

# Maintenance of Naval Power Plants

*Data Science and Management Project*



**Alice Cavazzuti**

*Matr. 213661*



**Afzal Shaesta Jabbeen**

*Matr. 212675*



# Descrizione del Dataset e Integrità Strutturale



**Contesto:** Dati generati da un simulatore ad alta fedeltà di un impianto di propulsione navale (Gas Turbine).



## Dimensioni

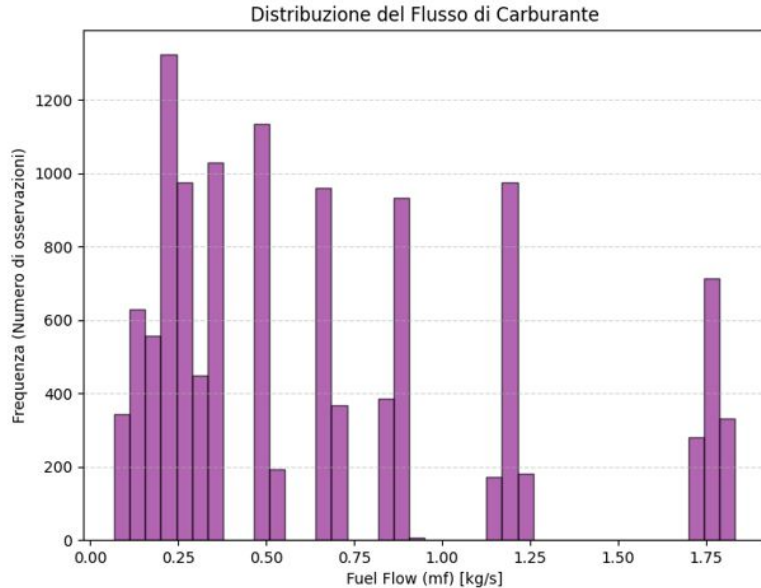
- **11.934** istanze (righe).
- **18** variabili (colonne): 16 sensori di input + 2 coefficienti di decadimento (target).



## Verifica Qualità del Dato

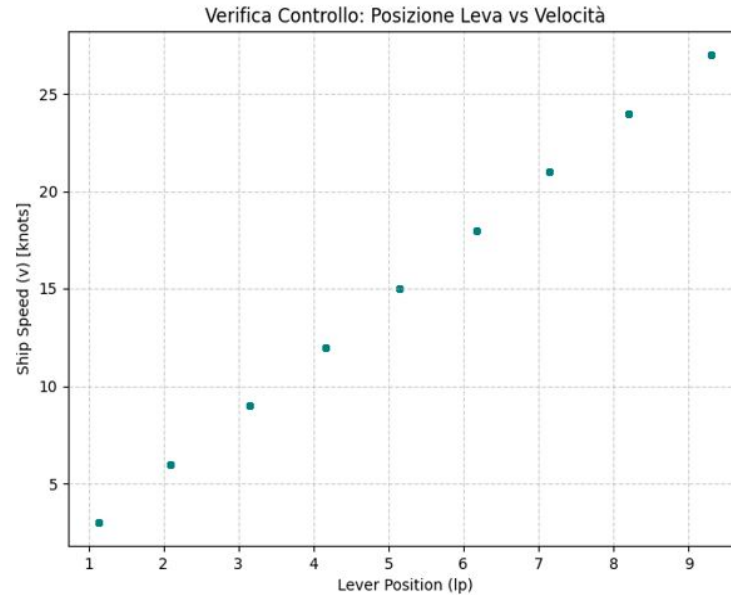
- **Completezza:** Assenza totale di valori mancanti (**0 Null Values**). Non è stata necessaria imputazione dati.
- **Formato:** Tutte le feature sono numeriche (**float64**), pronte per l'elaborazione algoritmica.

# Esplorazione dei Dati



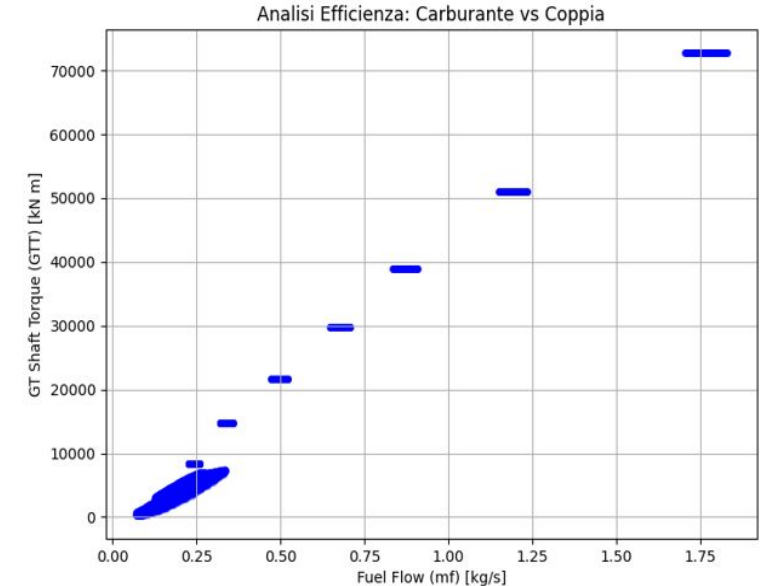
## Regimi Operativi

La distribuzione multimodale di *mf* (*Fuel Flow*) conferma che la nave lavora su specifici punti di carico discreti (es. crociera, alta velocità) e non in modo casuale.



## Verifica della Logica di Controllo

Il grafico evidenzia una correlazione strettamente monotona tra la Posizione della Leva (*lp*) e la Velocità della Nave (*v*)



## Efficienza Termodinamica

Il grafico mostra una relazione lineare tra il Flusso di Carburante (*mf*) e la Coppia all'Albero (*GTT*), il che significa che ogni incremento di energia chimica si traduce in un aumento proporzionale di potenza meccanica.

# Controllo di Coerenza dei Dati - Outlier



## 2 APPROCCI



### Controllo nel dataset stesso:

Controllo nel dataset stesso:  
Individuazione di valori molto diversi  
dalla media dei valori per quel  
parametro

01

02

### Confronto valori nel dataset con valori Reali:

ricerca su internet del range normale di valori  
di ciascuna variabile per una nave militare,  
Controllo con codice Python se effettivamente  
i valori nel dataset rientrano in questo range,  
Identifico il numero di outliers per ogni  
variabile



# Creazione del Database



01

Dimensional  
Fact Model

02

Star  
Schema

03

Star Model

04

Creazione  
database  
naval\_vessel

MySQL 8.0

05

Caricamento dati dal  
file csv  
nel database  
naval\_vessel

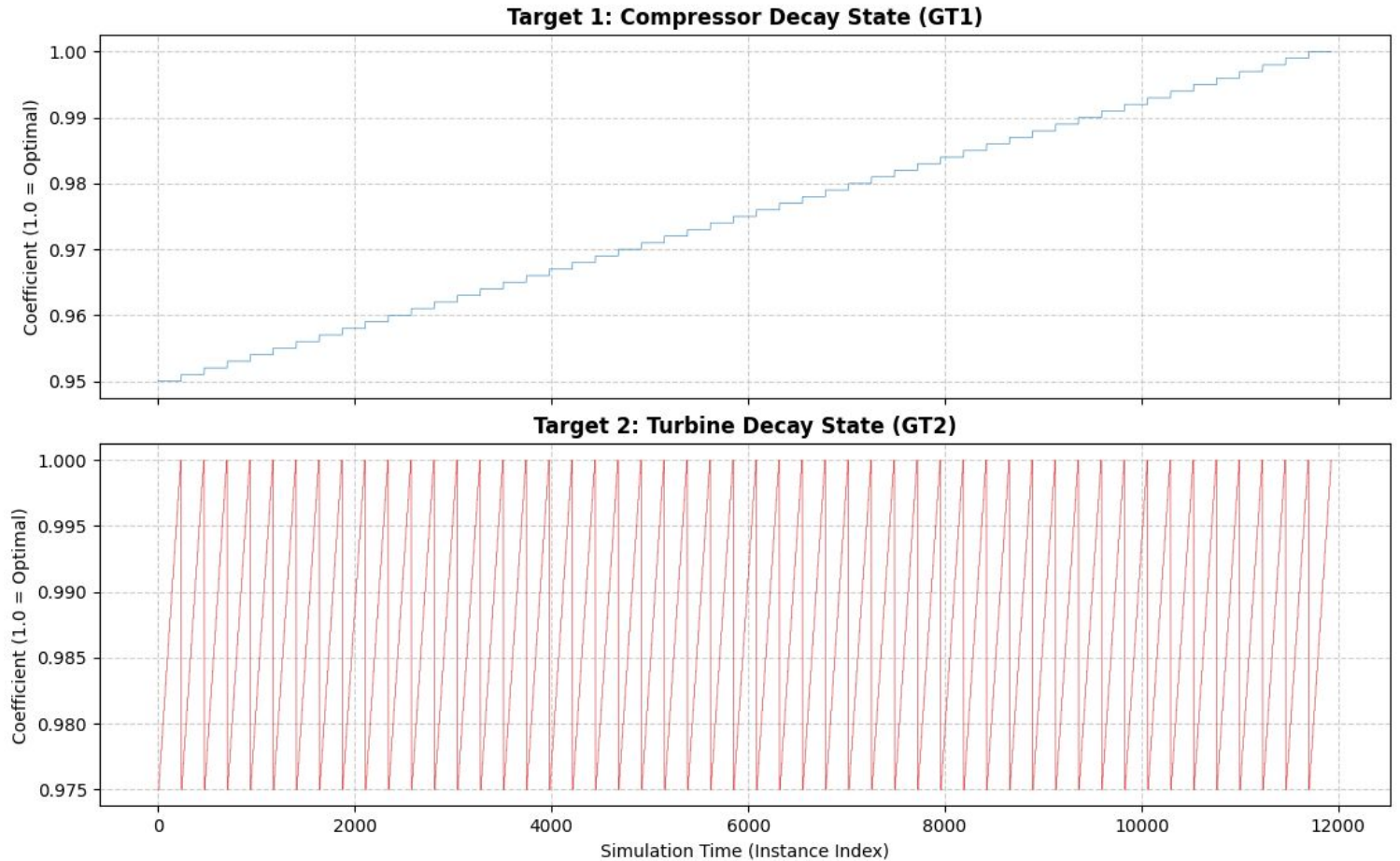
Pycharm



# Analisi Temporale del Decadimento



- Obiettivo:** Analizzare l'evoluzione dell'usura di Compressore (GT1) e Turbina (GT2).
- GT1 (Compressore):** Mostra un andamento a scala. Il simulatore testa un livello di usura fisso prima di passare al successivo.
- GT2 (Turbina):** Oscilla coprendo tutto il range [0.975 - 1.0] per ogni singolo gradino del compressore.
- Conclusione:** Il dataset copre sistematicamente tutte le combinazioni di guasto possibili



# Predizione del Decadimento del Compressore e della Turbina



DATA

70% train set

30% test set

Sul train set l'**algoritmo impara le regole** per dedurre il valore del decadimento del compressore date le altre colonne in input, (stesso procedimento anche per decadimento della turbina)  
L'algoritmo **utilizza il modello di regressione lineare**

Sul 30% dei dati usati per testare il modello (utilizzando le altre colonne in input)  
**Il modello predice il decadimento del compressore/della turbina**

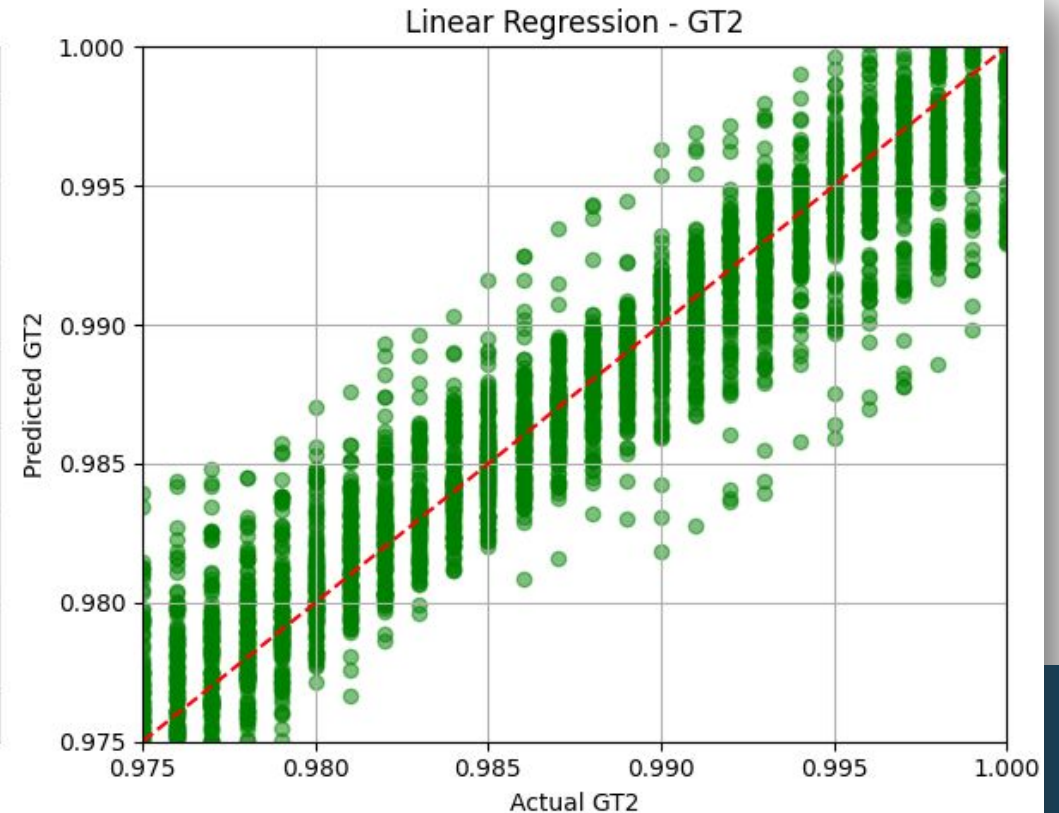
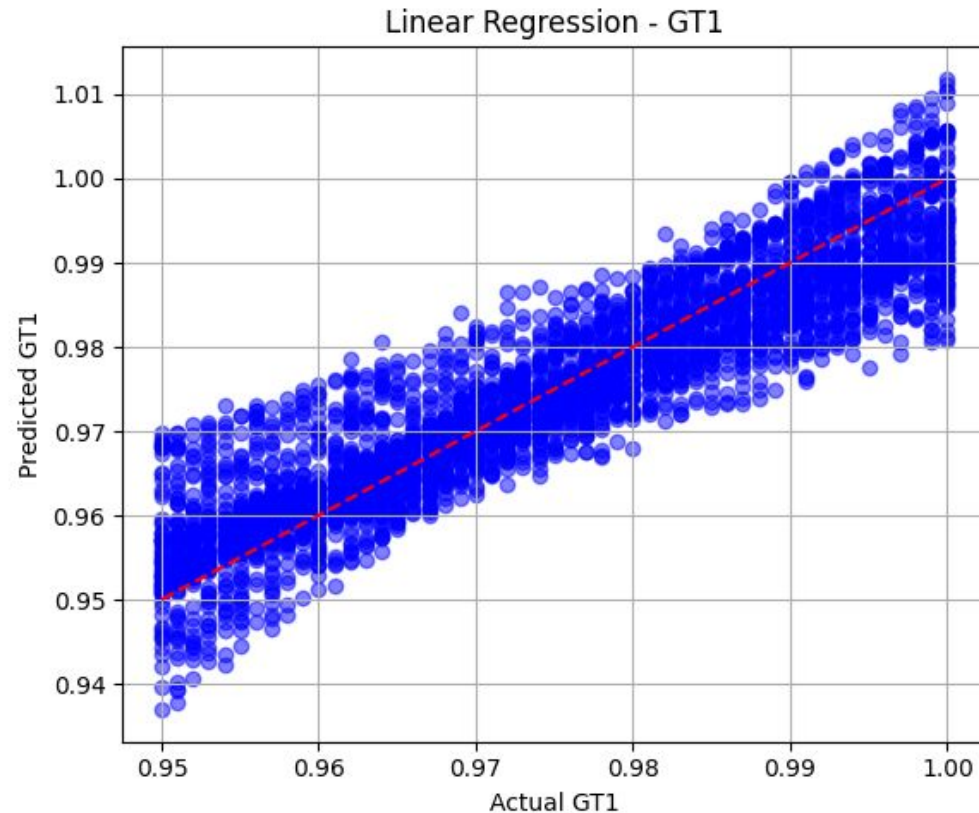
**si confrontano i valori predetti** del decadimento del compressore/della turbina **con i valori reali**

E quindi si calcola l'errore, tramite gli indicatori **MAE** e **MSE**

**Linear  
Regression**



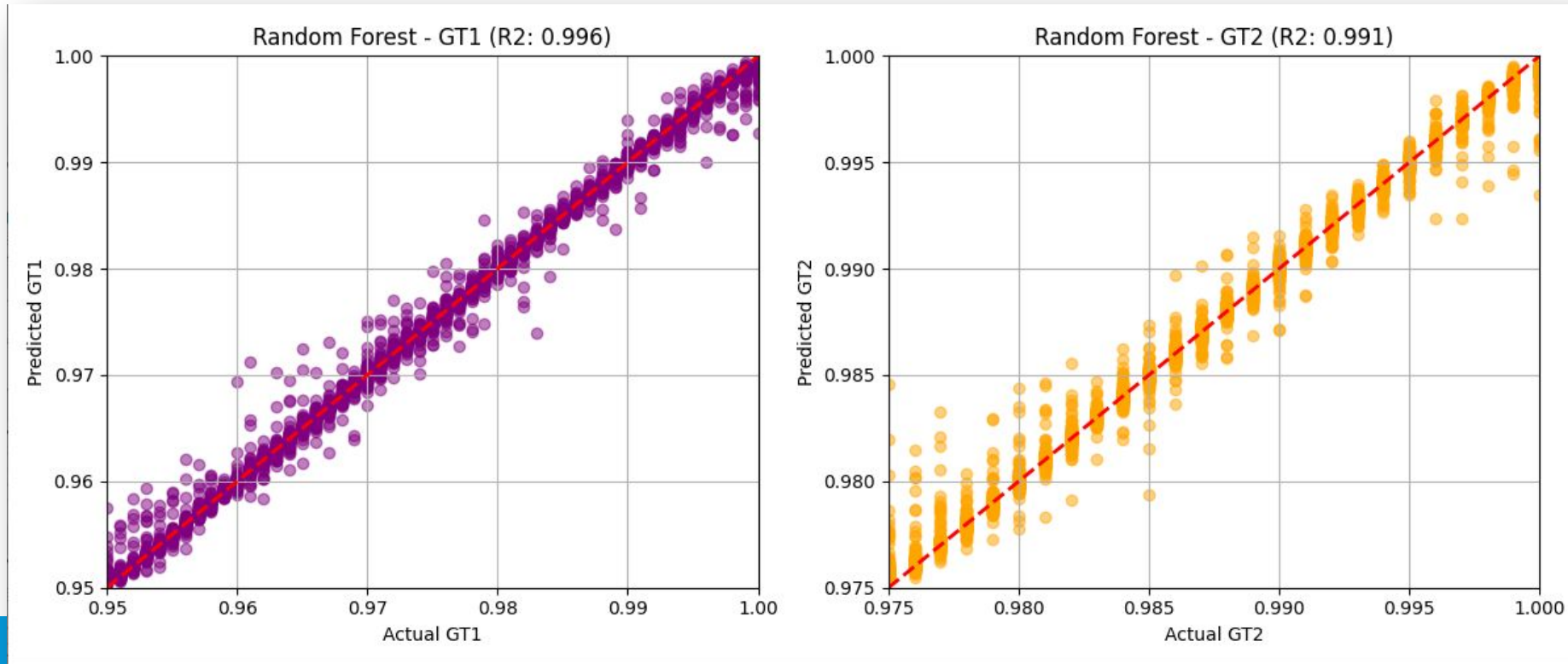
# Linear Regression



Il modello Linear Regression è STABILE e AFFIDABILE



# Random Forest



**MODELLO :** Utilizzo di un algoritmo a 100 alberi decisionali per catturare le relazioni non lineari.

•**ALTA PRECISIONE:** I grafici mostrano una dispersione quasi nulla (i punti seguono perfettamente la diagonale rossa).

## RISULTATI:

- **$R^2$  Score:** 0.99
- Errore (MAE):** ridotto a 0.0005.



Grazie per l'attenzione!