

# Valutazione dei campioni sconosciuti tramite Support Vector Machine (SVM):

Per poter selezionare un significativo numero di campioni per il test finale, è stato necessario condurre una fase aggiuntiva di addestramento del modello. Essendo il numero di campioni appartenenti alle classi degli elettrodomestici leggermente inferiore a 200 mila, abbiamo deciso di ampliare il numero di campioni etichettati come "altro". Ciò ci ha permesso di selezionare circa 240 mila campioni, presi randomicamente, da mettere da parte. Questo processo è avvenuto dopo la feature extraction e prima dello scaling in previsione del test finale.

Inoltre nella fase di Processamento dei dati è stato eliminato il meccanismo di pulizia tramite il riconoscimento dei cicli, questo è servito a rendere quanto più autentici i dati per la fase di test.

Dopo aver completato l'addestramento e ottenuto il modello, l'abbiamo applicato ai dati precedentemente riservati. La granularità dei dati analizzati era di 1 minuto, quindi per valutare le reali capacità dell'algoritmo, abbiamo esteso le classi predette a tutti i secondi presenti in quel minuto di analisi.

Successivamente, abbiamo eseguito un merge utilizzando il parametro DateTime. Questa operazione ci ha permesso di associare le previsioni (relative ai dati selezionati in modo casuale) agli istanti di campionamento del file sorgente. Abbiamo così aggiunto una colonna al file originale contenente i valori predetti, facilitando il calcolo del recall e dell'F1-score tramite Sklearn.

Il file che si occupa del test si trova nella repository principale ed è chiamato: **"test\_code.py"**.

Classe	Recall	F1-score
altro	1.0	0.98
washing_machine	0.9	0.94
oven	0.99	0.99
dishwasher	0.93	0.96

Nell'immagine qui presente sono riportati i risultati ottenuti dal Test con granularità di un minuto.

Test effettuato su 4000 campioni.

Classe	Recall	F1-score
altro	0.91	0.95
washing_machine	0.99	0.73
oven	1.0	0.99
dishwasher	0.98	0.96

In questa immagine sono riportati i risultati ottenuti dal Test con granularità di un secondo.

Test effettuato su 240000 campioni.

Conclusione:

Media Recall: 99%   Media F1-score: 89%

washing_machine	0.99	0.73
oven	1.0	0.99
dishwasher	0.98	0.96

Il valore della F1-score appartenente alla classe washing\_machine risulta più basso poiché i campionamenti intra minuto con etichetta altro aumentano i falsi positivi della classe lavatrice abbassandone la precisione.

