

# Text Sumarization Service

Francesco Verona

August 23, 2025

## 1 Introduzione

Lo scopo dell' applicazione è di effettuare riassunti di testi passati come input. I riassunti generati potranno essere basati su approccio extractive/abstractive ed avere una certa lunghezza in base alle impostazioni decise dall' utente. L' obbiettivo di un buon riassunto è quello di estrarre le informazioni più importanti da un testo di partenza e di esporle in maniera efficace e comprensibile. L' importanza spesso può dipendere dall' ambito che tratta il testo di partenza e spesso solo esperti in tale settore sono capaci di individuarne gli aspetti essenziali.

### 1.1 Data Analysis

Il primo problema nell' affrontare questo compito è stata la ricerca di un dataset adeguato. Il dataset utilizzato per questa analisi è stato *cnn\_dailymails* *aperilsuovastoutilizzosunumerosiprogrammi*

### 1.2 Data cleaning

Per prima cosa abbiamo ripulito gli articoli dalle intestazioni giornalistiche come ad esempio "LONDRA (CNN) —" che non erano utili al nostro scopo poiché ci dobbiamo concentrare sul contenuto.