Modeling of household energy usage through analysis of spatial environmental conditions
(12741- Project Proposal)

Yash Gokhale

M.S. Chemical Engineering, Dec '20

Motivation

- Estimating energy requirement of an area is important to allocate resources properly
- Energy use depends on a number of factors such as number of members in house, income, number of appliances, etc
- Analysis has been conducted on these factors
- Imperative to understand if environmental conditions influence energy usage
- Required to analyze the seasonal variance in energy usage

Database under consideration
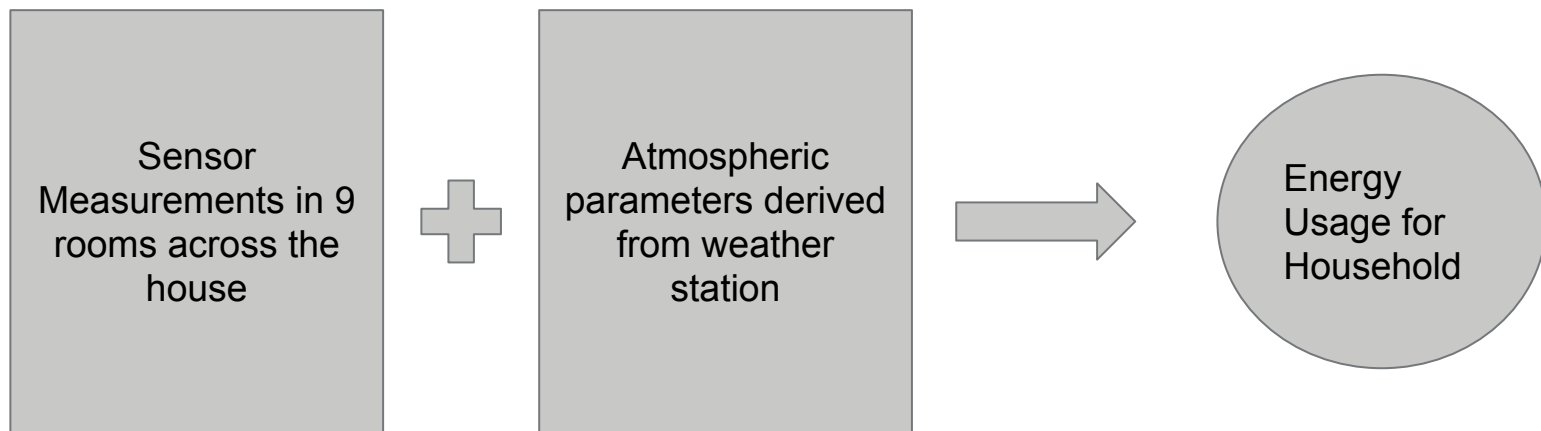
About the Dataset: (Recordings from Belgium, shared publicly)
Number of features: Date+(Temperature & RH)*9 rooms + 6 weather measurements = **25 features**
Output variable: Energy Usage
Number of data points: 19735

Data Source: https://www.kaggle.com/loveall/appliances-energy-prediction

Solution & approach to solve the problem

- Develop a predictive model of the 25 features to estimate the energy requirement in an household
Relevance to course: <u>Time series Linear Regression, Outlier Detection, Time independent Linear Regression</u>
- Regroup the energy output in 3 categories based on usage and remodel the dataset with categorically embedded features
Relevance to course: <u>Run SQLite Queries to filter the data as per requirement (Develop a Python function which can output .sql databases through user inputs)</u>

Questions to be tackled

- Is the dataset sufficient to develop a robust model?
- Is the energy usage really seasonal (time-dependent) or no relation can be established?
- Which features (out of the 25) are important?
- Is correlating atmospheric parameters a good idea? (Modeling with and without those parameters)
- Can the database be further simplified through custom made .sql queries?

Carnegie
Mellon
University