SCHEDE PROFILO ALLEGATO AL BANDO CdR 1/2025

Posizioni ammesse a finanziamento elenco Centro - Nord

CUP	*J93C25000440001*
CODICE PROFILO	**PROFILO A)2**
NUMERO DI POSIZIONE DEL PROGETTO DI RICERCA	2
POSIZIONI MESSE A SELEZIONE	1
DURATA CONTRATTO DI RICERCA	24 MESI
SEDE OPERATIVA	Area Science Park - Campus di Padriciano Padriciano, 99 34149- Trieste
SEDE OPERATIVA DIPARTIMENTO	RIT- Istituto Ricerca per l'Innovazione Tecnologica Ufficio Laboratorio di Data Engineering - LADE Struttura Ricerca e Innovazione- Area Science Park
SETTORE SCIENTIFICO DISCIPLINARE	IINF-01/A; IINF-05/A
PROFILO DEL CONTRATTISTA DA ASSUMERE	Il ricercatore dovrà avere esperienza pregressa nell'analisi di modelli di reti neurali profonde, con particolare attenzione allo studio delle loro rappresentazioni interne attraverso metodi geometrici e tecniche di interpretabilità. Saranno valutate positivamente competenze nell'analisi degli spazi latenti dei modelli, nell'applicazione di approcci interpretabili per lo studio dei circuiti computazionali e nella caratterizzazione dell'interazione tra modalità differenti (immagini, testo, video). È richiesta inoltre una buona conoscenza di strumenti e framework per l'addestramento e l'analisi di modelli di reti neurali profonde.
DESCRIZIONE DEL PROGETTO DI RICERCA	La ricerca in Intelligenza Artificiale (IA) sta producendo modelli sempre più complessi e integrati, in grado di elaborare simultaneamente più modalità di dati, come immagini, testo e video. Tali modelli multimodali dimostrano eccellenti abilità di generalizzazione, ma il loro funzionamento interno rimane in gran parte opaco, rendendo difficile comprendere come sviluppino capacità di ragionamento su informazioni eterogenee. Questo progetto sviluppa nuove tecniche di interpretabilità per analizzare le proprietà geometriche delle rappresentazioni apprese negli spazi latenti dei modelli multimodali, affiancandole a metodi ispirati alla mechanistic interpretability per identificare e studiare circuiti computazionali emergenti. Un aspetto cruciale, ma complesso, è comprendere l'interazione tra le diverse modalità, per ottimizzare il ragionamento multimodale, migliorando performance, efficienza energetica e sicurezza e rendendo l'IA più trasparente e affidabile.
RESPONSABILE DELLA RICERCA	Alberto Cazzaniga