Domain:

Non-intrusive load monitoring (NILM), or energy disaggregation, aim to break down a household's aggregate electricity consumption into individual appliances.

In order to evaluate such energy disaggregation algorithms, it is necessary to collect data sets containing both the household aggregate power demand (the input of the algorithm)and the power demand of each individual appliance(the ground truth against which the algorithm is compared).

The domain of this problem is signal processing in power system, an area on which I worked on for 6 year by analyzing the power and current signal in high voltage cable for the identification of partial discharges.

Data:

The dataset for this project is one month of data from the UK-DALE dataset (thttp://jack-kelly.com/data/), which stores appliances and aggregated data from 5 different houses over a 26 month time-span.

I will take into account 5 type of appliances (fridge, washing machine, dish washer, kettle and microwave).

Features:
Main features: for each appliance, power and current values are recorded as time-series, with 6 second sampling. For the aggregate the sampling is 1 second.
For each appliance further features will be evaluated and used for the classification, such as:
- Representative period of voltage and current.
- Voltage – Current trajectory.
- Representative cycle of current and voltage.
- Real and reactive power
- Power harmonics
Other features are building meta-data (e.g. : age of construction, number of people living here, …)

Known unknowns:
- Feature extraction and selection