

Challenge context

BCM Energy is a start-up based in Lyon and created in 2015. BCM operates on the whole value chain of renewable electricity, from production assets management on electricity markets (Epexspot, EEX) to electricity supply to final consumers through the brand Planète OUI:

1. BCM Energy is active continuously on the electricity markets,
2. BCM Energy is responsible of its own balancing perimeter,
3. BCM Energy also manages capacity certificates within its own certification perimeter and guarantees of origin in its register.

The diverse priority areas of growth are supported by a trading team with more than fifteen years of experience of the various electricity markets, developing state-of-the-art financial analysis modeling. Planète OUI, created in 2007, is one of the first French green electricity supplier. The company supplies tens of thousands of homes and professionals and covers more than 95 % of metropolitan France (Enedis network). Planète OUI promotes an ecology constructive and made available for all. It has integrated BCM Energy's perimeter in 2017. The supplier has to offer green electricity supply with prices adapted to the consumption profiles of its clients. Indeed, the information of disaggregation consumption could help reduce electricity consumption and so reduce the electricity bill with customized advice or control of appliances of our clients (with their agreement).

Challenge goals

The goal is to train an algorithm to replace many monitoring systems which are too intrusive and too expensive. This challenge is known as NILM (Nonintrusive load monitoring) or NIALM (Nonintrusive appliance load monitoring). The aim of the challenge is to find the part of electric consumption in one household dedicated to 4 appliances (washing machine, fridge_freezer, TV, kettle). There are no time constraints. The past and the future are known.

Data description

The first line of the input contains the header, the columns are separated by ',' and decimals by decimal point. The columns are:

1. the "time_step": date measured each minute (format yyyy-MM-ddTHH :mm :ss.Z)
2. the "consumption": household consumption (W) measured each minute
3. the "visibility": distance at which it is possible to clearly distinguish an object (km) measured once per hour
4. the "temperature": temperature (°C) measured once per hour
5. the "humidity": presence of water in the air (%) measured once per hour
6. the "humidex": index used to integrate the combined effects of heat and humidity measured once per hour
7. the "windchill": an index that expresses the subjective feeling of cold or heat as a function of measured temperature, wind and humidity (°C) measured once per hour
8. the "wind": wind speed (km/h) measured once per hour
9. the "pressure": applied perpendicular to the surface of an object per unit area (Pa) measured once per hour

The first line of the output contains the header, the columns are separated by ',' and decimals by decimal point. The columns are:

1. the "time_step": date measured each minute (format yyyy-MM-ddTHH :mm :ss.Z)
2. the "washing_machine": washing machine power (W) measured each minute
3. the "fridge_freezer": fridge freezer power (W) measured each minute
4. the "TV": TV power (W) measured each minute
5. the "kettle": kettle power (W) measured each minute

The train set contains 417 599 values with 10 231 missing values (2,44%) for "consumption", "washing_machine", "fridge_freezer", "TV", "kettle" and the test set contains 226 081 values with 24 719 missing values (10,93%).