised.

Following completion of training and assessment. Linear, Lasso, Ridge, Elasticnet, Gradient Boosting, Random Forest, and XGBoosting techniques were employed. We checked and compared numerous matrices and determined that the Random Forest Regressor produces the best results.

I used the ExtraTreeRegressor and used Hyperparameter Tuning to improve accuracy and prevent overfitting. I also used Gridsearch cross validation to find the optimal parameter. These parameters improved our model's prediction performance.

Finally, we performed Model Explainability using ELI5, which assisted us in understanding the participation of features and their impact on the target variables.

To summarise, the ExtraTreeRegressor has proven to be the best model for our dataset.