

UNIVERSITÀ POLITECNICA DELLE MARCHE
FACOLTÀ DI INGEGNERIA

Dipartimento di Ingegneria dell'Informazione
Corso di Laurea Magistrale in Ingegneria Informatica e dell'Automazione



**PROGETTO DI MANUTENZIONE PREVENTIVA PER LA ROBOTICA E
L'AUTOMAZIONE INTELLIGENTE**

**🌀 C2 PHM Asia Pacific 2023 - Sviluppo di un modulo per la
diagnosi di guasti all'interno di un sistema di propulsione**

Docente

Freddi Alessandro

A cura di

De Grazia Davide,
Piergallini Enrico,
Giuliani Mauro

ANNO ACCADEMICO 2024-2025

1	Introduzione	1
1.1	Scenario	1
1.2	Tasks da svolgere	2
2	Data exploration	3
3	Feature engineering	7
4	Tasks implementati	10
4.1	Task 1	10
4.2	Task 2	11
4.3	Task 3	13
4.4	Task 4	14
4.5	Task 5	15
5	Risultati e confronto sul Test Set	18
5.1	Task 1	18
5.2	Task 2	19
5.3	Task 3	19
5.4	Task 4	20
5.5	Task 5	21
5.6	Calcolo dello score ottenuto	21
5.6.1	Metriche di valutazione della competizione	21
5.6.2	Punteggio finale	22

Elenco delle figure

2.1	Serie Temporale	3
2.2	Serie Temporale per singoli casi	4
2.3	Prima fase di apertura per il sensore P2	4
2.4	Fase di chiusura	4
2.5	Apertura e chiusura del sensore	5
2.6	Deviazione standard sensore P2	5
2.7	Trasformata di Fourier sensore P2	6
4.1	Schema operativo Task 1	10
4.2	Top 20 Feature selezionate tramite ANOVA F-value	11
4.3	Schema operativo del Task 2	12
4.4	Schema operativo del Task 3	13
4.5	Top 10 feature selezionate secondo ANOVA F-value (Task 3)	13
4.6	Schema operativo Task 4	14
4.7	Feature principali selezionate con metodo ANOVA per il Task 4	15
4.8	Schema operativo Task 5	16
4.9	Modelli di regressione a confronto sulla valvola SV3 in fase di validazione . .	16
5.1	Matrice di confusione Task 1 relativa al testing set	18
5.2	Matrice di confusione Task 2 sul testing set	19
5.3	Matrice di confusione Task 3 relativa ai dati di testing	20
5.4	Matrice di confusione Task 4, testing set	20

Il progetto PHM Asia Pacific 2023 si concentra sullo sviluppo di un modulo diagnostico per il sistema di propulsione spaziale. Il sistema è stato simulato dalla Japan Aerospace Exploration Agency (**JAXA**) e include scenari di guasti alle elettrovalvole, contaminazione da bolle d'aria e anomalie sconosciute. Il compito consiste nell'identificare e classificare tali anomalie, utilizzando tecniche avanzate di machine learning e selezione delle feature diagnostiche, al fine di migliorare la precisione nella diagnosi e nella gestione dei guasti nei veicoli spaziali.

1.1 Scenario

Si considera un sistema di propulsione spaziale in cui il fluido impiegato è acqua pressurizzata a 2 MPa, espulsa tramite quattro elettrovalvole (SV1 – SV4), che simulano i propulsori. I sensori di pressione (P1 – P8) rilevano i dati temporali con una frequenza di campionamento di 1 kHz su un intervallo di 0-1200 ms.

L'apertura e la chiusura delle elettrovalvole segue uno schema periodico ben preciso all'interno della finestra temporale di 1200 ms:

- Prima apertura tra 0,0 e 0,1 s;
- Prima chiusura tra 0,1 e 0,4 s;
- Seconda apertura tra 0,4 e 0,5 s;
- Seconda chiusura tra 0,5 e 0,8 s;
- Terza apertura tra 0,8 e 0,9 s;
- Terza chiusura tra 0,9 e 1,2 s.

Questo alternarsi di cambiamenti genera fluttuazioni di pressione, causate dal fenomeno del *colpo d'ariete*, all'interno del sistema di propulsione.

Le anomalie incluse nel dataset comprendono:

- **Anomalia da bolle:** La presenza di bolle d'aria nei tubi modifica la velocità del suono, alterando lievemente le fluttuazioni di pressione. È necessario rilevare la comparsa di bolle e determinare la loro posizione. Le posizioni possibili per le bolle sono: BV1 e BP1–BP7. La quantità di bolle è costante per tutti i casi.

- **Guasti alle elettrovalvole:** I guasti alle elettrovalvole comportano una riduzione del volume di fluido che passa attraverso l'elettrovalvola, in quanto queste si aprono a un grado compreso tra 0% e 99%. Nel corso del lavoro sarà necessario determinare quale elettrovalvola ha subito il guasto ed il relativo grado di apertura di quest'ultima.
- **Anomalia sconosciuta:** Possono verificarsi guasti imprevisti che devono essere distinti da quelli noti (bolle o guasti alle elettrovalvole). Alcuni guasti sconosciuti sono inclusi nei dati di test e devono essere identificati senza confonderli con altre anomalie.
- **Differenze tra veicoli spaziali:** Poiché le elettrovalvole presentano differenze individuali, come i tempi di apertura e chiusura, i dati acquisiti mostrano variazioni tra i veicoli. Nel contesto della competizione, sono considerati quattro veicoli spaziali (Spacecraft No.1 - Spacecraft No.4). I dati di addestramento includono i risultati per i veicoli No.1 – No.3, mentre i dati di test provengono dai veicoli No.1 e No.4.

Il modulo diagnostico dovrà classificare le anomalie, identificare le posizioni delle bolle e distinguere tra guasti noti e sconosciuti.

1.2 Tasks da svolgere

Il progetto prevede una serie di task sequenziali che devono essere completati per sviluppare il modulo diagnostico. I task principali sono i seguenti:

- **T1 - Determinare se i dati di test sono normali o anomali:** Per tutti i dati di test, è necessario identificare se presentano anomalie (di qualsiasi tipo) o no.
- **T2 - Classificazione dell'anomalia:** Per i dati identificati come anomali, determinare se si tratta di un'anomalia da contaminazione da bolle, guasto alle elettrovalvole o guasto sconosciuto.
- **T3 - Identificazione della posizione delle bolle:** Per i dati identificati come contaminazione da bolle, determinare la posizione della bolla tra otto possibili posizioni: BV1 e da BP1 a BP7.
- **T4 - Identificazione del guasto alle elettrovalvole:** Per i dati identificati come guasto alle elettrovalvole, determinare quale delle quattro elettrovalvole (SV1–SV4) ha subito il guasto.
- **T5 - Predizione del rapporto di apertura dell'elettrovalvola:** Per l'elettrovalvola identificata come guasta, prevedere il rapporto di apertura (compreso tra 0% e 100%).

Per l'analisi di data exploration, viene selezionato un sensore specifico dal dataset di addestramento, e viene scelto un caso tra i 178 disponibili nei dati di training forniti dalla competizione. La Figura 2.1 mostra il plot della serie temporale che rappresenta l'andamento del sensore selezionato per il caso specifico.

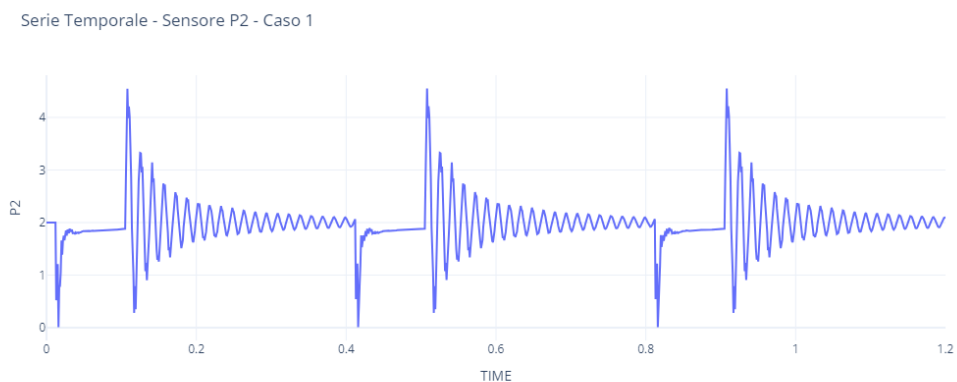


Figura 2.1: Serie Temporale

Successivamente, vengono confrontati i segnali relativi a tre condizioni operative: *Normal*, *Fault* e *Anomaly*. Il grafico in Figura 2.2 visualizza le serie temporali del sensore scelto, utilizzando colori distinti per ciascuna condizione.

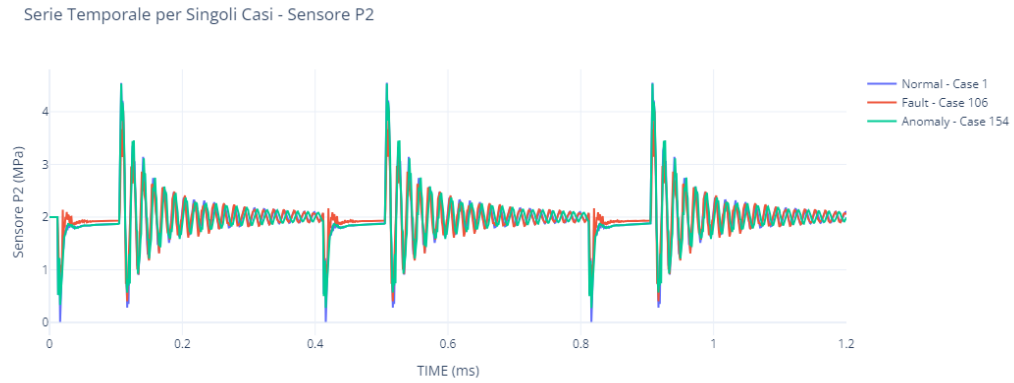


Figura 2.2: Serie Temporale per singoli casi

Infine, il comportamento del sensore viene analizzato in dettaglio nelle fasi di apertura e chiusura, considerando una finestra temporale ristretta. La Figura 2.3 mostra l'andamento della pressione nel tempo per ciascuna condizione nella prima fase di apertura (0.0-0.1ms).

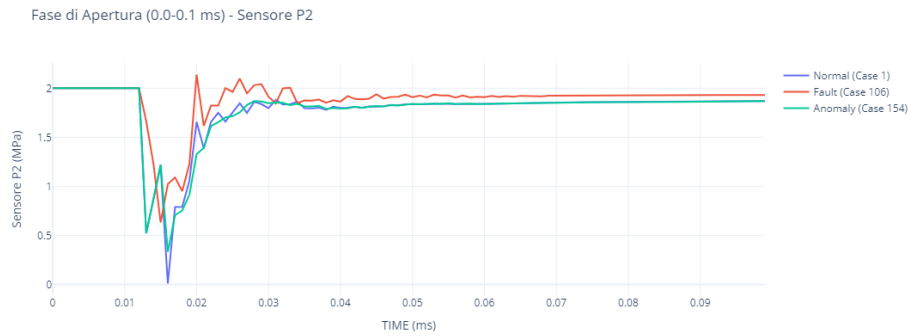


Figura 2.3: Prima fase di apertura per il sensore P2

Viene poi definita una finestra temporale per la fase di chiusura, come 0.1-0.4 millisecondi, con la possibilità di scegliere altre finestre temporali. Per ogni condizione operativa (*Normal*, *Fault*, *Anomaly*), viene selezionato un caso rappresentativo e tracciata la relativa serie temporale. Ogni condizione è rappresentata con un colore distinto: blu per *Normal*, rosso per *Fault*, e verde acqua per *Anomaly*. Il grafico in Figura 2.4 mostra il comportamento del sensore durante la fase di chiusura scelta, nelle diverse condizioni simulate.

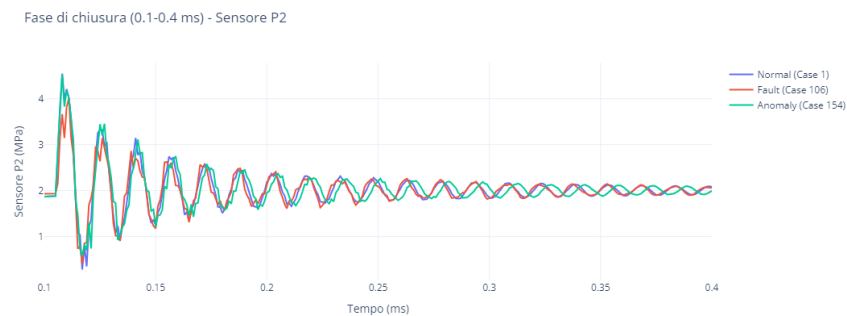


Figura 2.4: Fase di chiusura

Per analizzare la fase di apertura e chiusura del sensore scelto, viene selezionato un sensore specifico (ad esempio, P2) e definita una finestra temporale, come 0.000–0.400 secondi. Vengono poi selezionati i casi rappresentativi per ciascuna condizione operativa (*Normal*, *Fault*, *Anomaly*) e tracciati i relativi segnali. Ogni condizione è colorata in modo diverso: blu per *Normal*, rosso per *Fault*, e verde acqua per *Anomaly*. Il grafico, mostrato nella Figura 2.5, rappresenta l'andamento del sensore durante la fase di apertura e chiusura nella finestra temporale selezionata, evidenziando le differenze tra le condizioni operative.

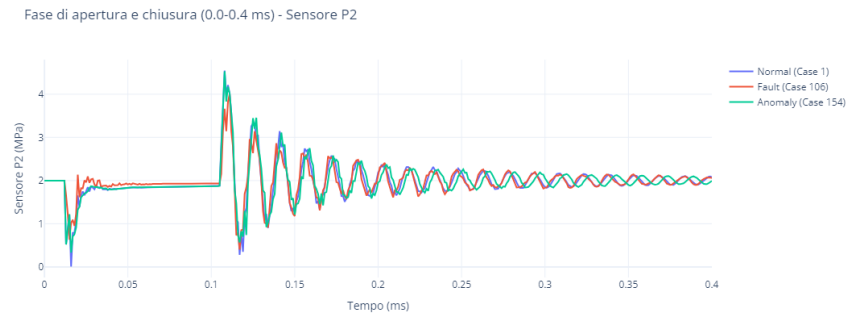


Figura 2.5: Apertura e chiusura del sensore

L'analisi del sensore P2, in Figura 2.6, ha rivelato variazioni significative nella sua deviazione standard in funzione del numero di casi. Inizialmente, il sensore mostra una notevole stabilità, con una deviazione standard che si mantiene costante. Tuttavia, si osserva un progressivo aumento della variabilità, culminando in un picco significativo nella fase finale dell'analisi. Questo comportamento suggerisce un cambiamento nelle condizioni operative o un'instabilità del sensore stesso.

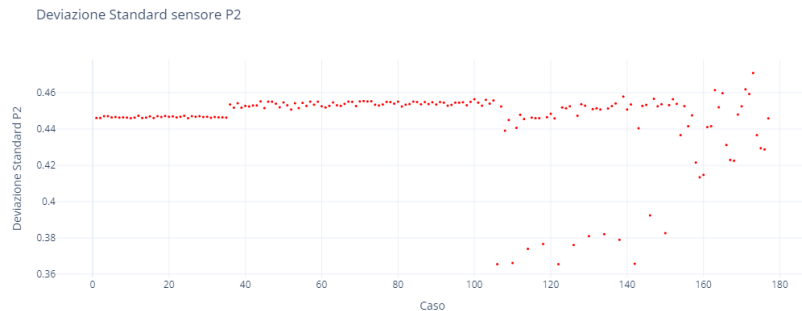


Figura 2.6: Deviazione standard sensore P2

L'analisi spettrale del sensore P2, ottenuta tramite la Trasformata di Fourier e mostrata in Figura 2.7, rivela le frequenze dominanti in diverse condizioni.

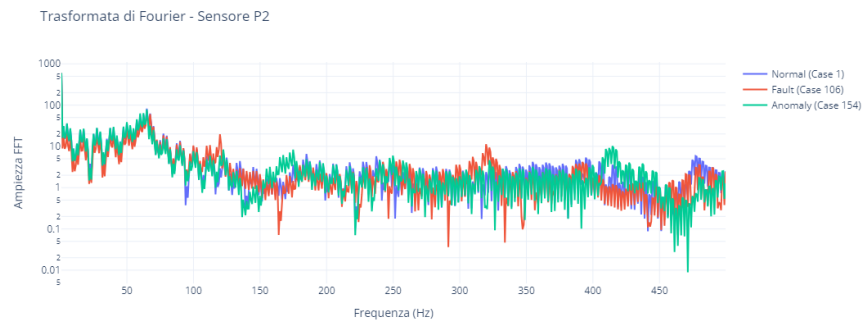


Figura 2.7: Trasformata di Fourier sensore P2

Il grafico appena mostrato ha come obiettivo evidenziare le frequenze dominanti del sistema in tre diverse condizioni operative: normale, guasto e anomalia. La trasformazione nel dominio della frequenza consente di identificare pattern ricorrenti o anomali non facilmente visibili nel dominio temporale. Differenze nello spettro di ampiezza possono essere utilizzate come indicatori per la diagnosi automatica di guasti o deviazioni dal comportamento atteso.

La feature engineering rappresenta il processo mediante il quale i dati grezzi vengono trasformati ed elaborati per estrarre un insieme di variabili significative da utilizzare come input per un modello di machine learning. In questo contesto, l'analisi è stata condotta sia nel dominio del tempo che in quello delle frequenze.

- Nel dominio del **tempo**, sono state considerate statistiche descrittive, come media, varianza e asimmetria, oltre a metriche più avanzate per caratterizzare la dinamica del segnale.
- Nel dominio della **frequenza**, invece, l'analisi è stata effettuata tramite trasformate, come la Fourier, al fine di individuare pattern e componenti frequenziali rilevanti per la classificazione.

L'obiettivo finale è quello di migliorare la capacità predittiva del modello, selezionando le caratteristiche più informative.

Sia $x = \{x_1, x_2, \dots, x_N\}$ un segnale monodimensionale discreto, composto da N campioni temporali. Le seguenti metriche statistiche sono calcolate nel dominio del tempo e sono comunemente utilizzate per l'estrazione di features da segnali 1D. Nella Tabella 3.1 sono riportate le formule esplicite associate.

Metrica	Formula
Media (μ)	$\mu = \frac{1}{N} \sum_{i=1}^N x_i$
Mediana	$\text{Mediana} = \begin{cases} x_{\frac{N+1}{2}}, & \text{se } N \text{ è dispari,} \\ \frac{x_{\frac{N}{2}} + x_{\frac{N}{2}+1}}{2}, & \text{se } N \text{ è pari} \end{cases}$
25 ^{esimo} percentile (P_{25})	$P_{25} = \min\{x \in \mathbb{R} : F(x) \geq 0.25\}$, dove $F(x) = \frac{1}{N} \sum_{i=1}^N \mathbf{1}(x_i \leq x)$
75 ^{esimo} percentile (P_{75})	$P_{75} = \min\{x \in \mathbb{R} : F(x) \geq 0.75\}$
Varianza (σ^2)	$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$
Integrale lineare	$\int_a^b x(t) dt$
Minimo	$\min\{x_1, x_2, \dots, x_N\}$
Massimo	$\max\{x_1, x_2, \dots, x_N\}$

Tabella 3.1: Statistiche nel dominio del tempo per un segnale 1D con formule esplicite per il calcolo delle features.

Sia $x(t)$ un segnale 1D continuo o campionato in funzione del tempo, e sia $S(f)$ lo *spettro di potenza* del segnale ottenuto tramite la Trasformata di Fourier. L'analisi nel dominio della frequenza consente di osservare la distribuzione spettrale dell'energia del segnale sulle varie componenti frequenziali. Nella Tabella 3.2 sono riportate alcune metriche comunemente utilizzate per descrivere tale distribuzione.

Metrica	Formula
Valore di picco	$P_{\text{peak}} = \max\{S(f)\}$, dove $S(f)$ rappresenta lo spettro di potenza.
Frequenza di picco	$f_{\text{peak}} = \arg \max_f \{S(f)\}$, ovvero la frequenza alla quale si verifica il valore massimo.
Somma dello spettro di potenza	$\sum_f S(f)$, la somma di tutti i valori dello spettro di potenza.
Deviazione standard dello spettro di potenza	$\sigma_S = \sqrt{\frac{1}{N} \sum_f (S(f) - \mu_S)^2}$, con $\mu_S = \frac{1}{N} \sum_f S(f)$.
RMS nel dominio della frequenza	$\text{RMS} = \sqrt{\frac{1}{N} \sum_f S(f)}$, la radice quadrata della media dei valori di $S(f)$.

Tabella 3.2: Metriche nel dominio della frequenza per un segnale 1D con formule esplicite.

La selezione delle caratteristiche (*features*) per ciascun task è stata effettuata tenendo conto del modello adottato e delle prestazioni ottenute nei test preliminari. Un'analisi approfondita è stata condotta per individuare le variabili più significative, utilizzando metodi statistici per determinare il loro impatto sul risultato finale.

Questo processo ha consentito di ottimizzare le sue prestazioni, migliorando la precisione delle previsioni e riducendo l'errore complessivo.

Una volta terminato quindi il processo di Feature Engineering, il `dataframe` aggregato, costruito a partire dai segnali grezzi, viene suddiviso in **tre finestre temporali distinte** corrispondenti alle tre fasi di apertura e chiusura delle valvole. Ciascuna finestra è trattata come un'istanza indipendente dello stesso caso, e da ogni `Case` originario vengono quindi generate tre righe nel `dataframe`.

Questo approccio ha permesso di **triplicare il numero di campioni** a disposizione per l'addestramento e migliorare la generalizzazione del modello, pur mantenendo la coerenza tra i dati e la loro etichetta.

4.1 Task 1

Il Task 1 si propone di classificare ogni caso come *normale* oppure *anormale*, attraverso un modello di **classificazione binaria**. Il flusso operativo è illustrato in Figura 4.1.

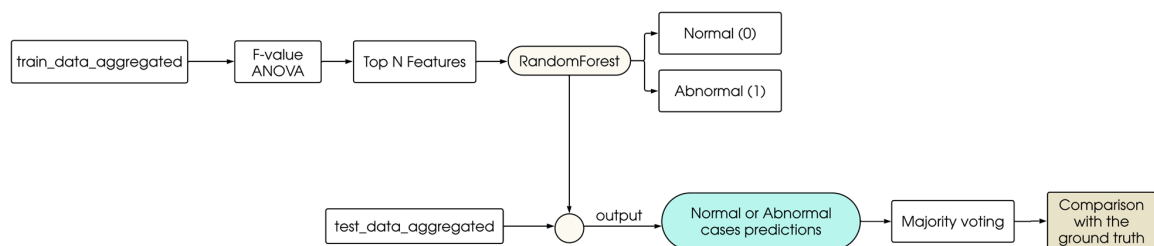


Figura 4.1: Schema operativo Task 1

Il primo passo consiste nella **selezione delle feature più rilevanti** tramite il metodo ANOVA F-value, che valuta la capacità di ogni variabile indipendente di discriminare tra le due classi del target. Tale valutazione si basa sul confronto tra la varianza tra gruppi e la varianza all'interno dei gruppi.

Le venti feature con i punteggi F-value più elevati sono riportate in Figura 4.2.

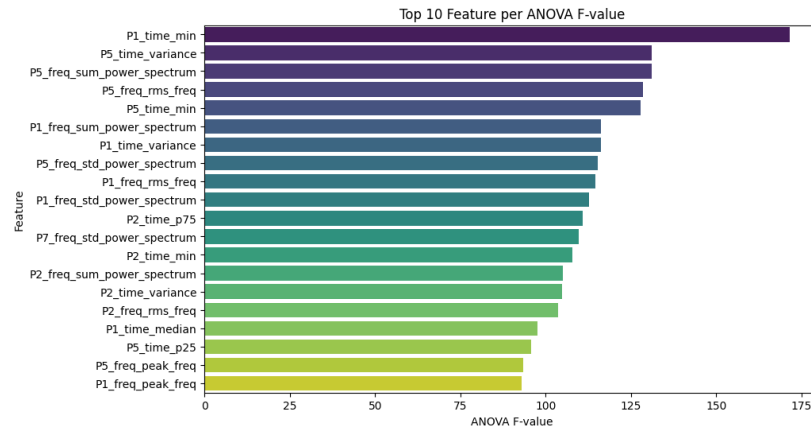


Figura 4.2: Top 20 Feature selezionate tramite ANOVA F-value

Successivamente, il dataset è stato suddiviso in training set (80%) e validation set (20%), e sulle feature selezionate è stato addestrato un modello di classificazione `Random Forest`.

La valutazione del modello sul set di validazione ha evidenziato performance eccellenti, come mostrato nella Tabella 4.1:

Metrica	Classe 0 (Normal)	Classe 1 (Abnormal)
Accuracy	0.99	
Precision	0.98	1.00
Recall	1.00	0.98
F1-score	0.99	0.99

Tabella 4.1: Metriche di valutazione sul validation set (Task 1)

I risultati ottenuti confermano l’elevata capacità del modello nel distinguere tra condizioni normali e anomale.

Infine, per quanto riguarda le predizioni sui dati di test, è stata adottata una strategia di *majority voting* sulle finestre temporali di ciascun *Case*, al fine di assegnare un’unica etichetta aggregata. Tali predizioni sono state confrontate con il *ground truth* per stimare l’efficacia del modello su dati non etichettati.

L’analisi dei risultati verrà discussa nella Sezione 5.1.

4.2 Task 2

Il Task 2 si concentra sulla classificazione dei *casi anormali* (già individuati nel Task 1) in tre categorie distinte:

- **Unknown** (Classe 1): Guasti sconosciuti, non osservati nei dati di addestramento.
- **Anomaly** (Classe 2): Anomalie causate da contaminazione da bolle d’aria (*BP Anomalies*);
- **Fault** (Classe 3): Guasti alle elettrovalvole (*SV Faults*);

Il flusso del lavoro svolto è rappresentato in Figura 4.3.

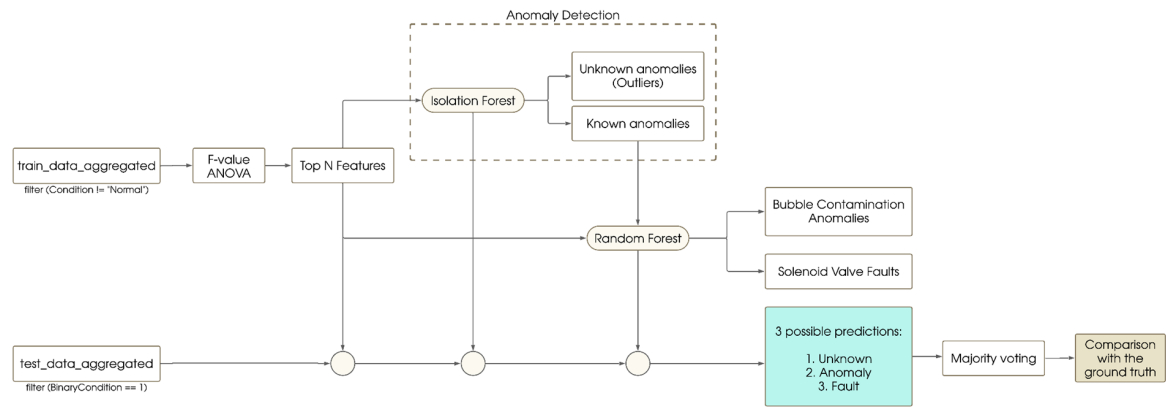


Figura 4.3: Schema operativo del Task 2

Un’importante difficoltà iniziale riguarda l’assenza, nel dataset etichettato, di esempi associati alla classe *Unknown*. Ciò impedisce l’addestramento diretto di un classificatore a tre classi. Inoltre, il dataset mostra uno sbilanciamento tra le classi *Anomaly* e *Fault*, con la seconda decisamente più rappresentata.

Per affrontare tale sbilanciamento, si è adottata una strategia combinata:

- **Undersampling** della classe maggioritaria (*Fault*);
- **SMOTE (Synthetic Minority Oversampling Technique)** per generare campioni sintetici della classe minoritaria (*Anomaly*).

Questa combinazione consente di bilanciare efficacemente il dataset evitando sia la perdita eccessiva di dati reali, sia l’eccessiva introduzione di dati artificiali.

A seguito di questa prima fase di preprocessing, sono stati effettuati i seguenti passaggi:

1. **Selezione delle feature:** è stata applicata la tecnica ANOVA F-value per identificare le feature più rilevanti rispetto al target, al fine di ridurre la dimensionalità del problema.
2. **Gestione dei guasti sconosciuti (Unknown):** poiché la classe *Unknown* non è presente nei dati di training, si è scelto di ricorrere a un Anomaly Detector non supervisionato per rilevarli. In particolare, è stato utilizzato un Isolation Forest, impostando una soglia di decisione dinamica basata sul 4° **percentile** dei punteggi ottenuti sul training set. Il modello ha mostrato un’accuratezza del **97.7%** nella rilevazione di outlier sul set di validazione, dimostrando elevata efficacia.
3. **Predizione sui dati di test:** il modello di Anomaly Detection è stato applicato al test set per identificare i guasti sconosciuti. Poiché ogni *Case* è associato a tre finestre temporali, è stato adottato un *majority voting* per assegnare un’unica etichetta consolidata a ciascun caso. I casi etichettati come *Unknown* sono stati esclusi dalla fase successiva.
4. **Classificazione binaria tra Anomaly e Fault:** sui restanti casi è stato addestrato un Random Forest, utilizzando le feature precedentemente selezionate, con l’obiettivo di discriminare tra le due classi residue. Anche in fase di inferenza, è stato impiegato il *majority voting* per garantire una predizione coerente per ogni *Case*. Le prestazioni del classificatore binario sul set di validazione sono sintetizzate nella Tabella 4.2.

Metrica	Classe 2 (Anomaly)	Classe 3 (Fault)
Accuracy	1.00	
Precision	1.00	1.00
Recall	1.00	1.00
F1-score	1.00	1.00

Tabella 4.2: Metriche di valutazione su Anomaly e Fault (Task 2)

I risultati ottenuti sul test set saranno analizzati in dettaglio nella Sezione 5.2.

4.3 Task 3

Il Task 3 ha come obiettivo l'identificazione della posizione della contaminazione da bolle d'aria all'interno del sistema. Le possibili localizzazioni sono le seguenti: BV1, BP 1, BP 2, BP 3, BP 4, BP 5, BP 6 e BP 7.

Il flusso del lavoro svolto è rappresentato in Figura 4.8.

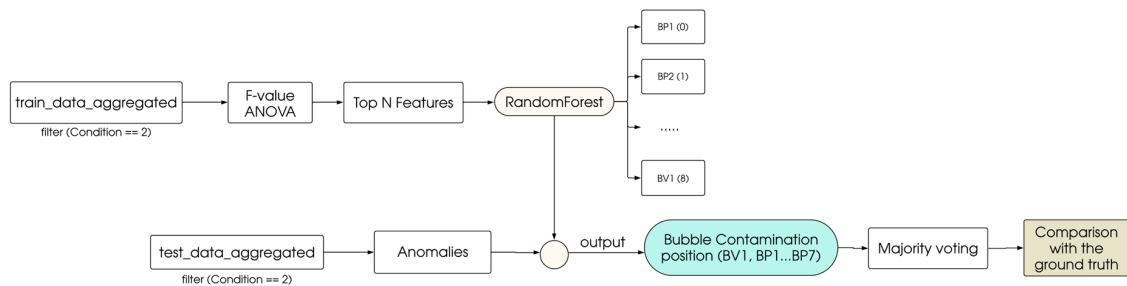


Figura 4.4: Schema operativo del Task 3

Come nei task precedenti, il primo passo consiste nell'effettuare una **selezione delle feature** utilizzando il criterio ANOVA F-value. Le feature sono state ordinate in base al loro punteggio e sono state rappresentate graficamente in Figura 4.5.

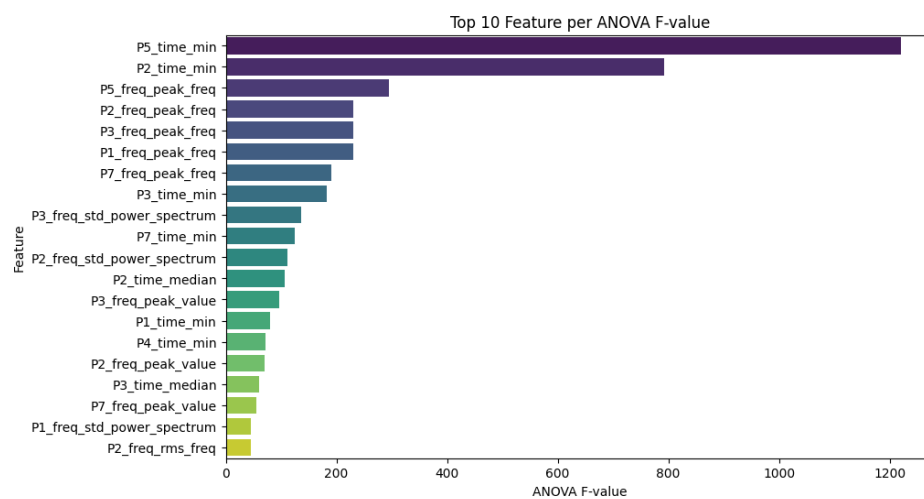


Figura 4.5: Top 10 feature selezionate secondo ANOVA F-value (Task 3)

A questo punto il dataset viene suddiviso in training set (80%) e validation set (20%), ed è stato addestrato un classificatore `Random Forest` in modalità multiclasse per la previsione della posizione dell'anomalia.

La valutazione del modello sul validation set ha prodotto le metriche riportate nella Tabella 4.3, evidenziando un'accuratezza complessiva del 93%.

Classe	Precision	Recall	F1-score
1 (BP1)	0.67	1.00	0.80
2 (BP2)	1.00	1.00	1.00
3 (BP3)	1.00	1.00	1.00
4 (BP4)	1.00	1.00	1.00
5 (BP5)	1.00	1.00	1.00
6 (BP6)	1.00	1.00	1.00
7 (BP7)	1.00	1.00	1.00
8 (BV1)	1.00	0.50	0.67
Accuracy	0.93		
Macro avg	0.96	0.94	0.93
Weighted avg	0.96	0.93	0.93

Tabella 4.3: Metriche di valutazione sul validation set per il Task 3

Il modello mostra buone capacità di generalizzazione, riuscendo ad individuare correttamente la maggior parte delle posizioni della contaminazione.

L'ultima fase prevede l'applicazione del modello al *test set*, con lo scopo di generare le predizioni finali. Come nei task precedenti, per ciascun *Case* viene adottata una strategia di *majority voting*, al fine di consolidare le tre predizioni (una per ciascuna finestra temporale) in un'unica etichetta definitiva.

I risultati ottenuti sul test set saranno analizzati in dettaglio nella Sezione 5.3.

4.4 Task 4

Il Task 4 presenta una struttura operativa molto simile a quella del Task precedente, concentrandosi però sulla classificazione precisa della posizione del guasto tra le quattro valvole presenti nel sistema di propulsione (SV1, SV2, SV3 e SV4). Dopo aver individuato i casi anomali e classificato le diverse tipologie di guasto nelle fasi precedenti, l'obiettivo di questa fase è determinare con esattezza quale valvola sia colpita dal malfunzionamento.

Lo schema operativo relativo al Task 4 è illustrato in Figura 4.8.

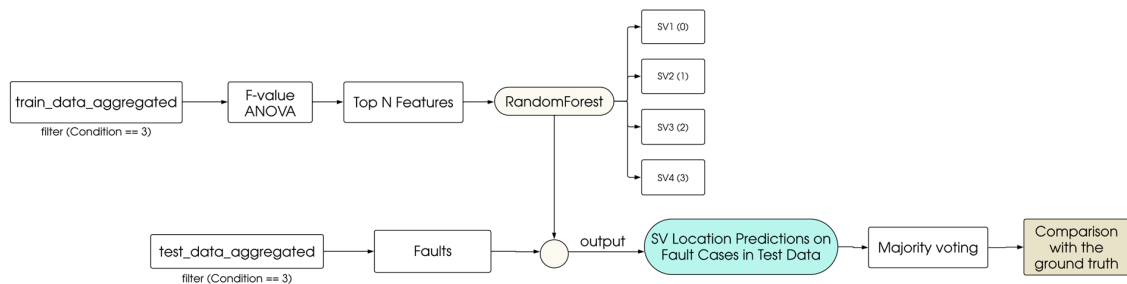


Figura 4.6: Schema operativo Task 4

Al fine di ottimizzare la classificazione, è stata effettuata una selezione delle feature più rilevanti utilizzando il metodo statistico ANOVA F-Value. Le feature selezionate con maggiore rilevanza sono illustrate nella Figura 4.7.

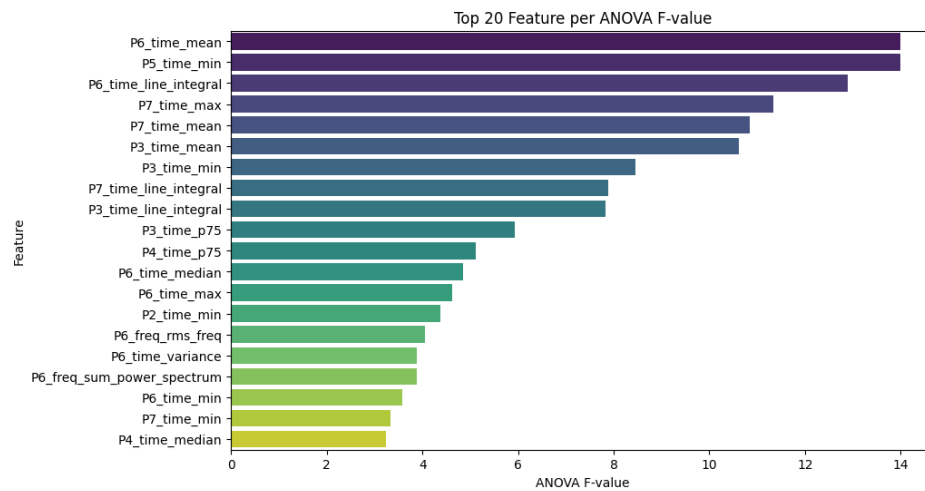


Figura 4.7: Feature principali selezionate con metodo ANOVA per il Task 4

Sulla base delle feature selezionate, è stato addestrato un modello di classificazione multiclasse basato anche qui sull'algoritmo Random Forest. Il dataset è stato suddiviso come di consueto in un training set (80%) e un validation set (20%).

Le prestazioni del modello addestrato sono state valutate sul set di validazione attraverso le seguenti metriche:

Classe	Precision	Recall	F1-score
SV1	1.00	0.86	0.92
SV2	0.88	1.00	0.93
SV3	0.89	1.00	0.94
SV4	1.00	0.86	0.92
Accuracy	0.93		
Macro avg	0.94	0.93	0.93
Weighted avg	0.94	0.93	0.93

Tabella 4.4: Risultati della classificazione sul set di validazione per il Task 4

L'ultima fase consiste nell'applicazione del modello al test set, con l'obiettivo di ottenere le predizioni finali. Anche in questo caso viene utilizzata una strategia di *majority voting* per ogni Case.

I risultati sul test set verranno analizzati nella Sezione 5.4.

4.5 Task 5

Il quinto ed ultimo task ha come obiettivo la predizione del rapporto di apertura (da 0% a 99%) per i casi in cui la valvola è identificata come guasta (*SV Fault*).

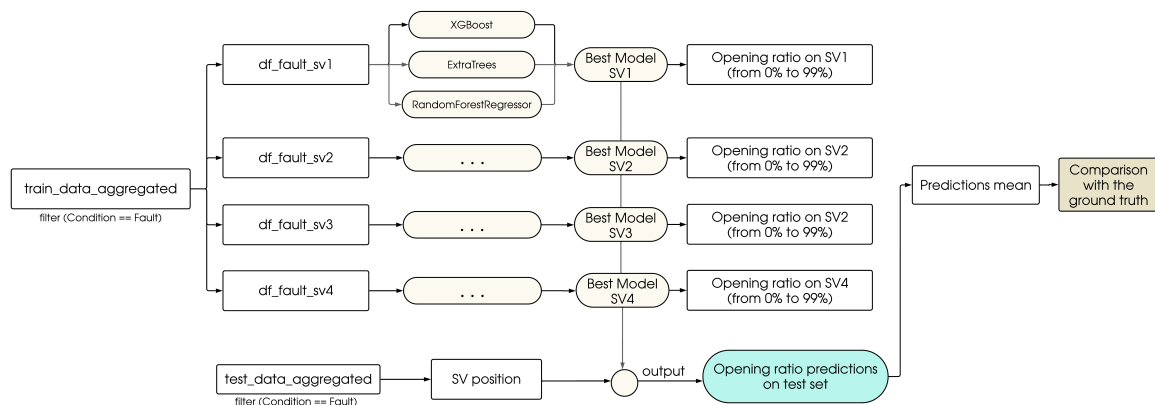


Figura 4.8: Schema operativo Task 5

Dato che ogni tipo di guasto (SV1, SV2, SV3, SV4) coinvolge meccanismi fisici differenti e può influenzare in modo diverso le variabili di sistema, la relazione tra le feature e il target (apertura della valvola) cambia sensibilmente a seconda del guasto considerato.

Per questo motivo, si è scelto di addestrare **quattro regressori indipendenti**, uno per ciascuna classe di fault.

Per ciascuno dei quattro guasti quindi, si procede nel seguente modo:

1. Si selezionano le migliori feature utilizzando il metodo `f_regression`, che valuta la correlazione lineare tra ciascuna variabile indipendente e il target, essendo questo un problema di regressione (a differenza dei precedenti Task non può essere applicato il metodo ANOVA F-value in quando il target non è una variabile categorica ma continua).
2. Vengono addestrati tre modelli di regressione: `XGBoost Regressor`, `Extra Trees Regressor` e `Random Forest Regressor`.
3. Per ognuno dei tre modelli addestrati, vengono calcolate delle metriche di valutazione per misurare la bontà delle predizioni sul set di validazione, utilizzando indicatori classici per problemi di regressione come il MAE, il RMSE e l' R^2 Score.
4. Viene selezionato il modello che ha riportato le migliori prestazioni in termini di accuratezza predittiva. Per i casi in questione, l'`Extra Trees Regressor` ha mostrato le performance più elevate per i guasti associati a SV1, SV2 e SV3, mentre per il guasto SV3 il miglior risultato è stato ottenuto con l'`XGBoost Regressor`.

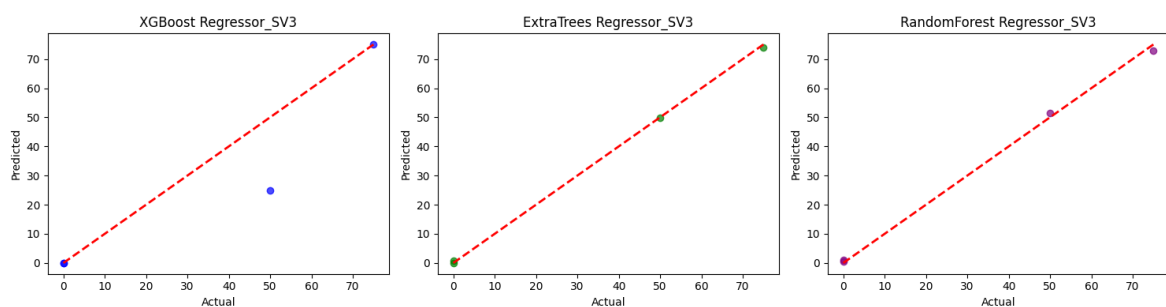


Figura 4.9: Modelli di regressione a confronto sulla valvola SV3 in fase di validazione

Terminata la fase di addestramento quindi, i dati di testing vengono importati e filtrati per i soli casi predetti come guasti (nel Task 4). A questo punto, ad ogni istanza viene applicato il regressore specifico per ottenere la predizione corrispondente.

Tuttavia, differenza dei primi quattro task, per ciascun *Case* non è possibile applicare una strategia di *majority voting* per consolidare le tre predizioni in un'unica etichetta, poiché non si tratta di predizioni categoriche. Per questo motivo, si è optato per una soluzione più semplice e diretta: l'aggregazione delle tre predizioni relative a ciascun *Case* tramite media aritmetica.

Le metriche ottenute sui dati di testing, assieme ai rapporti di apertura delle valvole predetti nello stesso dataset, sono presenti nella Sezione 5.5.

Risultati e confronto sul Test Set

Nel seguente capitolo sono presentati i risultati ottenuti, ovvero le predizioni ottenute sui dati di test all'interno di ogni singolo Task. Per quanto riguarda il ground truth invece, esso è stato fornito dalla competizione attraverso il file *answer.csv*.

Le tabelle riportate di seguito permettono quindi di effettuare un confronto diretto tra le predizioni generate e le condizioni effettive, consentendo una valutazione puntuale dell'accuratezza del modello. Ove rilevante, verranno riportate anche metriche sintetiche di performance per ogni Task.

5.1 Task 1

I risultati ottenuti nel primo Task sono descritti dalla seguente matrice di confusione (Figura 5.1), con relativa *Accuracy*:

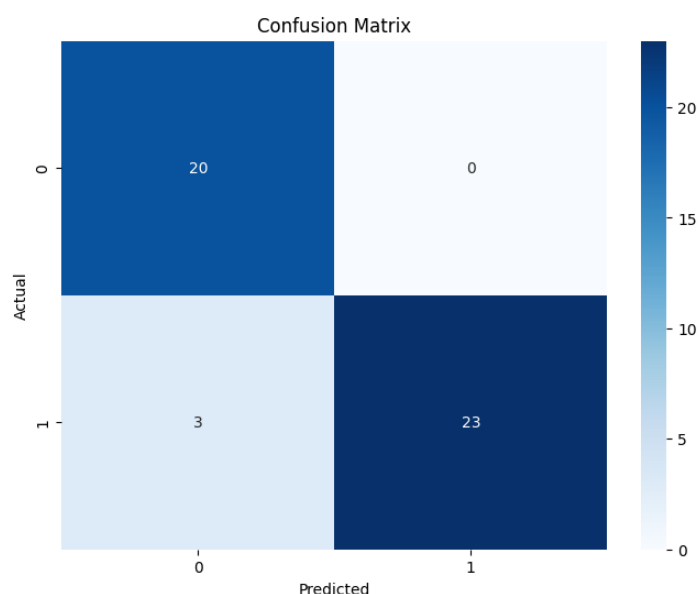


Figura 5.1: Matrice di confusione Task 1 relativa al testing set

Accuracy: 0.93

I 3 casi predetti erroneamente sono descritti nella tabella che segue.

Si noti come in tutti e tre i casi, i *Case* sono stati classificati come *Normal*, quando sono in realtà casi di *Abnormal*.

Spacecraft N.	Case	Predicted Condition	Expected Condition
1	199	0	1
4	205	0	1
4	211	0	1

5.2 Task 2

I risultati complessivi ottenuti nel secondo Task sui dati di test sono descritti dalla seguente matrice di confusione (Figura 5.2), con relativa *Accuracy*:

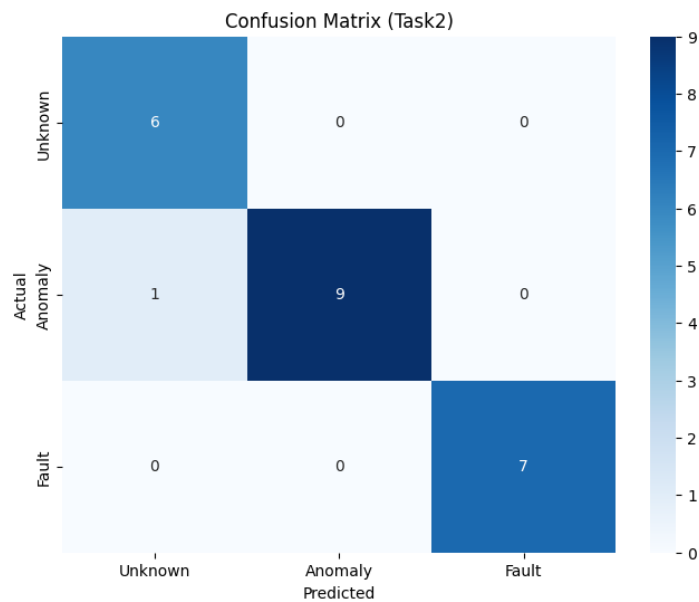


Figura 5.2: Matrice di confusione Task 2 sul testing set

Accuracy: 0.96

L'unico *Case* classificato erroneamente è il seguente:

Spacecraft N.	Case	Predicted Condition	Expected Condition
4	204	1	2

5.3 Task 3

I risultati complessivi ottenuti nel terzo Task sui dati di test sono descritti dalla seguente matrice di confusione (Figura 5.3), con relativa *Accuracy*:

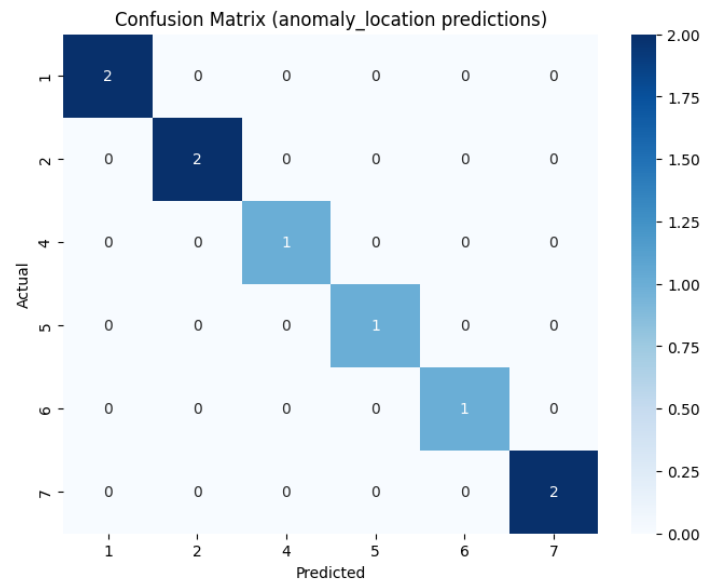


Figura 5.3: Matrice di confusione Task 3 relativa ai dati di testing

Accuracy: 1.00

Per questo Task non vi sono stati *Case* classificati erroneamente.

5.4 Task 4

I risultati complessivi ottenuti nel quarto Task sui dati di testing sono descritti dalla seguente matrice di confusione (Figura 5.4), con relativa *Accuracy*:

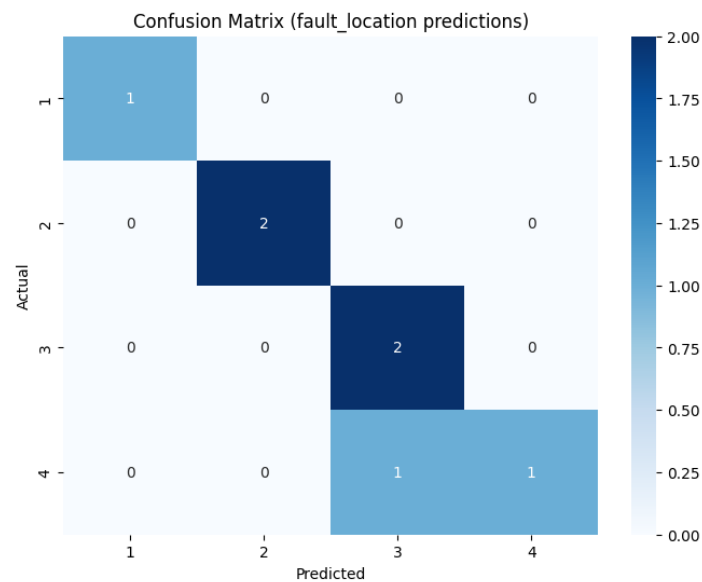


Figura 5.4: Matrice di confusione Task 4, testing set

Accuracy: 0.86

L'unico *Case* classificato erroneamente è stato il seguente:

Spacecraft N.	Case	Predicted SV Position	Expected SV Position
4	214	3	4

5.5 Task 5

Le metriche ottenute dai regressori sui dati di testing sono state le seguenti:

Metrica	Valore
Mean Absolute Error (MAE)	11.7
Mean Squared Error (MSE)	194.0
Root Mean Squared Error (RMSE)	13.9
R^2 Score	0.70

Tutte le metriche sono state implementate utilizzando le funzioni messe a disposizione dalla libreria `sklearn.metrics`, applicate direttamente alle colonne del DataFrame contenente i risultati delle predizioni sul test di set.

Per i *Case* di test correttamente classificati come *SV Fault* nel Task 2, le percentuali di apertura delle valvole sono state le seguenti:

Spacecraft N.	Case	Opening ratio predicted (%)	Opening ratio expected(%)
1	179	26	22
1	181	75	76
1	188	24	5
1	190	49	46
4	202	55	44
4	212	63	70
4	214	64	24

Si noti che nel **Case 214** vi è un errore di 40 punti percentuale, questo risultato è spiegato dal fatto che nel Task precedente, questo campione era stato classificato come Fault nella SV3, anziché nella SV4.

Nel procedimento svolto all'interno del quinto Task quindi, al caso in questione è stato applicato il regressore specifico del Task 3, producendo così una predizione parecchio distante dal valore atteso.

5.6 Calcolo dello score ottenuto

5.6.1 Metriche di valutazione della competizione

Il sistema di valutazione della competizione assegna punteggi parziali in base alla correttezza delle predizioni effettuate durante il processo. Di seguito sono riportate le metriche utilizzate:

Tabella 5.1: Punteggi assegnati per ciascuna fase

Fase di valutazione	Punteggio
Class. la condizione (Task 1)	10 pt (se corretta)
Class. il tipo di anomalia (Task 2)	10 pt (se corretta)
Identificazione della posizione della bolla (Task 3)	10 pt (se corretta)
Identificazione della valvola SV (Task 4)	10 pt (se corretta)
Rapporto di apertura (Task 5)	$\max(- \text{truth} - \text{prediction} + 20, 0)$

NB: Per i casi relativi allo **Spacecraft 4**, tutti i punteggi assegnati vengono **raddoppiati**, in quanto considerati più complessi da analizzare.

5.6.2 Punteggio finale

Il modulo di diagnosi sviluppato ha ottenuto un punteggio percentuale complessivo dell'**85.82%**.