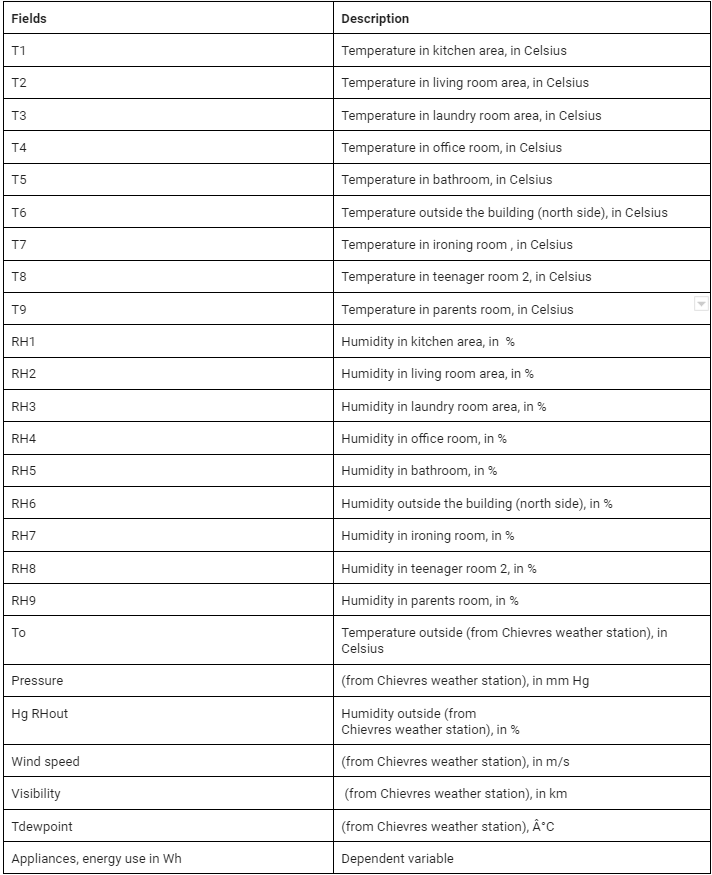
**Appliance Energy Prediction**

**Project Description**

**Business Context**

The data set is at 10 min for about 4.5 months. The house temperature and humidity conditions were monitored with a ZigBee wireless sensor network. Each wireless node transmitted the temperature and humidity conditions around 3.3 min. Then, the wireless data was averaged for 10 minutes. The energy data was logged every 10 minutes with m-bus energy metres. Weather from the nearest airport weather station (Chievres Airport, Belgium) was downloaded from a public data set from Reliable Prognosis (rp5.ru) and merged together with the 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. You need to predict the energy use of appliances.

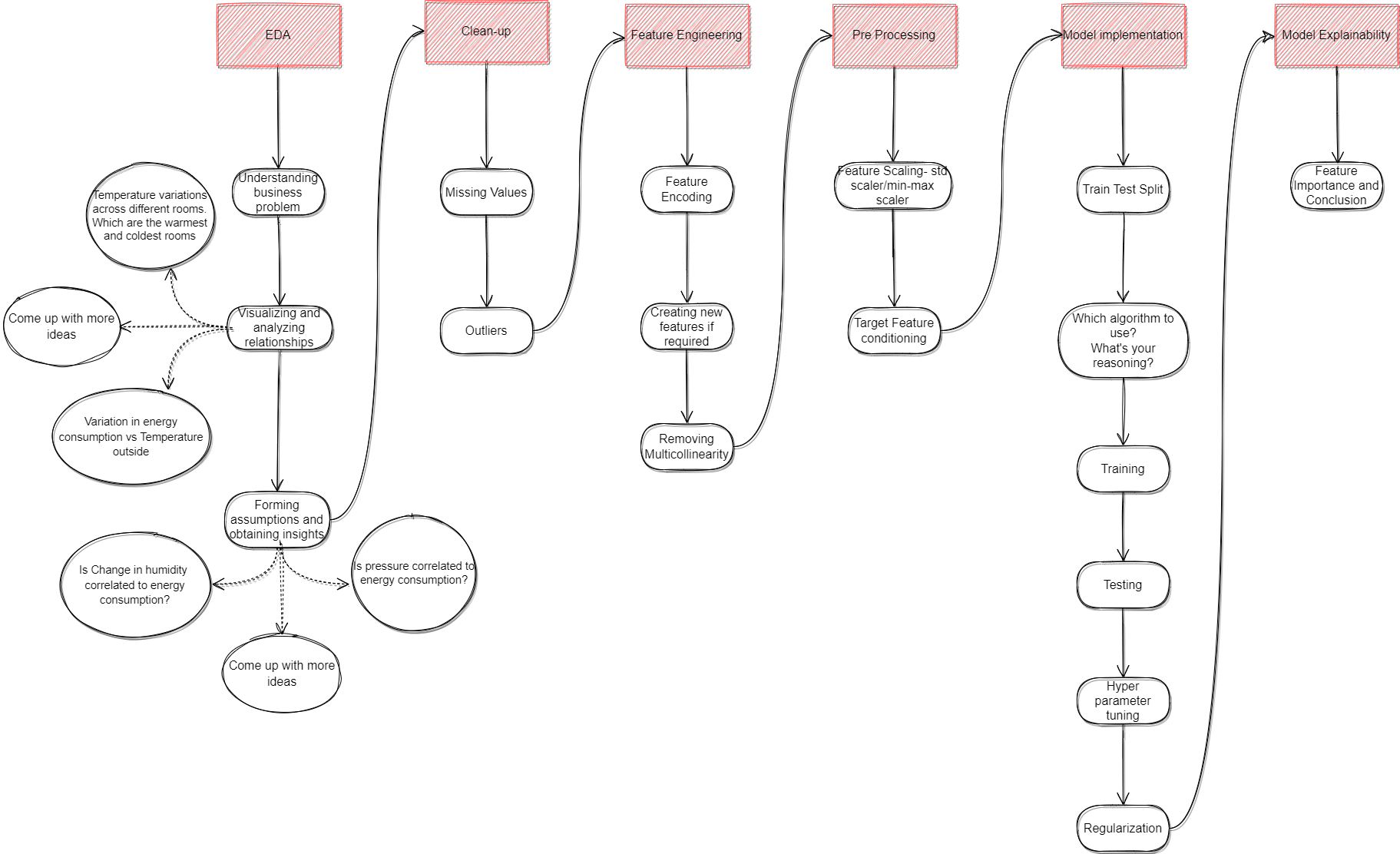
**Data Description:**



**Main Libraries to be Used:**

* Pandas for data manipulation, aggregation
* Matplotlib and Seaborn for visualisation and behaviour with respect to the target variable
* NumPy for computationally efficient operations
* Scikit Learn for model training, model optimization, and metrics calculation

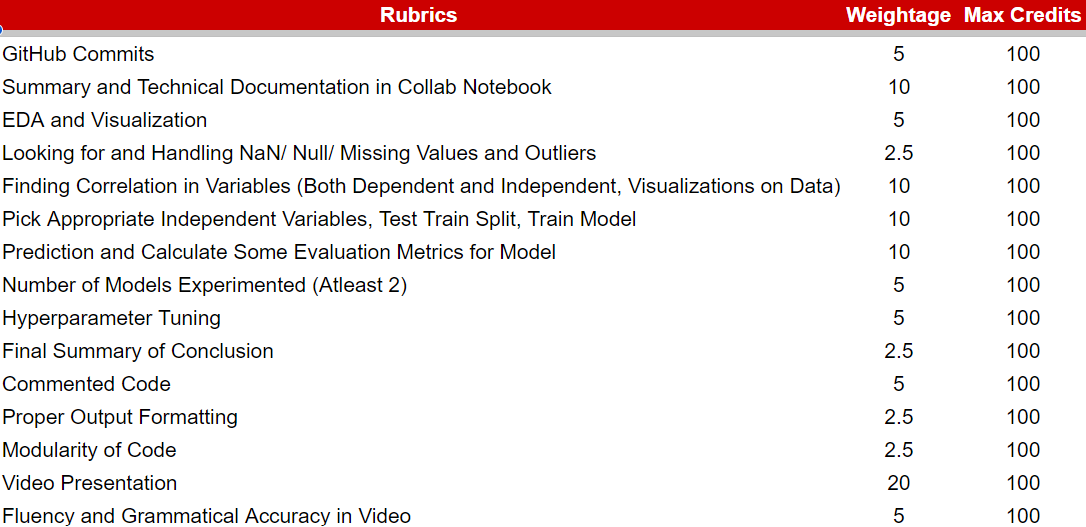
**Project Architecture:**



**Project Evaluation Criteria**

* Efficient EDA
* Encoding if necessary.
* Feature selection, new feature creation
* Dealing with multicollinearity if any
* Feature scaling
* Understanding the target feature and its distribution
* Modeling - use at least two algorithms
* Evaluation and improvement of model.
* Feature Importance and Conclusion
* Understanding of how your project is useful to stakeholders?

**Rubrics**



**Data File**

<https://drive.google.com/file/d/1F8b3o67GpdIvEsHEVNZCD5tmQksW4vuJ/view?usp=sharing>

**Project Template**

<https://colab.research.google.com/drive/1pLHuk5ml8brfq76USc0n0suZU4FLxCqd?usp=sharing>