

Dipartimento di Matematica e Informatica

Corso di Laurea Triennale in Informatica

*Lorenzo Maria Alberto Paoria*

Titolo

Relazione Progetto Finale

Relatore: Emiliano Tramontana

Chiar.mo Prof. ---

Correlatori:

Gabriella Verga

Alessandro Midolo

Anno Accademico 2023 – 2024

# Indice

[Indice 3](#_Toc175745904)

[Introduzione 4](#_Toc175745905)

[Capitolo 1 - Metodologia 6](#_Toc175745906)

[1.1 - Titolo. 6](#_Toc175745907)

[1.1.1 - Titolo. 6](#_Toc175745908)

[1.1.2 - Titolo. 6](#_Toc175745909)

[1.2 - Titolo. 6](#_Toc175745910)

[Capitolo 2 - Risultati. 7](#_Toc175745911)

[2.1 - Post con codice. 7](#_Toc175745912)

[2.2 - Post senza codice. 10](#_Toc175745913)

[Capitolo 3 - Conclusioni 14](#_Toc175745914)

[Riferimenti 15](#_Toc175745915)

# Introduzione

Questa tesi si propone di confrontare le risposte fornite dagli utenti di StackOverflow con quelle generate da sistemi di intelligenza artificiale avanzati, come ChatGPT, al fine di valutare la capacità di tali tecnologie nel fornire soluzioni tecniche di alta qualità. L'obiettivo principale è analizzare la pertinenza, l'accuratezza e l'efficacia delle risposte prodotte dall'AI rispetto a quelle fornite da esperti umani, per comprendere i limiti e le potenzialità delle moderne tecnologie di linguaggio naturale nel supportare la risoluzione di problemi complessi nel campo dell'informatica.

Nel seguito, verranno illustrati i contenuti dei capitoli della tesi.

* Font per i Titoli dei capitoli: Cambria 16 (Stile Titolo 1).
* Font per i Titoli dei sottocapitoli: Cambria 14 (Stile Titolo 2),
* Si prosegue con Titolo 3, ecc…
* L’utilizzo degli Stili “Titolo” permette di generare l’Indice automaticamente.
* Ogni capitolo deve terminare con una Interruzione di pagina.
* Font per il testo semplice: Calibri 12 (Stile Normale).
* Interlinea 1,5.
* Margine Sinistro 4 cm. Destro, Superiore e Inferiore 3 cm.
* Spaziatura paragrafo 12 pt dopo.
* Figure: Salvo particolari motivazioni, si suggerisce di inserire le figure “In linea con il testo” e centrate. Inoltre, bisogna aggiungere la didascalia con l’opportuno strumento in modo da numerarle automaticamente e poterle citare nel testo tramite Riferimenti incrociati. Vedere esempio.
* Bibliografia: la bibliografia va inserita come ultima sezione, dopo le conclusioni. Tutte le fonti vanno inserite con lo strumento Gestione fonti e citati con lo strumento Riferimenti. La bibliografia potrà così essere generata automaticamente. Vedere esempio.
* Si possono usare note a piè di pagina.
* Si possono inserire ringraziamenti.
* Si può essere flessibili rispetto alle regole sopraccitate.

# Metodologia

Questo è il capitolo 1. Questa è una citazione [1]

## Titolo.

Questo è un sottocapitolo del capitolo 1

### Titolo.

Questo è un sotto-sotto capitolo del capitolo 1.

### Titolo.

Questo è un altro sotto sotto capitolo del capitolo 1

## Titolo.

Questo è un altro sottocapitolo del capitolo 1.

# Risultati.

Una volta analizzati tutti i post e completate le richieste all’IA, abbiamo prodotto i risultati per i diversi tipi di catalogazione, principalmente suddivisi in post con e senza codice.

I risultati sono stati illustrati attraverso grafici ed esempi per facilitarne l'interpretazione.

## Post con codice.

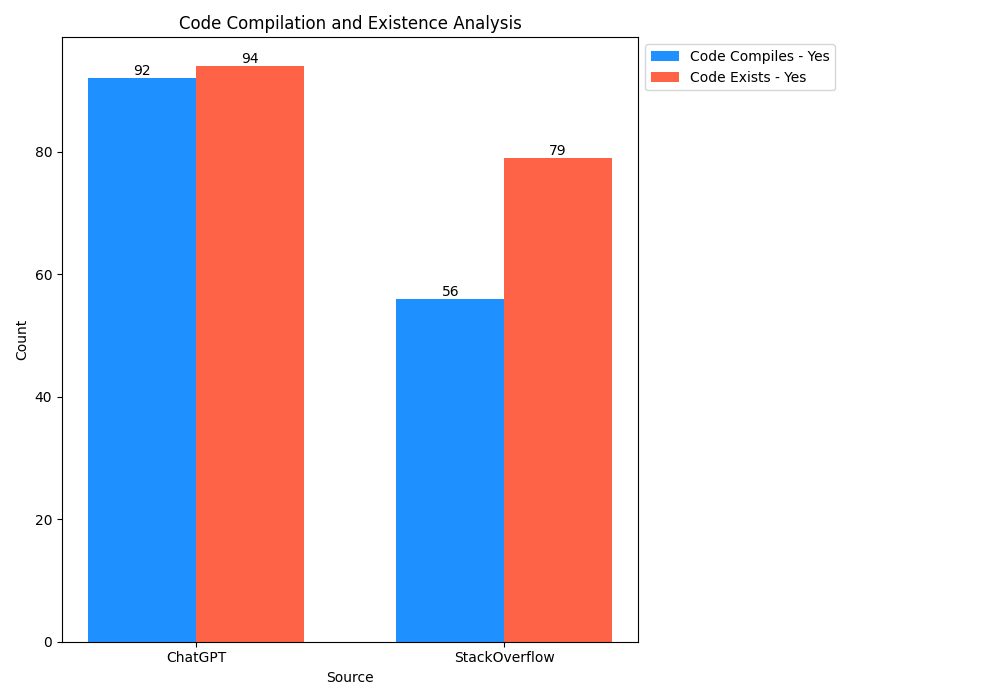
Per Per i post contenenti codice presenti nelle risposte, è stata valutata l'eseguibilità del codice o, nel caso di frammenti decontestualizzati, la correttezza semantica rispetto alla domanda. L'analisi è stata condotta su un dataset di 218 domande. Dai risultati è emerso che i post contenenti codice, già selezionati durante la creazione del dataset, erano:

* ChatGpt: 94 instanze
* StackOverflow: 79 instanze

Di queste:

* ChatGpt: 92 instanze compilano con successo
* StackOverflow: 56 instanze compilano con successo

Tabella 1- Grafico della situazione



Per fornire un esempio testuale dei post analizzati e dei risultati ottenuti, presenteremo un esempio di un post contenente codice e la relativa analisi.

### Esempio post con codice

Domanda:

"ID":"937", "Question": "I know that they are implemented extremely unsafely in C/C++. Can't they be implemented in a safer way? Are the disadvantages of macros really bad enough to outweigh the massive power they provide? "

#define TWO 1+1

#define min(x, y) (((x) < (y)) ? (x) : (y))

Risposta StackOverflow:

Anche se questo codice è compilabile, può comportarsi in modo inaspettato quando viene utilizzato in espressioni. Pertanto, quando l'IA viene interpellata sulla sua correttezza semantica, lo considera non semanticamente corretto rispetto a quanto richiesto dalla domanda.

Risposta ChatGpt:

inline int square(int x) {

return x \* x;

}

ChatGPT sconsiglia l'uso delle macro in C++ e raccomanda invece di adottare un approccio basato su funzioni inline. Questo codice si compila correttamente ed è in linea con la domanda, poiché propone una soluzione alternativa ma più sicura rispetto a una che potrebbe comportare dei rischi.

## Post senza codice.

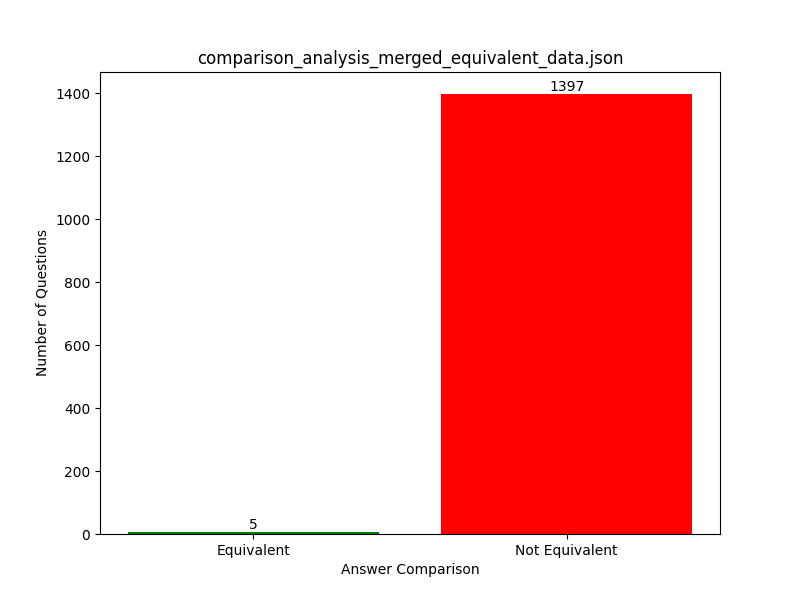
Per i post che non contengono codice, sono state raggruppate diverse categorie, tra cui:

* Post con un numero di caratteri superiore a limitChar
* Post con un numero di caratteri inferiore a limitChar
* Post contenenti un termine specifico

Sono state analizzate le equivalenze tra le migliori risposte di StackOverflow e le risposte di ChatGPT per tutti questi tipi di post. L'analisi è stata condotta su un dataset di 1402 domande e dai risultati è emerso che:

* 1397 domande non avevano risposte equivalenti
* 5 domande avevano risposte equivalenti

Tabella 2- Grafico della situazione



### Esempio post senza codice

Per illustrare il tipo di analisi effettuata e i risultati ottenuti, presenteremo un esempio che include una domanda con due risposte equivalenti.

"Question": "Where I work, employees use a third-party desktop program for their clients. This program saves data to a flat file. My colleague wants to write a Java program that uploads that flat file to a remote server, opens the desktop program when the flat file is downloaded from a Web site, and checks if the desktop program is running or not by looking at the Windows processes. He keeps calling this helper/utility program a \"wrapper.\" But it doesn't wrap anything! I tried to clear it up with him, but he said, \"Well, I call it a wrapper.\" He now has everyone in the company calling it a \"wrapper.\" What would you call it? I say that it's a helper program or utility program. "

Domanda:

"ChatGpt answer": "You're correct in suggesting that the term \"wrapper\" may not be the most accurate description for the program your colleague wants to create. Typically, a \"wrapper\" refers to software that encapsulates or interfaces with another underlying program or library,

Risposta di ChatGpt:

"Stack Overflow best answer": "I don't think there's an official definition!  In my mind, a wrapper program is a program that, in general, starts before the wrapped program starts, and exits no earlier than after the wrapped program exits.  A wrapper program would probably also manage the lifetime of the wrapped program (start it for you, stop it for you.)

Risposta di StackOverflow:

often providing a simpler or enhanced interface. In the case you've described, the Java program doesn't seem to be wrapping another piece of software but rather performing a variety of tasks\u2014uploading files, launching an application, and checking processes.  Given its functionality, you could accurately describe it as a \*\*\"utility program\"\*\* or a \*\*\"helper program.\"\*\* These terms more clearly convey the program\u2019s role as a tool that assists users in performing specific tasks instead of encapsulating or modifying another program.  You might also consider terms like \*\*\"automation script\"\*\* or \*\*\"file management tool,\"\*\* depending on how your colleague envisions the program functioning and its primary purpose. Whatever you decide to call it, highlighting its utility may help clarify its purpose within the company."

La risposta di ChatGPT è più dettagliata e orientata a trovare una terminologia più precisa per descrivere il programma specifico, mentre la risposta di Stack Overflow è più generale e focalizzata sull'aspetto funzionale del software. Le due risposte non sono equivalenti poiché adottano approcci e prospettive differenti.

I recommend against arguing too much over the names of things, though.  The important thing is to design software that works well.  Sometimes belaboring terminology can aid in this goal; sometimes it is a hobgoblin. "

# Conclusioni

# Riferimenti

|  |
| --- |
|  |