



<CodeHex16>

unipd.codehex16@gmail.com

# Coordinamento Sprint 6

## Verbale di riunione interno

<b>Data</b>	06/01/2025
<b>Ora</b>	15:30-18:00
<b>Versione</b>	0.1.0

### Ordine del giorno

- Revisione Sprint 5
- Coordinamento per lo Sprint 6
- Condivisione progressi documenti RTB
- Scelta del modello LLM
- Schedulazione incontro con Ergon

### Presenze

Matteo Bazzan	Relatore	2h30m
Luca Ribon	Verificatore	2h30m
Filippo Sabbadin	Verificatore	2h
Francesco Fragonas	Verificatore	2h
Yi Hao Zhuo	Verificatore	2h30m
Luca Rossi	Verificatore	1h

## Indice

<b>1. Revisione Sprint 5 .....</b>	<b>1</b>
<b>2. Scelta del modello LLM .....</b>	<b>1</b>
<b>3. Schedulazione incontro con l'azienda Ergon .....</b>	<b>2</b>
<b>4. Coordinamento Sprint 6 .....</b>	<b>3</b>
<b>5. Azioni future .....</b>	<b>3</b>
<b>6. Fonti .....</b>	<b>3</b>

## 1. Revisione Sprint 5

Durante la riunione è stato analizzato il lavoro svolto nello **Sprint 5**, in particolare sono stati condivisi i progressi e risolto alcuni dubbi sui seguenti documenti :

- [Analisi dei Requisiti\\*](#)
- [Norme di Progetto\\*](#)
- Piano di Qualifica
- [Piano di Progetto\\*](#)

Inoltre è stato condiviso lo studio sui vari [LLM\\*](#) in commercio.

## 2. Scelta del modello LLM

Durante lo Sprint 5 è stata condotta un'analisi dei modelli LLM disponibili si è deciso di utilizzare [GPT-4o mini\\*](#) per i primi test. In caso di performance non in linea con i requisiti, si valuterà un modello più performante tra le offerte di [OpenAI\\*](#).

È stato scelto GPT-4o mini perché:

- **Comprensione:** Fornisce buone capacità di comprensione del linguaggio naturale, utili allo sviluppo di un [chatbot\\*](#).
- **Costo:** Presenta un costo molto contenuto:
  - \$0.150/1M [input tokens\\*](#)
  - \$0.600/1M [output tokens\\*](#)
- **Token\*:** Ha una finestra di contesto ([context window\\*](#)) di 128,000 tokens, abbastanza per avere contesto sufficiente per rispondere correttamente alle domande poste al chatbot
- **Velocità:** Fornisce un [output speed\\*](#) di 125.1 tokens al secondo, più veloce della media degli altri modelli

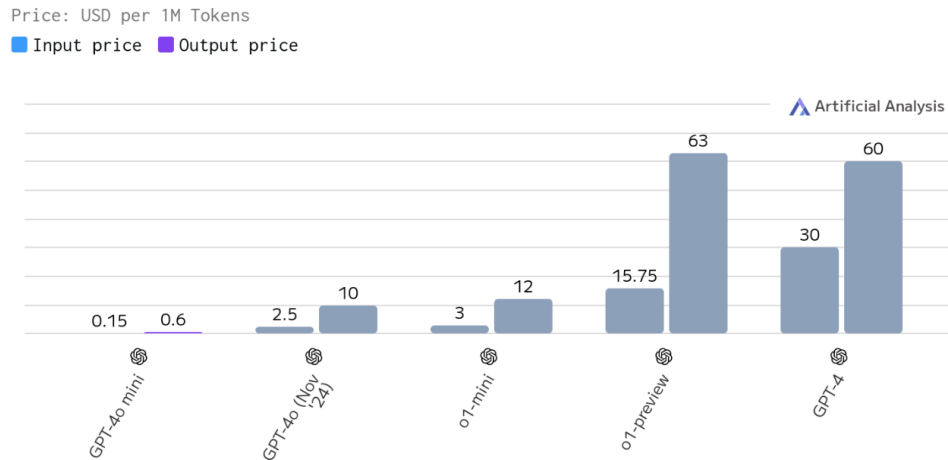


Figura 1: Confronto tra i costi di input e output per un 1M di token di modelli OpenAI

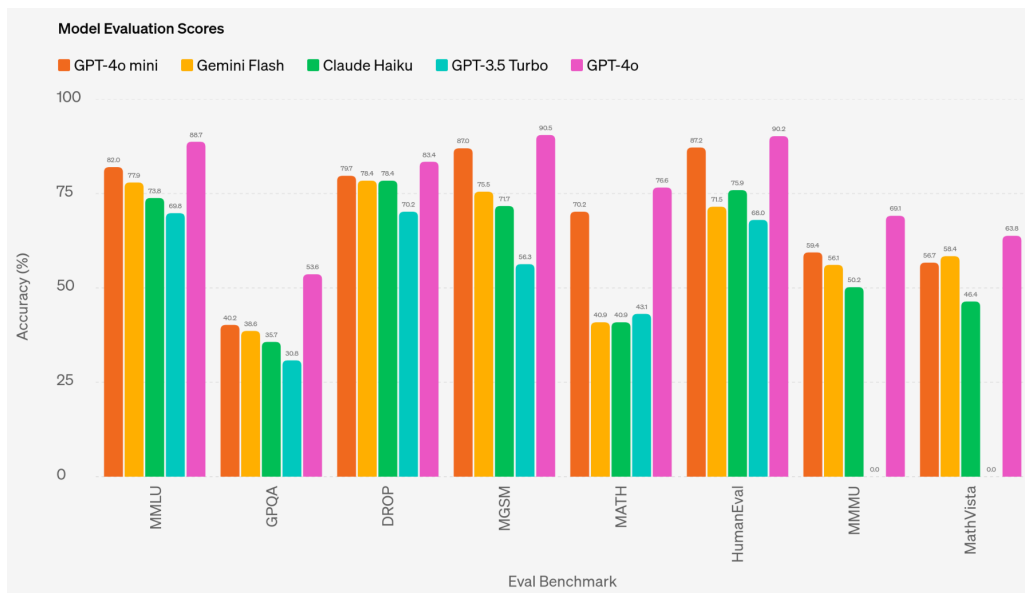


Figura 2: Confronto delle prestazioni di GPT-4o mini con altri modelli concorrenti nelle categoria con cui viene valutata l'intelligenza degli LLM

### 3. Schedulazione incontro con l'azienda Ergon

È necessario organizzare un incontro con l'azienda Ergon per:

- condividere gli [use case\\*](#) finora individuati
- illustrare la scelta del modello LLM e richiedere le credenziali di accesso alla piattaforma di OpenAI, così da avviare i primi test tramite le [API\\*](#).

## 4. Coordinamento Sprint 6

- I **nuovi ruoli** per lo Sprint 6 sono:

Membro	Ruolo Sprint 5	Ruolo Sprint 6
Matteo Bazzan	Progettista	Verificatore
Francesco Fragonas	Responsabile	Progettista
Gabriele Magnelli	Amministratore	Responsabile
Luca Ribon	Analista	Amministratore
Luca Rossi	Analista	Verificatore
Filippo Sabbadin	Verificatore	Analista
Yi Hao Zhuo	Verificatore	Verificatore

## 5. Azioni future

- Incontro con l'azienda Ergon
- **Aggiornamento Glossario**
- **Sistemare paginazione template typst**
- **Creazione diagrammi UML per gli Use Case**
- **Continuare stesura piano di qualifica**
- **Continuare stesura norme di progetto**
- **Aggiornare piano di progetto con sprint 5**
- **Iniziare PoC per LLM e Backend**: Inizio dello studio e sviluppo della parte del PoC riguardante il backend e gli LLM

## 6. Fonti

- GPT-4o mini - OpenAI
- GPT-4o mini - Artificial Analysis