

Proposta di progetto - Big Data in Public and Social Services

Gruppo composto da: Cosimo Simone Farallo, Julia Bui Xuan, Michele Salvaterra

Nel settore sanitario, l'accesso alle informazioni sulle sedi in cui eseguire specifiche analisi mediche rappresenta una sfida, così come la comprensione del reale significato delle analisi stesse da parte dell'utente finale. In questo contesto, il presente progetto mira a creare un servizio di supporto in ambito sanitario per la regione Lombardia, al fine di fornire agli utenti informazioni dettagliate sulle analisi mediche e di aiutarli a individuare le strutture sanitarie più adeguate alle loro esigenze.

L'obiettivo principale del progetto consiste nell'implementazione di un chatbot. L'idea consiste nel fornire in input l'elenco delle analisi mediche da eseguire, attraverso la ricetta medica (formato pdf o png) o anche semplicemente attraverso una query di testo. L'output invece, dovrà contenere una lista di ambulatori, accompagnata da una serie di informazioni dettagliate, e la descrizione delle analisi stesse.

I dataset necessari per svolgere il lavoro sono i seguenti:

1. Prestazioni ambulatori:
<https://www.dati.lombardia.it/Sanit-/Prestazioni-Ambulatoriali/d4mg-9zw3>
2. Elenco strutture ambulatori:
https://www.dati.lombardia.it/Sanit-/Specialistica-Ambulatoriale_Elenco-Strutture-Lomba/kyrf-v8bb
3. Informazioni sulle analisi: Wikipedia o altra documentazione sulla quale eseguire il fine tuning

Elenchiamo di seguito gli strumenti che verranno attualmente impiegati: la tecnologia OCR (Optical Character Recognition) su file in formato PDF e JPG, i tool di Wikipedia e DuckDuckGo, i modelli di OpenAI per la chat, e il Tool Stuff per la conversazione. In seguito, in base alle necessità, verranno considerati altri strumenti. Infine, sarà in fase di valutazione, decidere se integrare o meno l'intero sistema su una piattaforma online dotata di un'interfaccia intuitiva.