

Analisi dei requisiti

Gruppo	Alt+F4
Data	13 dicembre 2024
Versione	v0.1



Registro delle Modifiche

Versione	Data	Autore/i	Verificatore/i	Descrizione
v0.1	13 dicembre 2024	Pedro Leoni	Enrico Bianchi	Sezioni Descrizione del prodotto

Indice

1	Descrizione del prodotto	3
1.1	Obiettivi del prodotto	3
1.2	Funzioni del prodotto	3
1.3	Caratteristiche utente	3

1 Descrizione del prodotto

1.1 Obiettivi del prodotto

L'obiettivo del prodotto è permettere di poter valutare in modo automatizzato le prestazioni di un Large Language Model su un $dataset_G$ composto da un insieme di coppie domanda-risposta. Questo permette di ridurre/eliminare la verifica umana diminuendo i costi di test delle prestazioni su questi modelli. Il sistema prodotto può quindi essere usato per:

1. Confrontare diversi Large Language Model.
2. Confrontare diverse scelte implementative di un Large Language Model.

Il prodotto verrà quindi utilizzato per prendere decisioni sull'integrazione e l'implementazione di Large Language Model.

1.2 Funzioni del prodotto

Le funzioni principali del sistema sono:

- Gestione del dataset: inserimento, modifica e eliminazione delle coppie di domanda e risposta su cui eseguire il test.
- Persistenza del dataset: caricare e memorizzare il dataset.
- Sottoporre le domande al modello da testare: interfacciarsi al modello esterno tramite un $API\ REST_G$ secondo lo standard *OpenAPI 3.1G*.
- Eseguire il test: deve riuscire a valutare la correttezza/verosomiglianza delle risposte ricevute dal modello sottoposto al test in base alle risposte attese.
- Mostrare i risultati del test: i risultati devono essere mostrati usando una dashboard che mostri i risultati in modo sintetico e una seconda parte che permetta l'analisi completa dei singoli risultati.

1.3 Caratteristiche utente

Gli utenti target del sistema sono le figure professionali che possiedono conoscenze informatiche specifiche nell'ambito dell'intelligenza artificiale. In particolare:

1. programmatori e progettisti che devono prendere decisioni sull'integrazione di Large Language Model esistenti.
2. ricercatori e sviluppatori di Large Language Model che devono studiare il comportamento dei modelli.