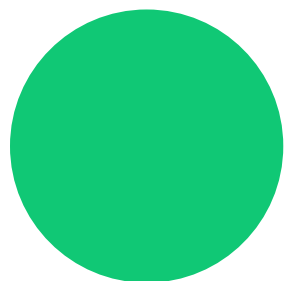


Machine Learning.

Progetto di Paolo Acone



Indice contenuti



Obiettivo del progetto

Preprocessing dei dati

Modello utilizzato

Risultati e conclusioni

Notebook completo

Obiettivo

Il progetto ha lo scopo di costruire un modello di Machine Learning per prevedere la **probabilità di sopravvivenza dei passeggeri del Titanic**.

Dataset: Titanic_sub

Features principali:

- Sesso (Maschio/Femmina)
- Età
- Classe del biglietto
- Porto di imbarco



START2IMPACT

Preprocessing dei dati

- **Separazione del target**, ovvero la colonna che ci indica se una persona è sopravvissuta o meno, dalle altre caratteristiche
- **Codifica** delle variabili categoriche:
 - Sesso → Label Encoding (0 = femmina, 1 = maschio).
 - Porto di imbarco → One-hot encoding per trasformare la variabile in numerica
- **Conversione** dei dati in intero per compatibilità con il modello
- **Suddivisione del dataset**:
 - Training set (75%) - di cui il 25% destinato al Validation Set
per scegliere la profondità ottimale.
 - Test set (25%) → per valutare le prestazioni finali

Modello utilizzato

Decision Tree

Ottimizzazione del modello

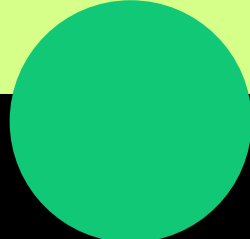

- Testato l' iparametro k , rappresentante la massima profondità dell'albero, con valori: 2, 5, 10, 25 e profondità massima sui dati del **set di validazione**

In particolare otteniamo che la versione ottimizzata del modello si ha con il valore:

- **$k=10$**

Osservazioni principali

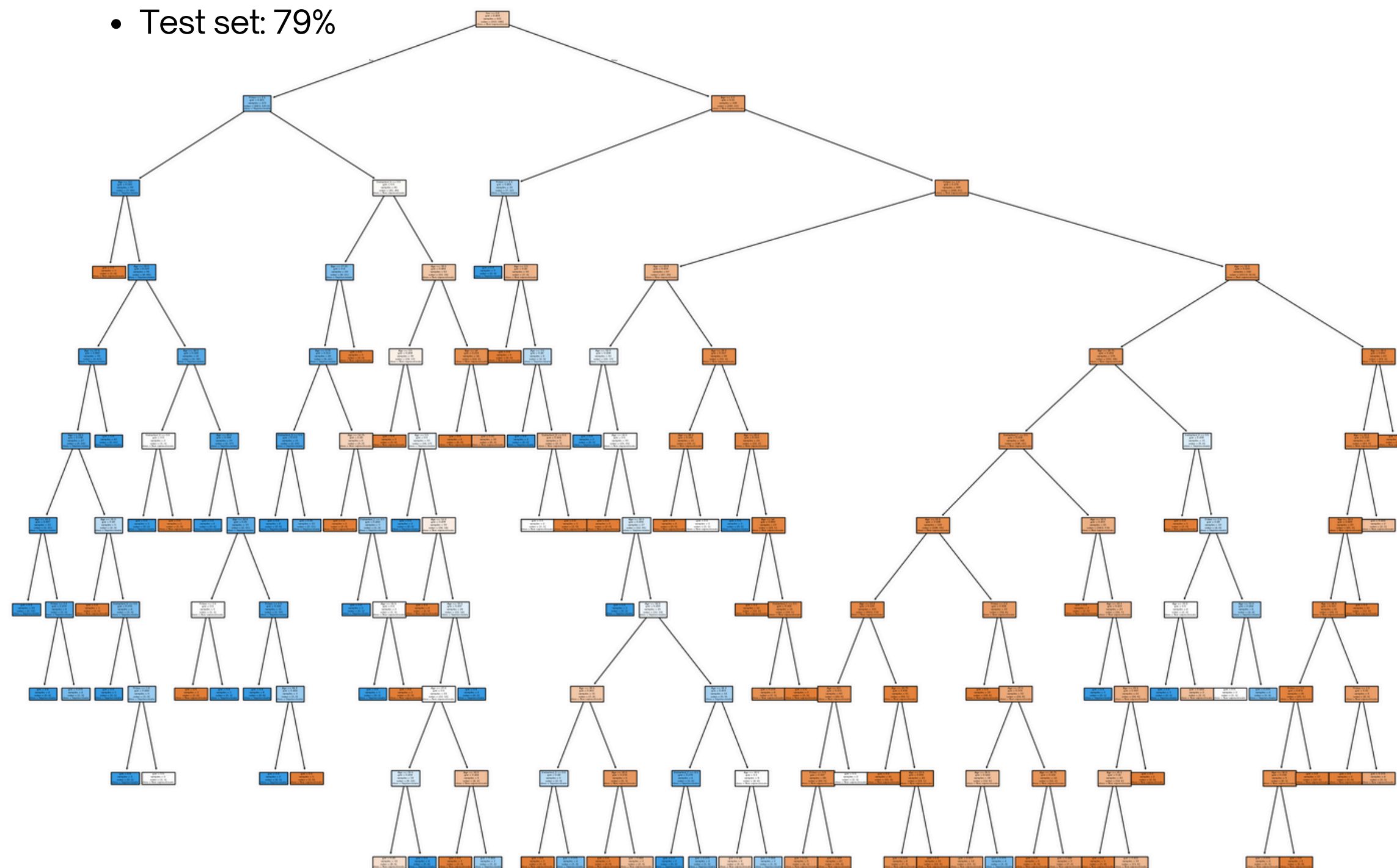
- Il valore ottimale di k pari a 10 ci dà un modello abbastanza complesso
- L'albero ha un cospicuo numero di nodi, con un'analisi dettagliata delle variabili;
- La struttura sottolinea l'importanza di sesso, classe e età, con una segmentazione articolata;
- Il modello mantiene bene l'accuratezza sui dati di test, siamo sul 79% rispetto all'82% risultante dai test di validazione.



Risultati e Conclusioni

✓ Accuratezza del modello con **k=10**

- Validation set: 82%
- Test set: 79%



Albero decisionale



Il notebook completo con codice e analisi è
disponibile ai seguente link

[Link al notebook](#)

Notebook Completo.

Grazie
per l'attenzione

Paolo Acone