

Capstone Project Submission

Name - Ajit Kumar Toppo
Email - ajitkumartoppo@gmail.com

Github Link:- ([appliance_energy_prediction](https://github.com/iam-ajitkumartoppo/appliance_energy_prediction))
(https://github.com/iam-ajitkumartoppo/appliance_energy_prediction)

Drive Link:-
https://drive.google.com/drive/folders/1dAGdB_3-urEVsUpYKzzAFf4gXieoZe2j?usp=sharing

https://drive.google.com/drive/folders/1dAGdB_3-urEVsUpYKzzAFf4gXieoZe2j?usp=share_link

Appliance Energy Prediction

context:

Energy prediction is important to predict future energy needs to achieve demand and supply equilibrium, have control over cost, prevent energy wastage. The purpose of this research is to forecast the electricity consumption of a particular household in Belgium based on the temperature and humidity levels of various rooms in the facility and surrounding weather information. The data set is at 10 min for about 4.5 months. Our aim here is to predict the energy consumed by the appliances using the machine learning regression method.

approach:

First I began with understanding of the data. Dropped the column 'light' because it contains maximum values as 0, also the column 'date' which will not be useful for regression. Found out and removed the outliers present in column 'appliances'. To visualize the distribution of values in each column plotted histogram, to find correlation between the variables present plotted a heatmap. To convert the data into a standard format suitable for regression model I used standardscaler after which with the help of train_test_split from sklearn module splitted the dataset into 75% for training of model and 25% for the testing of model. Various models were used for training like ridge regression, support vector regression, k-nearest neighbours regression, Random forest regression, Gradient boosting regressor, Extra tree regressor. To find the best suitable model evaluation metrics like r-squared(r^2) and root mean square error(rmse). All the models were trained on the dataset and their test results were compared, hyperparameter tuning was performed on the best model using grid search cv to find the best result possible.

Conclusion:

Extra tree regressor proved to be the best model among all other model for the prediction of appliance energy consumed with r^2 score of 0.65 and rmse score of 0.59. The model can be improved by using independent variables which have good correlation with the dependent/target variable.