

## DESCRIPTION

Do you want to solve one of the hottest problems the electrical energy industry faces right now?

Distribution companies currently massively replace the old energy meters with the new so-called smart meters. One of the greatest challenges right now is how to get as much as possible from this data.

Your challenge is to make a supervised machine learning model that would recognize electrical appliances (refrigerator, oven, boiler, heat pump etc.) from smart metering energy time series data.

Energy disaggregation, also called Non-Intrusive Load Monitoring (NILM) is an approach where machine learning and statistical models are used to analyse the total energy consumption and determine the energy consumption signals originating from electrical appliances. A lot has been made on datasets with high-resolution data. Since smart meters measure consumed energy at 15 min intervals you should learn a model that can acquire original signals of electrical appliances from actual smart metering data.

## OUTCOME:

**Task 1:** Start with a freely available dataset called REDD [1]. Original dataset is recorded at 3s resolution. Resample data to 15 min and explore which devices can be recognized by learning a model with this low-resolution data.

**Task 2:** After learning a model on REDD dataset try to recognize electrical appliances from actual households smart metering data. When examining actual data you will not have labels about appliances.

Plot recognized behavioural patterns. Results have to be visually appealing and understandable to non-experts.

Python - Jupyter is the preferred environment.

External data about households appliances can be used (only for Task 2).

We are waiting for you to help us with your solution.

Links:

[1] REDD dataset description: <http://redd.csail.mit.edu/kolter-kddsust11.pdf>

[2] Data description (only low resolution data): <http://redd.csail.mit.edu/readme.txt>

Datasets descriptions:

Folder "*REDD\_low\_freq*" consists of 6 sub-folders. Each sub-folder includes datasets for one household. Please read the exact description in [2] under title "*Low-Frequency Power Data*".

Smart metering data for 100 households is in file "*hack\_elect\_data\_1.csv*". Consumers are anonymized and numbered from 0 to 99. Every row represents daily load time series in kW for particular consumer. Each consumer has load time series data for year 2016.