

Progetto Machine Learning 2019/2020

1 Problema

Il problema della disaggregazione dei consumi elettrici è noto in letteratura con l'acronimo NILM (Non-Intrusive Load Monitoring). Con il termine NILM si fa riferimento ad un processo volto a stimare, sulla base delle variazioni del consumo generale, il consumo energetico di un dato elettrodomestico. I sistemi NILM permettono quindi di monitorare il consumo elettrico all'interno di una abitazione senza richiedere l'utilizzo di smart meter dedicati ai singoli elettrodomestici. In letteratura, sono stati proposti diversi approcci atti alla risoluzione del problema della disaggregazione del consumo elettrico [1]. Negli ultimi anni, grazie al forte interesse per le tecniche di Machine Learning, il problema della disaggregazione dei consumi elettrici è stato affrontato ricorrendo ad algoritmi di Deep Learning.

2 Dataset

Il dataset contiene misurazioni al secondo per il consumo complessivo e per due elettrodomestici: frigorifero e lavastoviglie. Il training set (*main_train.csv*, *fridge_train.csv*, *dishwasher_train.csv*) [2] comprende le prime osservazioni che vanno dal 2019-01-01 00:00:00 al 2019-03-14 23:59:59. Il test set invece comprende le ultime osservazioni che vanno dal 2019-03-15 00:00:00 al 2019-03-31 23:59:59.

3 Obiettivo

L'obiettivo del progetto è risolvere il problema di disaggregazione dei consumi elettrici residenziali utilizzando tecniche di Machine Learning. L'algoritmo

dovrà ricevere in input il consumo complessivo dell’abitazione e restituire in output il consumo di frigorifero e lavastoviglie.

4 Valutazione

Il progetto può essere svolto individualmente o in gruppi composti da al massimo 2 componenti. Deve essere prodotta una breve relazione in cui si giustificano le scelte fatte. Per ogni elettrodomestico, la metrica di valutazione è l’energy based **F1** score sul test set:

$$P_i = \frac{\sum_{t=1}^T \min(y_i(t), \hat{y}_i(t))}{\sum_{t=1}^T \hat{y}_i(t)},$$
$$R_i = \frac{\sum_{t=1}^T \min(y_i(t), \hat{y}_i(t))}{\sum_{t=1}^T y_i(t)},$$
$$F1_i = 2 \frac{P_i \cdot R_i}{P_i + R_i},$$

dove T è l’orizzonte temporale del test set, $y_i(t)$ e $\hat{y}_i(t)$ sono rispettivamente il consumo vero ed il consumo predetto per l’elettrodomestico i .

5 Modalità di consegna

Al fine di poter valutare nel migliore dei modi i progetti è importante che gli script siano chiari e parzialmente commentati. Oltre alla relazione, bisogna consegnare il codice sorgente con gli script per l’addestramento dell’algoritmo. Il codice deve inoltre contenere una routine che ha la funzione di pre-processare, leggere da file il modello addestrato e valutare in modo automatico le prestazioni sul test set.

veronica.piccialli@uniroma2.it
antonio.maria.sudoso@uniroma2.it

6 Riferimenti

- [1] <https://arxiv.org/pdf/1703.00785.pdf>
- [2] <https://we.tl/t-xdH24Mczg3>