

# Capstone Project Submission

## Instructions:

- i) Please fill in all the required information.
- ii) Avoid grammatical errors.

### **Team Member's Name, Email and Contribution:**

1. G. V. Kapeesh Varma ([kapeeshvarma@gmail.com](mailto:kapeeshvarma@gmail.com))
  - Data Wrangling
  - Exploratory Data Analysis & Visualizations
  - Data Preprocessing
  - Regression Modelling
  - Conclusions

### **Please paste the GitHub Repo link.**

Github Link: <https://github.com/kapeeshvarma/Appliances-Energy-Prediction>

### **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)**

In this time of global uncertainty world needs energy and in increasing quantities to support economic and social progress and build a better quality of life, in particular in developing countries. But even today there are many regions which face constant power outages. These outages are primarily caused due to excess load consumed by home appliances. Hence, the ability to predict energy consumption can not only save money for end user but can also help in generating money by giving excess energy back to Grid (in case of solar panels usage).

In this project, I performed Regression analysis in order to predict Appliance energy usage based on data collected from various sensors. The dataset used is 'data\_application\_energy.csv'. It consists of home sensors data taken at 10-minute intervals for 4.5 months. Weather data was downloaded from a public data set from Reliable Prognosis (rp5.ru) and merged with this experimental data sets using the date and time column. Two random variables have been included in the data set for testing the regression models and to filter out non-predictive attributes (parameters).

Through different data exploration, visualizations, and various predictions, I found out many valuable insights such as Highest correlation factor, Energy consumption during Weekdays vs Weekends etc. Regression has been performed on this data using different models such as Linear Regressor, Ridge Regressor, Random Forest Regressor etc., and the performance metrics of these models have been calculated to find the best fit.