DICEMBRE 2024 MILANO

Hackathon.

L'Intelligenza artificiale al servizio dell'oncologia.

Event: Al Hackathon

Date: December 3° 2024

Location: Milan

TEAM 5:

Mentor: Flavio Sartori

Giovanni Tornaghi Mattia Birti Alessandro Forcina-Barrero Matteo Fogliata Selenia Caruso

Predire i Tempi di Sopravvivenza per i Pazienti con Leucemia Mieloide Acuta (AML)

Obiettivo del Progetto

- Sviluppare modelli predittivi per stimare i tempi di sopravvivenza dei pazienti, misurati in mesi, con la maggiore accuratezza possibile.
- Valutare i modelli utilizzando il **Concordance Index (C-index)**, una metrica comune per i compiti di predizione della sopravvivenza.

Dataset

Dataset creato da Humanitas

Problematiche riscontrate nel dataset e soluzioni

Diagnosi nei dati di Utilizzo mediana Missing dei dati test non presenti nei Accorpate per imputare valori dati di training Diagnosi dei Dati su diverse Standardizzazione Trasformazione one pazienti in formato scale di misura dei valori hot encoding testuale Alta collinearità tra Colonne non Creazione di una due features informative Eliminate nuova features relazione al fegato (ratio AST ALT) (costanti) (AST, ALT)

DICEMBRE 2024 HACKATHON.

PIPELINE DATI

Pre-processing e training dei dati

丛

Grid Search

Esplorare combinazioni di iperparametri per ottimizzare il modello

 Z_{c}

Cross-validation

Dividere il dataset in fold per migliorare la stima delle performance e ridurre l'overfitting 00

Stratification

Assicurare una distribuzione bilanciata delle classi nel target

<u>\$</u>

VIF

Misurare la multicollinearità tra le variabili indipendenti DICEMBRE 2024 HACKATHON.

FINAL C-INDEX SCORES ON PUBLIC TEST SET

The following are the results in output on public test set for each model chosen:



Base Cox PH model without any preprocess:

C-index = 0.77



Random Survival Forest with basic preprocess

C-index = 0.76



Gradient Boosting with basic preprocess

C-index = 0.74



Base Cox PH model with basic preprocess:

C-index = 0.70

EXPLAINABILITY: SHAP

response chemo Complete Response response chemo Unknown

Il modello allenato suggerisce che:

- Pazienti che hanno subito un trapianto sono soggetti a minor rischio rispetto a coloro che non lo hanno subito
- Pazienti con mutazione tp53, tipicamente associata a prognosi avverse, sono soggetti a maggior rischio rispetto a colore che non presentano la mutazione
- Pazienti che rispondono a un trattamento di chemioterapia sono soggetti a minor rischio rispetto a coloro che non rispondono alla terapia



