HOME APPLIANCE IDENTIFICATION FOR NILM SYSTEMS BASED

Theory

Load Disaggregation

Let A be the set of all known appliances in a house and $P(t_i)$ be the aggregate power of the entire house measured at time t_i. Without loss of generality, in the following, we denote $P(t_i)$ as $P(t_i) = P_i \geq 0$. Let ${P_i}^a \geq 0$ be the power load of appliance $a \in A$ at time instance t_i. Then,

$$P_i = \sum_{a=1}^{|A|} P_i + n_i$$

where n_i is the measurement noise that not submetered. The disaggregation task is for $i = 1, ..., T$ and $a \in A$, given the P_i to estimate the ${P_i}^a$.

Operational State Change

Naturally, $\Delta Pi = P_{i+1} - P_i$, $i = 1, ..., T$ and $\Delta {P_i}^a = {P_{i+1}}^a - {P_i}^a$, $i = 1, ..., T$ respectively correspond to the variation of the aggregate power and appliance a power measured at time t_i. The classification labels of each appliance a at time t_i is denoted as ${s^a}_i$. The ${s^a}_i$ is defined as following:

$${s^a}_i = 1, for\ |\Delta Pi| \geq Thr_a$$
$${s^a}_i = 0, for\ |\Delta Pi| \leq Thr_a$$

$where\ Thr_a \geq 0$ is a power threshold for appliance a. Thr_a is set to half of difference between mean values of appliance a's adjacent states, which is observed through training process. When Thr_a is set to 1, it means appliance a changed its operational state (e.g., switched on/off) at time t_i; When ${s^a}_i$ is set to 0, it means appliance a didn't change its operational state at time t_i .

Refer to this paper: http://aircconline.com/ijaia/V9N2/9218ijaia06.pdf

Data

In Data folder you have per second consumption data for 4 houses with following columns:

- Dataid: Unique id for the house
- Localminute: Timestamp for each data point
- Use: Total power consumption of house

- Air1: Consumption of Air conditioner only
- Microwave1: Consumption of Microwave only
- Refrigerator1: Consumption of Refrigerator only

Assignment

Your goal is to predict the operations state change of each appliance just from total energy energy data.

Some key ideas you can try:

1. From data given create appropriate training and testing data
2. Divide data into dependent and independent variables to do a predictive analysis using various ML/DL models
3. Present the information/insights derived for the analysis using various plots and tables.

Judging criterias:

Assignment will not be solely judged based on the accuracy scores of models used but also depth of the analysis, thinking process, decision making and quality of your code.

Submission Instruction

Please write your code preferably on jupyter Notebook like environment using python3. You can use you personal machine to run the code (See the installation instructions here) or you can use google colab (they provide free CPU and GPU with 12Gb ram) without any hassle of installing any packages. Instruction to connect google colab to google drive to access the data is give in sample solution notebook.

If you are not familiar with jupyter notebook you can also submit your solution in python script files, but make sure to add a separate report with your findings and explaining your methods.

- Keep you code clean and well documented
- Do not share this assignment with anyone else.
- **Do not copy code** from other applicants otherwise you will be **immediately removed** from this process and **blocked from company** for any future association.
- If you have any doubts in the assignment you can mail us.