

Descrizione del Progetto

Focus Generale:

Impiegare tecnologie di intelligenza artificiale per estrapolare e organizzare dati in maniera strutturata, creando una base di conoscenza specializzata per un sistema di Retrieval-Augmented Generation (RAG).

L'obiettivo del progetto:

L'obiettivo del progetto è sviluppare un sistema in grado di estrapolare informazioni da documenti complessi, come brochure contenenti testi, immagini, tabelle e figure, utilizzando tecnologie di intelligenza artificiale. Lo studio intende verificare se è possibile utilizzare modelli di AI alternativi al modello Vision di OpenAI, attraverso l'utilizzo di modelli open source o a costo inferiore, applicando tecniche di fine-tuning e ottimizzazione, con conseguente analisi delle performance e comparazione dei risultati.

Obiettivi Specifici:

- Determinare se il progetto può essere realizzato con strumenti gratuiti.
- Comparare le prestazioni ottenute con quelle del modello GPT-4 Vision.
- Sperimentare e implementare pratiche di ottimizzazione e fine-tuning per migliorare le performance dei modelli AI in questo contesto.

Outcome:

Il focus principale dell'esperienza sarà lo studio approfondito delle tecnologie di intelligenza artificiale alternative a GPT-4 Vision, comprendendo il loro funzionamento, le potenzialità e le eventuali limitazioni. Inoltre, si valuterà la possibilità di sviluppare un sistema entro contenuto comprensivo della tecnologia sopra descritta.

Consideriamo un sistema che, periodicamente, riceve nuovi documenti strutturati, come brochure contenenti testi, immagini, tabelle e figure. La base di conoscenza dell'assistente virtuale deve poter essere aggiornata con l'aggiunta di nuovi documenti tramite un'analisi automatizzata degli stessi, l'estrazione delle informazioni rilevanti e l'inserimento in un sistema di gestione della conoscenza (e.g. Chroma DB), alimentando così il database e consentendo un aggiornamento continuo e un training del modello AI.

Inoltre, si potrebbe prendere in considerazione la possibilità di affinamento della conoscenza pregressa tramite la scrematura delle informazioni diventate obsolete dopo un certo lasso di tempo, con un focus su come tali informazioni possano essere individuate, rimosse o corrette.