

8 luglio 2024

Recommendation System

Marketing Analysis

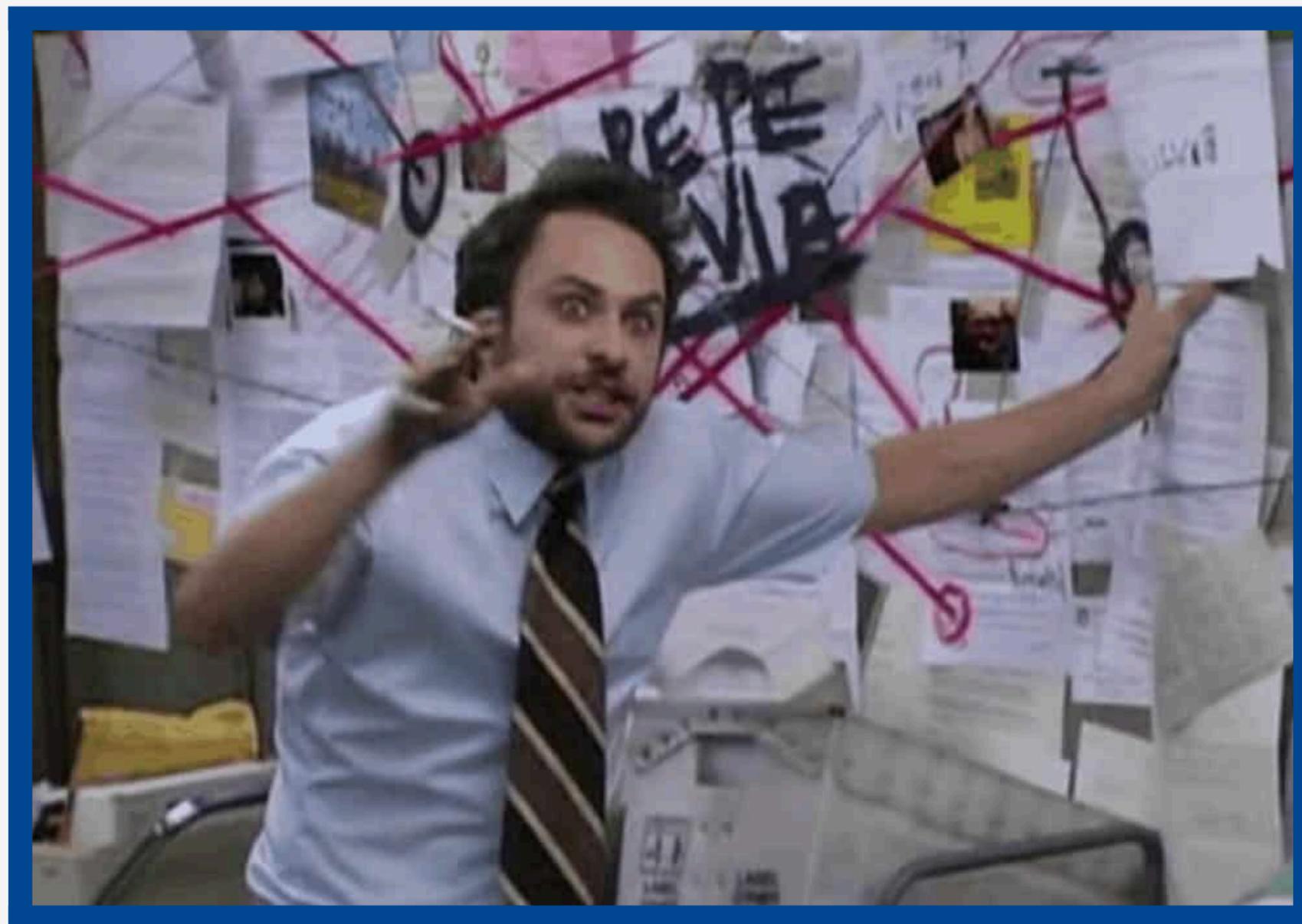
CESMA x  TIM

Roadmap ideale

- 1 ANALISI
ESPLORATIVA**
- 2 MODELLO**
- 3 RISULTATI**



Roadmap reale



Datasets



ACTIONS

Contiene tutte le chiamate con gli esiti delle proposte per offerta.



FEATURES

Contiene i profili aggiornati per ogni cliente, rappresentati da 64 componenti principali.

Analisi esplorativa



Ortogonalità delle 64 componenti principali



Correlazioni tra tutte le actions



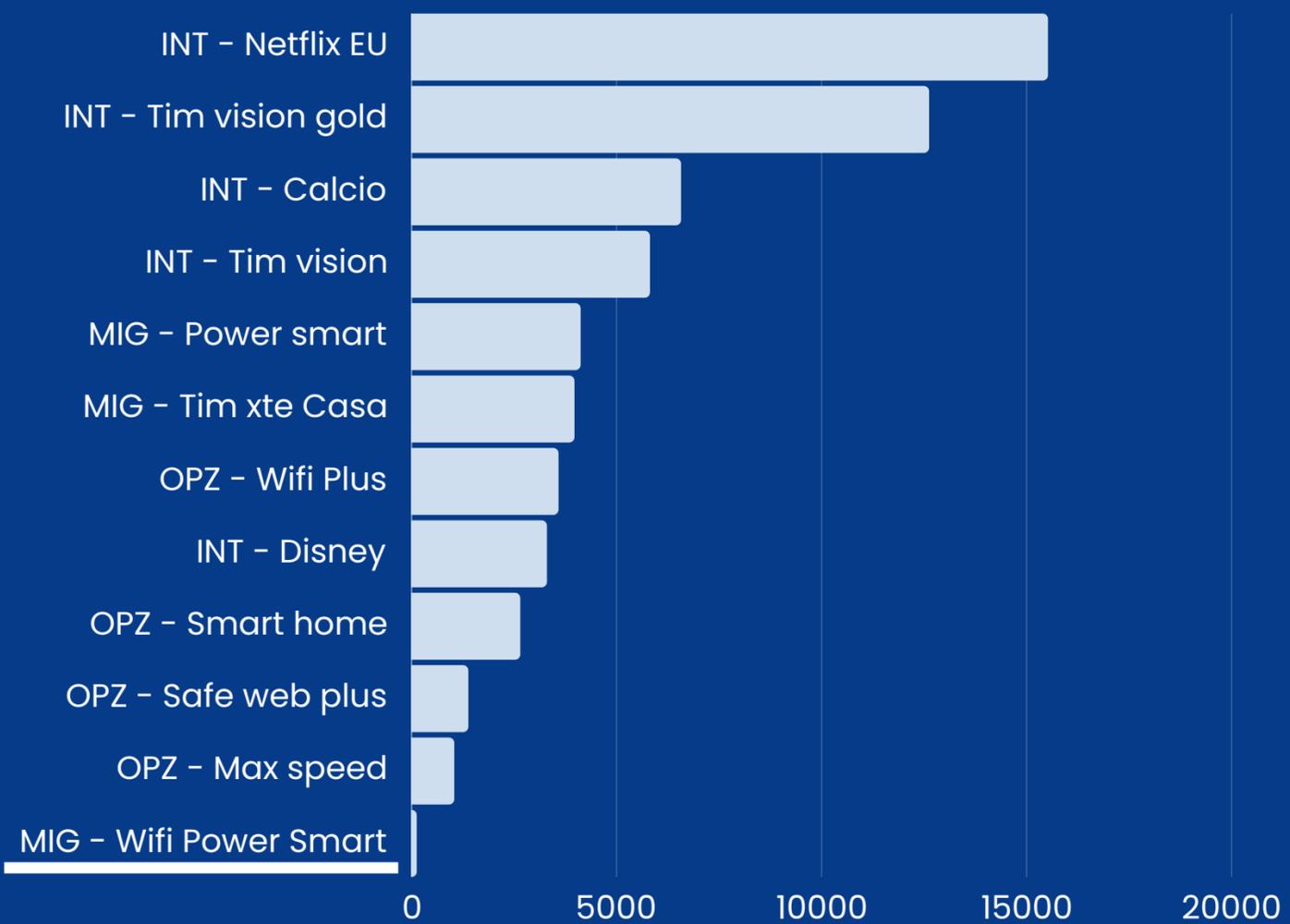
Distribuzione della variabile target



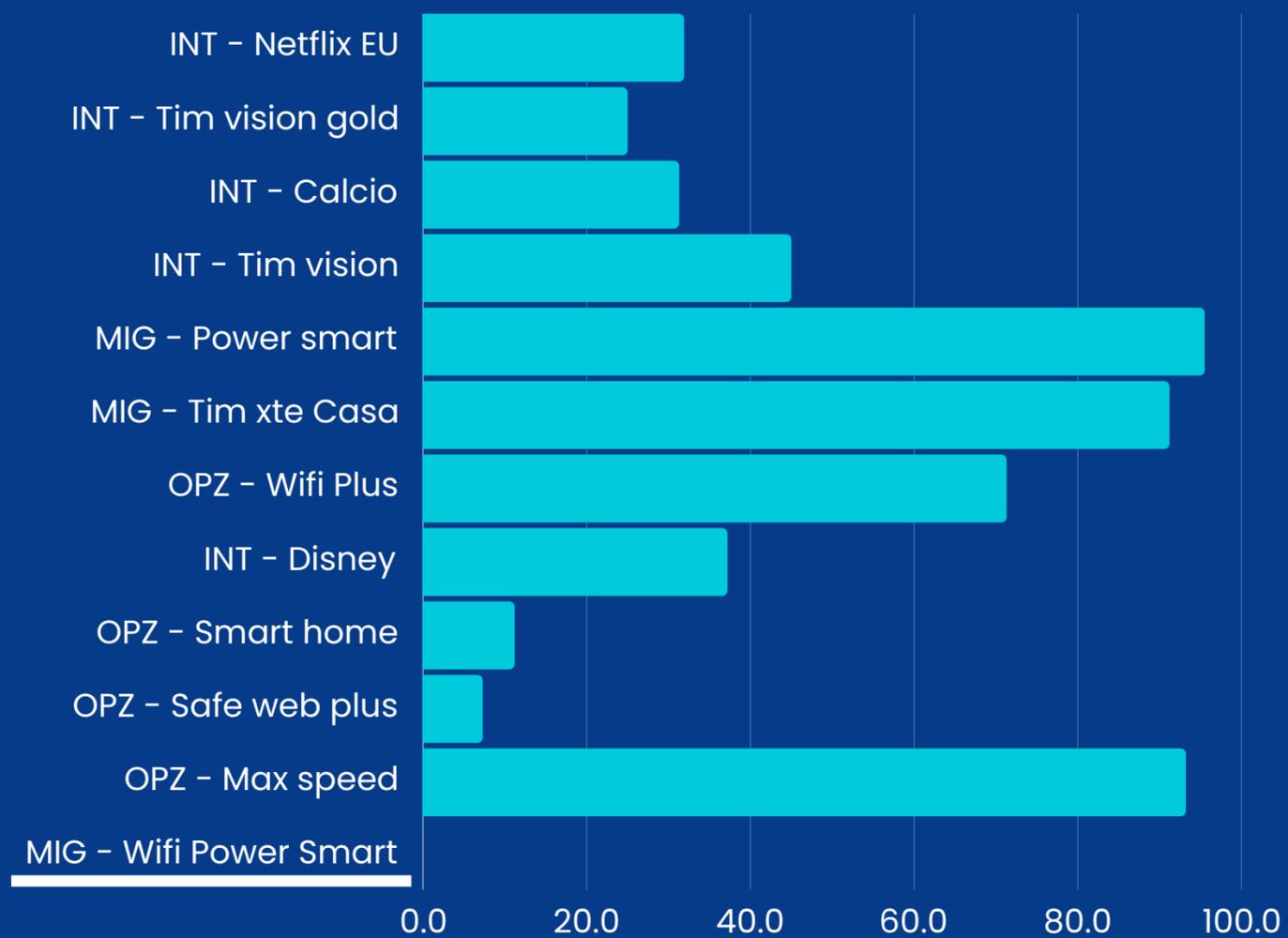
Calcolo delle distanze euclidee



Proposte

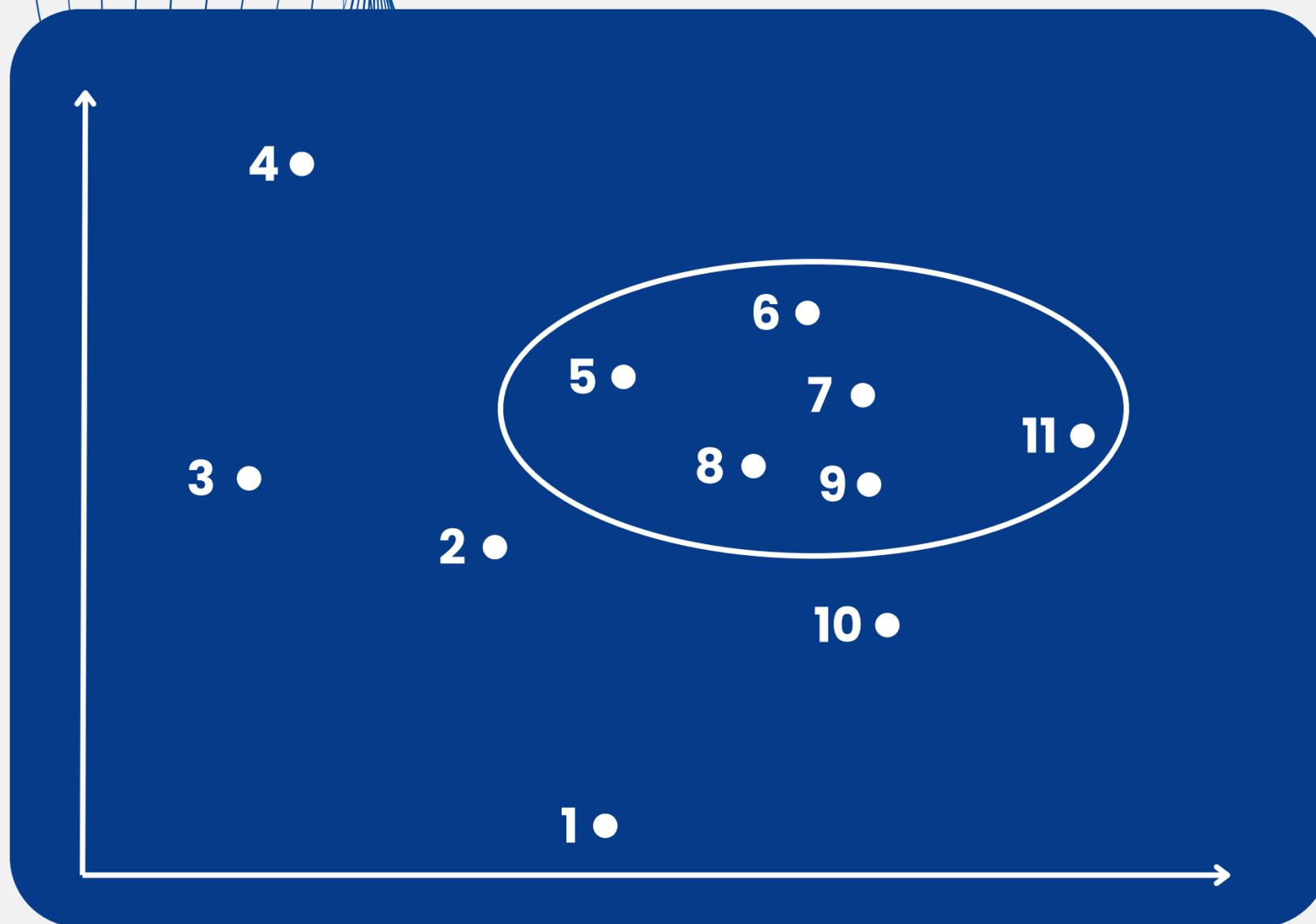


Accettati





Profili medi utenti



Si considerano gli utenti che hanno accettato una singola offerta.

- | | | | |
|---|--------------|----|-----------------|
| 1 | Smart Home | 7 | Netflix EU |
| 2 | Tim xte casa | 8 | Tim vision |
| 3 | Max speed | 9 | Tim vision gold |
| 4 | Power smart | 10 | Safe web plus |
| 5 | Disney | 11 | Wifi plus |
| 6 | Calcio | | |

Pre-processing

1

**Unione dei
datasets**

2

**Pulizia del
dataset**

3

**Creazione di
nuove variabili**

Unione dei datasets

60538
final obs

• Numero di telefono

• Data contatto x = Data contatto y

• Eliminazione servizio mai accettato

• One hot encoding di action

Nuovo dataset

Numero contatti per cliente

Dummy Weekend

Distanze euclidee

79
variabili

Modelli proposti



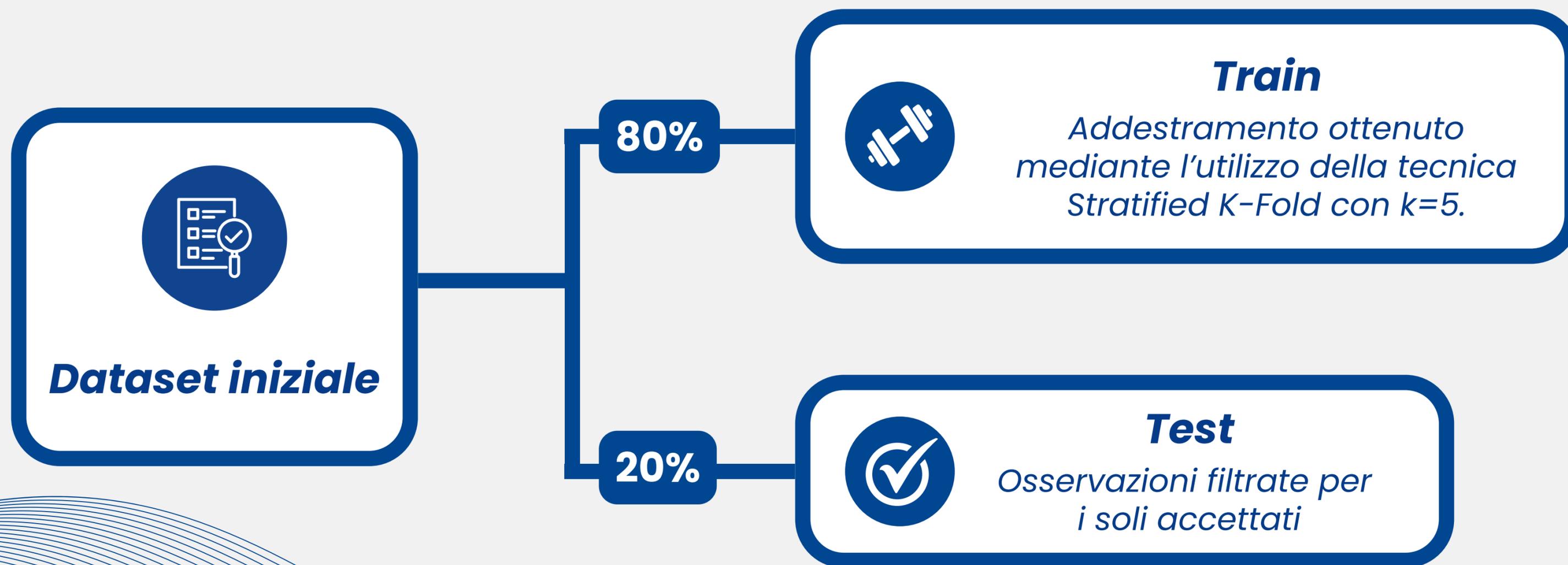
Modelli specializzati

Per riuscire a catturare le informazioni di ogni offerta, si sono divisi i dataset e si sono addestrati 11 modelli separatamente, per poi riunire i risultati e visualizzare tutte le probabilità in maniera ordinata.



Rete neurale

Train e test



Metodo



Ciclo FOR

Selezione dei sottodataset basata sul servizio proposto



Modello

Selezione del modello per la migliore previsione

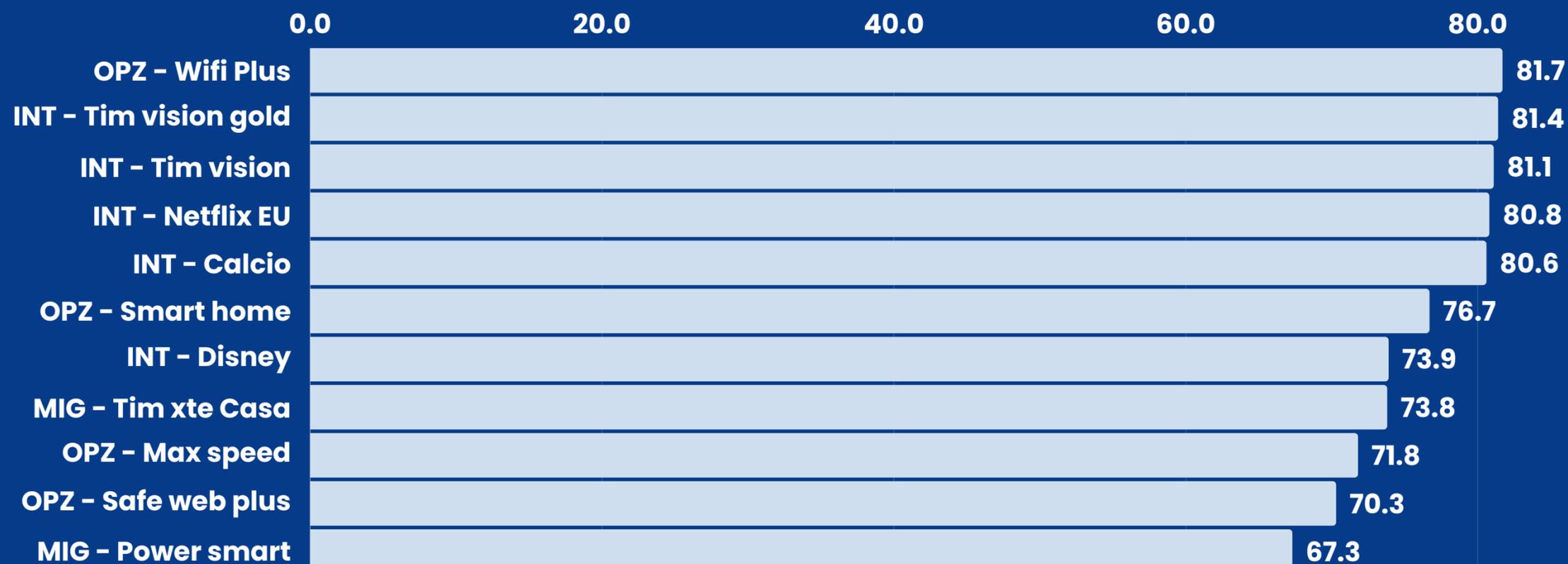


Fine Tuning

Ottimizzazione degli iperparametri tramite GridSearchCV e Optuna

Modelli specializzati

AUC - Training - Modello LightGBM





Metrica di selezione

NDCG@K

- *Misura l'efficacia di un sistema di raccomandazione.*
- *Valuta la qualità dei risultati ordinati in base alla rilevanza e alla posizione.*
- *Pondera la rilevanza dei risultati in base alla loro posizione.*
- *Confronta con l'ordinamento ideale per ottenere un punteggio tra 0 e 1.*

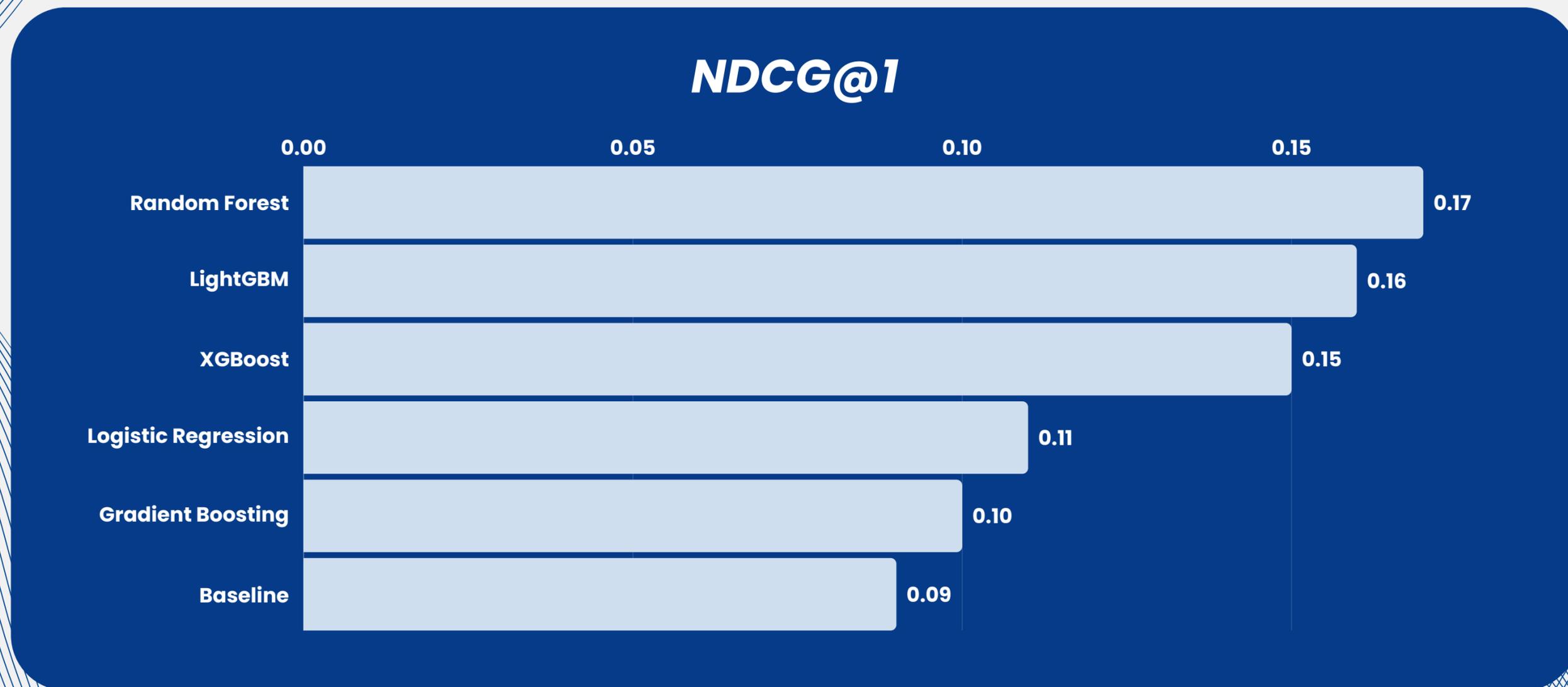
$$DCG@k = \sum_{i=1}^k \frac{rel_i}{\log_2(i+1)}$$

$$iDCG@k = \sum_{i=1}^k \frac{rel_i^{ideal}}{\log_2(i+1)}$$

$$NDCG@k = \frac{DCG@k}{iDCG@k}$$



Confronto tra modelli



Stima



Ranking probabilità

Stima delle 11 probabilità e creazione di un dataframe per il calcolo della matrice di valutazione.

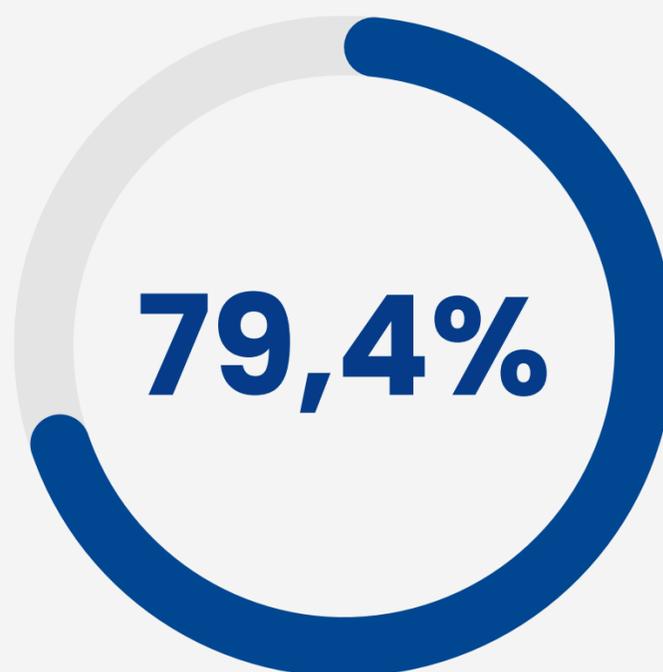


Risultati

In fase di addestramento del modello finale



AUC



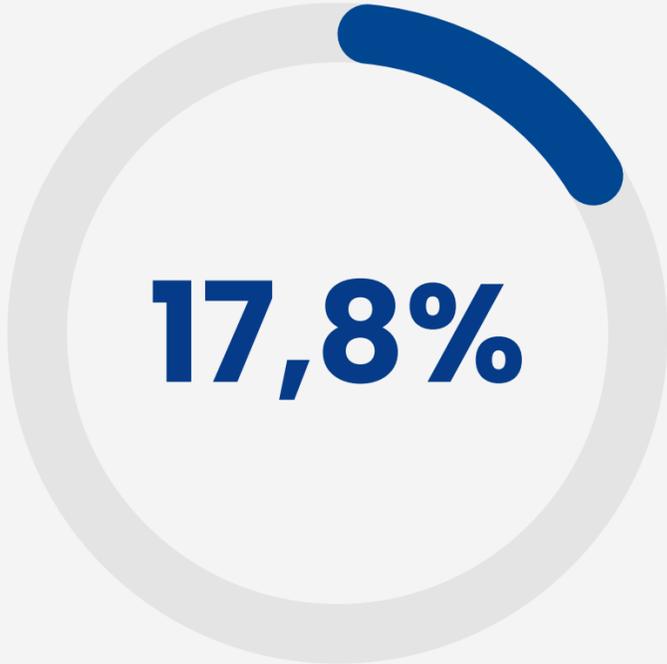
ACCURACY



NDCG@3

Risultati

In fase di test del modello finale

A donut chart representing the NDCG@1 metric. It features a light gray outer ring and a dark blue inner segment that covers approximately 17.8% of the circle. The percentage value "17,8%" is displayed in the center of the chart in a bold, dark blue font.

17,8%

NDCG@1

A donut chart representing the NDCG@3 metric. It features a light gray outer ring and a dark blue inner segment that covers approximately 29.9% of the circle. The percentage value "29,9%" is displayed in the center of the chart in a bold, dark blue font.

29,9%

NDCG@3

A donut chart representing the NDCG@5 metric. It features a light gray outer ring and a dark blue inner segment that covers approximately 39.4% of the circle. The percentage value "39.4%" is displayed in the center of the chart in a bold, dark blue font.

39.4%

NDCG@5

Prossimi passi



RACCOLTA DATI

Lavorare con una maggiore quantità di dati e migliorare l'interpretabilità.



MODELLISTICA

Esplorare e implementare ulteriori modelli.

**Grazie per
l'attenzione**

CESMA x  **TIM**

Team



**ABIRE
CHERIF**



**ELENA
BACIU**



**LEONARDO
ANDREASSI**



**ELENA
RICCIARELLI**



**LUCA
BATTAGLIONE**