

PROGRAMMA DEL CORSO DI LABORATORIO PROMPT ENGINEERING E LLM

SETTORE SCIENTIFICO

ING-INF/05 (IINF-05/A)

CFU

6

OBIETTIVI FORMATIVI PER IL RAGGIUNGIMENTO DEI RISULTATI DI APPRENDIMENTO PREVISTI NELLA SCHEDA SUA

L'obiettivo di questo corso è fornire la teoria, le tecniche, i suggerimenti e i trucchi necessari per padroneggiare l'ingegneria del prompt e costruire applicazioni basate su Large Language Models di qualità.

Obbiettivi formativi:

1. Fornire una comprensione fondamentale degli LLM, del loro funzionamento interno e della loro funzionalità come motori di completamento del testo.
2. Presentare un approccio di alto livello allo sviluppo di applicazioni basate su Large Language Models.
3. Introdurre le tecniche fondamentali per l'ingegnerizzazione dei prompt; in particolare, come reperire le informazioni sul contesto, classificarne l'importanza per il compito da svolgere, comporre il prompt in modo che risulti in completamenti di alta qualità e che fornisca la risposta desiderata.
4. Presentare tecniche avanzate per assemblare loop, pipeline e flussi di lavoro di inferenza basati su LLM, per creare agenti conversazionali e flussi di lavoro guidati da LLM.
5. Presentare le tecniche e le metriche per la valutazione degli LLM.

RISULTATI DI APPRENDIMENTO ATTESI

/**/

Conoscenza e capacità di comprensione

Comprendere i principi di funzionamento degli LLM;

Conoscere i principi di sviluppo di applicazioni basate su LLM;

Conoscere le tecniche fondamentali di composizione e ingegnerizzazione di prompt;

Conoscere le principali tecniche per assemblare flussi di lavoro guidati da LLM;

Conoscere i principali criteri e le principali metriche per la valutazione degli LLM.

Capacità di applicare conoscenza e comprensione

Formulare soluzioni strategiche ed operative riguardanti:

La progettazione di applicazioni basate su LLM;

L'ingegnerizzazione e la composizione di prompt per sistemi LLM;

La progettazione di flussi di lavoro e applicazioni basate su LLM.

Autonomia di giudizio

Sviluppare un approccio critico all'uso di LLM, in particolare:

La formulazione di giudizi su modelli, linguaggi e sistemi da adottare in base al dominio applicativo e al contesto di utilizzo;

Le strategie e le tecniche da adottare nella composizione di prompt per LLM.

Abilità comunicative

Acquisire specifiche competenze comunicative riguardo:

L'esposizione dei principi di funzionamento LLM in modo chiaro e articolato, a interlocutori specialisti e non;

La partecipazione a discussioni sulle tecnologie di intelligenza artificiale, in particolare quella generativa, e sulle tecnologie LLM.

Capacità di apprendimento

Sviluppare la capacità di analizzare concetti e tecniche fondamentali, integrando le differenti conoscenze apprese, per:

L'approfondimento dei temi trattati in modo autonomo, per aggiornare le conoscenze su tecnologie e metodologie emergenti nel campo dell'intelligenza artificiale;

L'analisi di nuovi problemi e bisogni, per ottimizzare la progettazione e l'utilizzo di applicazioni basate su intelligenza artificiale generativa e LLM.

PREREQUISITI

/**/

Non sono richieste conoscenze preliminari.

PROGRAMMA DIDATTICO: ELENCO VIDEOLEZIONI/MODULI

/**/

Di seguito l'elenco delle 30 videolezioni

1. I principi fondamentali del Prompt Engineering
2. Il secondo principio - La strutturazione del formato
3. Il terzo principio - Fornire degli esempi
4. Quarto principio - La valutazione della qualità

5. Quinto principio - La divisione del lavoro
6. Introduzione agli LLM
7. Le architetture Transformer
8. Il funzionamento di un LLM
9. Gli LLM per la generazione del testo
10. Il Prompt Engineering applicato a ChatGPT
11. Il chunking nei LLM
12. Dagli Instruct ai Chat Tools
13. Progettare applicazioni LLM
14. Gestione del contenuto nei sistemi basati su LLM
15. Progettare Prompt Efficaci
16. Function Calling
17. Meta Prompting
18. LangChain per il Prompt Engineering
19. Funzionalità avanzate del framework LangChain
20. Creazione e utilizzo di catene di LLM Chains e Agents
21. Lavorare sui testi con LangChain
22. RAG con LangChain
23. Utilizzo della Memoria
24. Integrazione tra Prompt Engineering e Diffusion Models
25. Modelli Open Source
26. Workflow con LLM
27. Generazione di immagini con Midjourney
28. Tecniche avanzate per generazione di immagini con Stable Diffusion
29. La progettazione di un'applicazione AI-driven
30. L'IA generativa per la progettazione industriale: intervista

ATTIVITÀ DIDATTICA EROGATIVA (DE)

//**

Le attività di didattica erogativa consistono, per ciascun CFU, nell'erogazione di 5 videolezioni corredate di testo e questionario finale.

- Il format di ciascuna videolezione prevede il video registrato del docente che illustra le slide costruite con parole chiave e schemi esemplificativi.
- Il materiale testuale allegato a ciascuna lezione corrisponde a una dispensa (PDF) composta da almeno 10 pagine con le informazioni necessarie per la corretta e proficua acquisizione dei contenuti trattati durante la lezione.

Attività di autoverifica degli apprendimenti prevista al termine di ogni singola videolezione consiste in un questionario costituito da 10 domande, a risposta multipla.

TESTI CONSIGLIATI

/**/

Gli studenti che intendono approfondire le tematiche del corso, integrando le dispense e i materiali forniti dal docente, possono consultare i seguenti volumi:

- J. Berryman, A. Ziegler, "Prompt Engineering for LLMs", O'Reilly Media Inc., ISBN 9781098156152
- J. Phoenix, M. Taylor "Prompt Engineering for Generative AI - Future-Proof Inputs for Reliable AI Outputs at Scale" O'Reilly Media Inc. ISBN 9781098153434

Si specifica che il testo consigliato è solo per approfondimento volontario, e non sarà oggetto specifico di esame. Il modello didattico è basato sullo studio delle dispense fornite del docente

MODALITÀ DI VERIFICA DELL'APPRENDIMENTO

/**/

L'esame può essere sostenuto sia in forma scritta che in forma orale.

Gli appelli orali sono previsti nella sola sede centrale. L'esame orale consiste in un colloquio con la Commissione sui contenuti del corso.

L'esame scritto consiste nello svolgimento di un test con 30 domande. Per ogni domanda lo studente deve scegliere una di 4 possibili risposte. Solo una risposta è corretta.

Sia le domande orali che le domande scritte sono formulate per valutare il grado di comprensione delle nozioni teoriche e la capacità di ragionare utilizzando tali nozioni. Le domande sulle nozioni teoriche consentiranno di valutare il livello di comprensione. Le domande che richiedono l'elaborazione di un ragionamento consentiranno di valutare il livello di competenza e l'autonomia di giudizio maturati dallo studente.

Le abilità di comunicazione e la capacità di apprendimento saranno valutate anche attraverso le interazioni dirette tra docente e studente che avranno luogo durante la fruizione del corso (videoconferenze ed elaborati proposti dal docente).

RECAPITI

/**/

- Prof. Andrea Generosi: andrea.generosi@unipegaso.it
- Prof. Salvatore Barone: salvatore.barone@unipegaso.it

OBBLIGO DI FREQUENZA

/**/

Obbligatoria online. Ai corsisti viene richiesto di visionare almeno l'80% delle videolezioni presenti in piattaforma.

AGENDA

In Informazioni Appelli, nella home del corso, per ogni anno accademico vengono fornite le date degli appelli.

ATTIVITÀ DIDATTICA INTERATTIVA (DI)

/**/

Le attività di Didattica interattiva consistono, per ciascun CFU, in due ore dedicate a una o più tra le seguenti tipologie di attività:

- Redazione di un elaborato
- Partecipazione a una web conference
- Partecipazione al forum tematico
- Lettura area FAQ
- Svolgimento delle prove in itinere con feedback

Per gli aggiornamenti, la calendarizzazione delle attività e le modalità di partecipazione si rimanda alla piattaforma didattica dell'insegnamento.