* **Dataset (descrizione dataset)**

Dataset per training

Labels.xlsx

contiene info su 3 navicelle spaziali riguardanti condizioni di fault, percentuale di apertura delle singole valvole e presenza o meno di una bolla

* SV1-4 riportano la percentuale di apertura delle 4 valvole per ogni navicella
* BP1-7, BV1 riportano la presenza di una bolla nelle tubazioni, in corrispondenza di 8 punti specifici

Case1-177.csv

Ogni file, riporta misurazioni di pressione campionati ad 1kHz su 0 – 1200 ms

* P1-7 sono le pressioni misurate dai singoli sensori

Dataset per testing

labels\_spacecraft.xlsx

indice sulle misurazioni effettuate per le navicelle 1 e 4

Case178-223.csv

Ogni file, riporta misurazioni di pressione campionati ad 1kHz su 0 – 1200 ms

* P1-7 sono le pressioni misurate dai singoli sensori





OBIETTIVI

[NOTE]

Anomaly – è conisderata come la presenza di una bolla nelle tubature durante le operazioni della navicella.

Fault – problema legato alla apertura della valvola per la propulsione della navicella.

Unknown anomaly – guasti o anomalie impreviste (sconosciute). Sono distinti dai Fault e Anomaly noti

1. Determinare per il dateset test ogni elemento con comportamento Normal e Abnormal.
2. Determinare se per ogni elemento Abnormal si tratta di un ‘Anomaly’ (presenza di bolla nelle tubature) ‘Fault’ (mal funzionamento di una valvola) o si tratta di ‘unknow fault’ (situazione sconosciuta).
3. Determinare per ogni ‘Anomaly’ la presenza della nelle 8 posizioni (BV1, BP1 – BP7).
4. Determinare per ogni ‘Fault’ la valvola che non funziona correttamente.
5. Determinare la percentuale di apertura per ogni ‘Fault’ predetto

RAGGIUNGIMENTO

[NOTE]

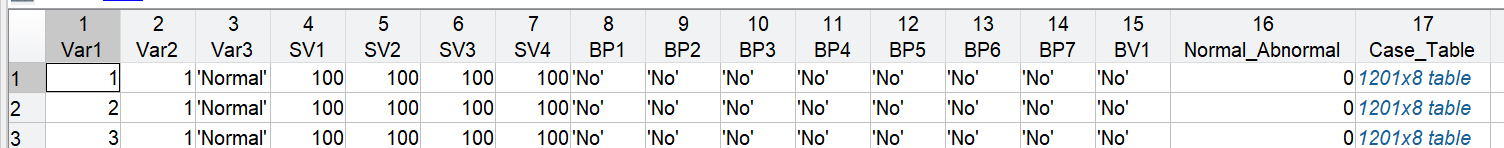
A seguire, per ogni obiettivo vengono riportate nel dettaglio le operazioni eseguite

**[1] Determinare per il dateset test ogni elemento con comportamento Normal e Abnormal**

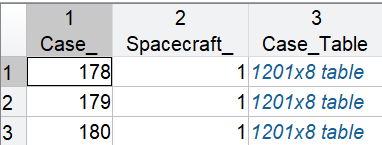
* **ETL (aggiunta time table e Normal\_Abnormal features)**

Consiste nella creazione di una script in matlab per caricare i dataset di training e testing, e seguire al file Label.xlsx sono stati aggiunti i vari file CSV ed una colonna che riporta l’intestazione ‘Normal\_Abnormal’.

Cosi da ottenere due tabelle distinte, una con 177 elementi per il training, dove per ognuno sono riportate le misurazioni di pressione ed è indicato se si tratta di una situazione Normal o Abnormal; ed una seconda tabella con 46 elementi per il testing.



Intestazione ‘table’ di training finale



Intestazione 'table' di test finale

* **Caricamento dataset su Deagnostic Feature Design**

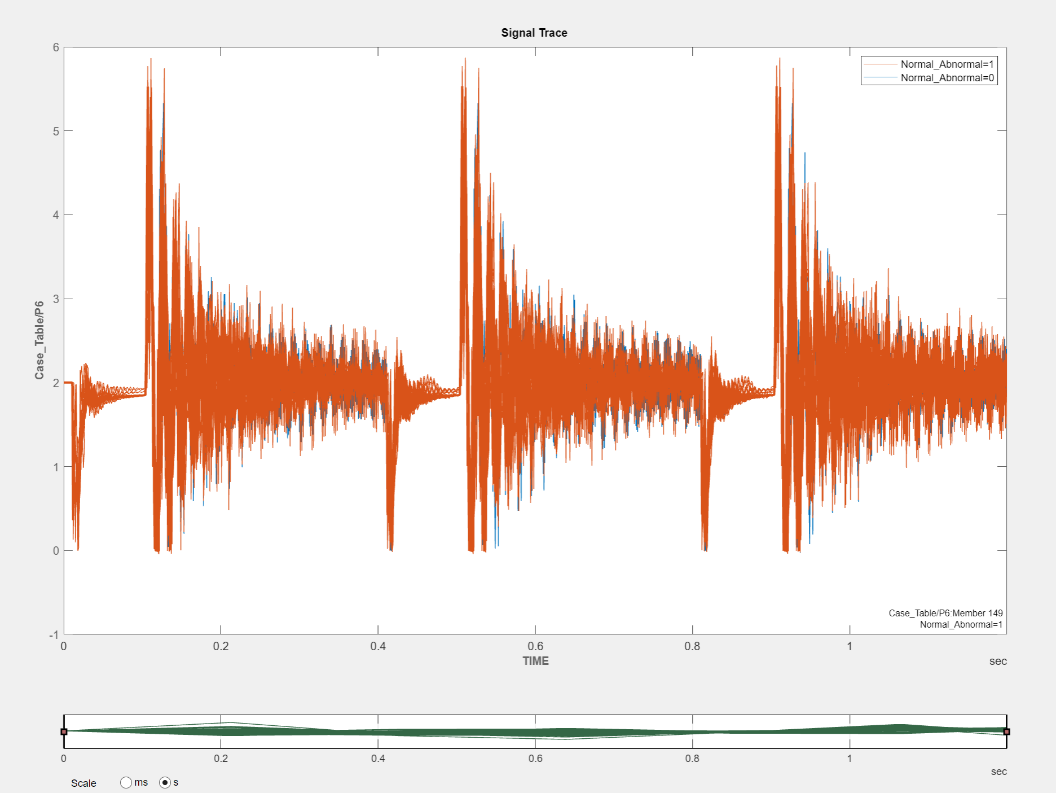
Nella fase successiva sono passato allo studio dell’estrazione e selezione features per costruire un modello classificatore.

Avviato Il DFD ho importato il dataset di train selezionando le variabili ‘Var1’ ‘Case\_Table’ e ‘Normale\_Abnormal’ impostandola come **Condition Variable**

* **Signal trace features P1-7**

Per i **Signal Trace**, ho eseguito l’operazione per i 7 segnali di pressione P1-7. Nel graficare i segnali ho impostato un raggruppamento per variabile ‘Narmal\_Abnormal’ e con 177 numero di curve.

[osservazione] tramite il signal trace per P4, P6 e P7 è facile individuare subito situazioni anormali (color blu)



Fesempio 'signal trace' ottenuto per P7

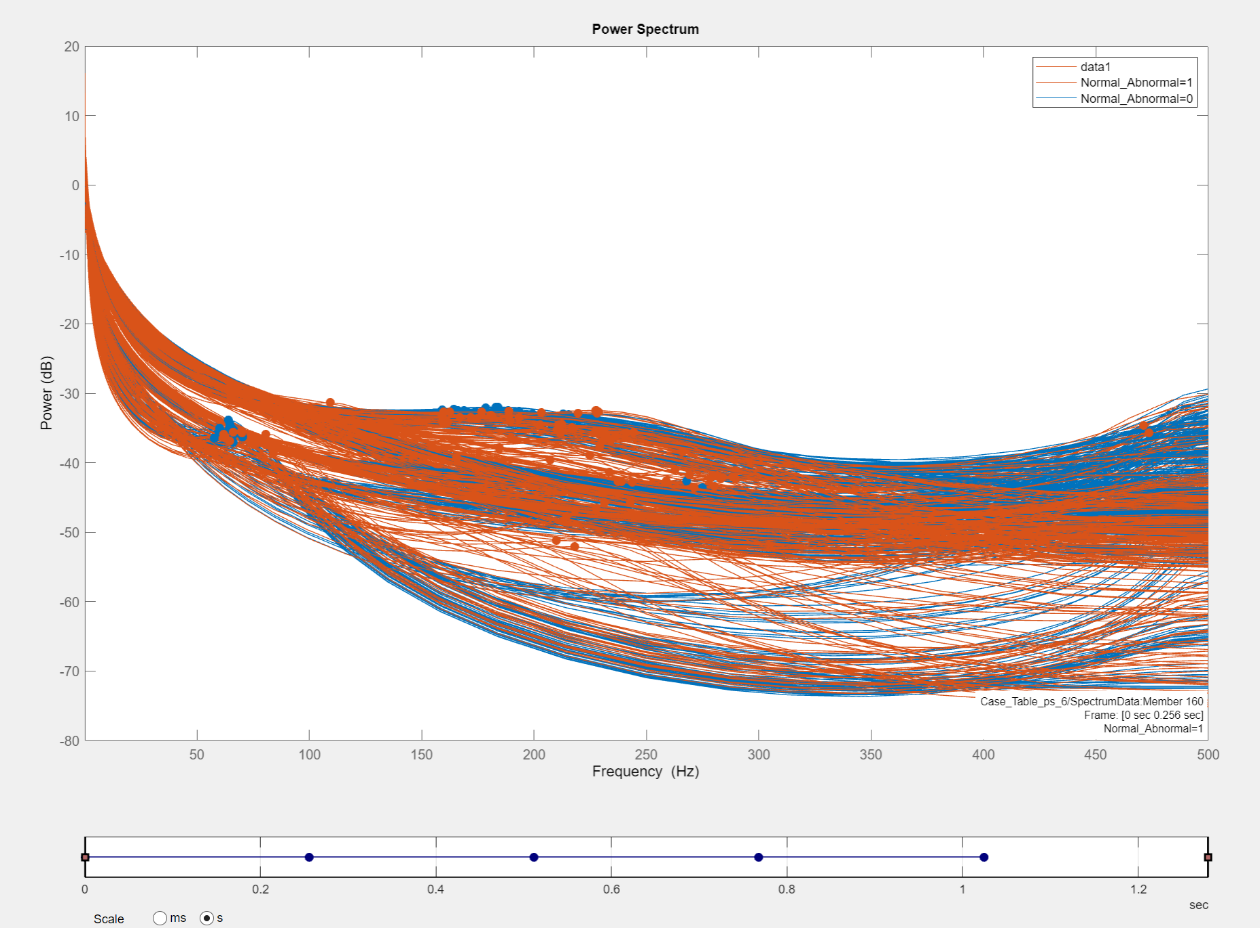
* **Estrazione features**

Per l’estrazione delle features nel dominio del tempo, ho usato il modulo **‘Signal Features’** fornito da DFD, selezionando tra i possibili settaggi solo le features incluse in **‘impulsive features’**.

L’operazione di estrazione delle features è stato eseguito per ogni segnale (P1-7).

Per l’estrazione delle features ho lavorato con una **‘Frame Policy’** che impone size e rate pari a 256 ms.

Nel **dominio spettrale**, è stato necessario eseguire un auto-regressione di ordine 4 per ogni segnale utilizzando un campionamento a 1kHz e una ‘Frame Policy’ con size e rate si 256 ms, e per ogni ‘Power Spectrum’ ottenuta ho raggruppato per Normal\_Abnormal e linearizzato l’asse delle X.



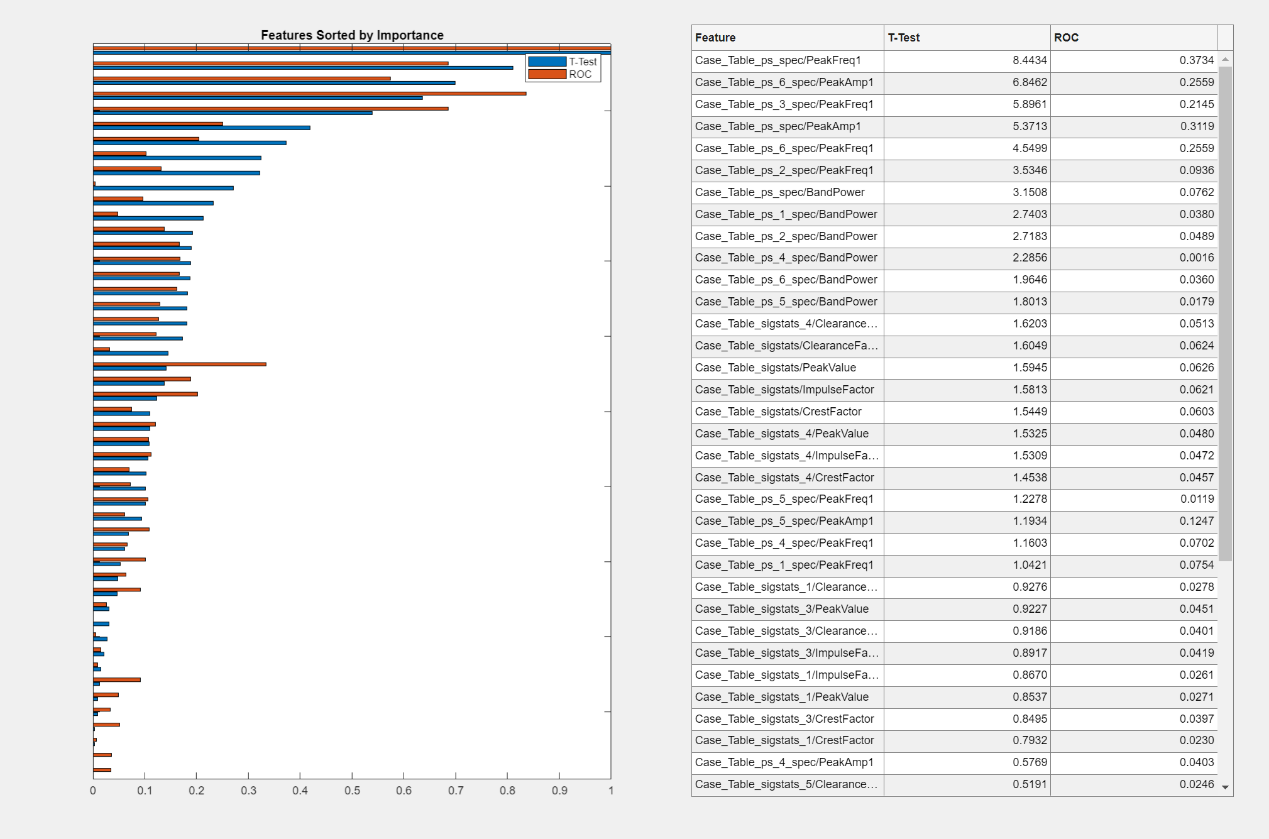
‘Power Spectrum’ attenuto su P7

Solo dopo aver creato i **SpectrumData** sono passato alla vera estrazione delle features nel dominio spettrale, usando il tool **Spectral Features** fornito da DFD settando Peak ad 1.

* **Ranking features**

Conclusa la fase di estrazione delle features sono passato alla selezione delle stesse, utilizzando il ‘Features Ranking’.

Nello scegliere la features ho utilizzato il T-test e ROC per classificarle cosi da potere avere un indice che mi aiuta nel selezionare le features ‘*migliori’.*

**

ranking features

* **Classifacre – classificator learner**

Con l’uso di FRAME POLICY con frame size e frame rate di 0.256, ottengo classificatori con accuratezza molto bassa

* **Matrice confusione**