Project Title: Prediction of factors affecting energy consumption by home appliances

Objective:

Developing and implementing a supervised learning model using linear and logistic regression on the dataset to predict the energy usage of appliances based on temperature, humidity and weather attributes.

The energy consumption in households can be primarily explained by two major factors - type and a number of appliances and usage of appliances by the occupants. Both these factors are interrelated. The use of these appliances by occupants leaves traceable signals near the vicinity of the appliance being used in form of changes in humidity, temperature, vibrations, noise and light. This dataset is about the inputs gathered from a particular house. Both indoor factors like humidity and temperature and external factors like weather, wind speed etc. are considered. In addition to that amount of usage of energy by lights was also factored in which could indicate occupancy in a particular room. The humidity and temperature inside the house were captured using sensors and outside weather-related data points was sourced from the closest weather station.

Data Source: The Dataset can be downloaded from UCI_ML_Repository.

Link: https://archive.ics.uci.edu/ml/datasets/Appliances+energy+prediction#

Description of Data:

- Dataset consists of 19735 observations and 29 variables. Description can be found by visiting the link.
- Dependent variable (y) is Appliances, energy use in Wh. The main features are temperature, humidity, and pressure. The features were monitored with a sensor and averaged for 10 minutes periods.

Data Summary :

Feature Range

Temperature   : -6 to 30 deg
Humidity   : 1 to 100 %
Windspeed   : 0 to 14 m/s
Visibility : 1 to 66 km
Pressure   : 729 to 772 mm Hg
Appliance Energy Usage : 10 to 1080 Wh

Teammates and Work Division:

Group: 3.6

Ajay Samala: Training Model, Data Analysis, Data Manipulation, Report
Meenakshi Shastri: Linear and Logistic Regression model, Data Visualization, Presentation
Siddhartha Shankar: Testing Model, Time Series Analysis, Presentation and Report review

Methodologies:

Linear regression; Logistic Regression; Time Series Analysis; Data Visualization ; Data Manipulation

Python Packages -NumPy, Pandas, Matplotlib; SciKit;