



UNIVERSITÀ
DI SIENA
1240

DIPARTIMENTO
INGEGNERIA DELL'INFORMAZIONE
E SCIENZE MATEMATICHE

SCHEDA ATTIVITÀ

INCARICO DI LAVORO AUTONOMO

<i>Titolo del progetto</i>	“Sviluppo software per integrazione di Language Model e adattamento progressivo in ambiente distribuito e decentralizzato”
<i>Soggetto proponente</i>	Prof. Marco Gori, Dipartimento di Ingegneria dell'Informazione e Scienze Matematiche
<i>Obiettivi e finalità</i>	Sviluppo e testing di soluzione software per integrazione di modelli di linguaggio (principalmente Small Language Models) in ambiente distribuito.
<i>Responsabili delle attività di progetto</i>	Prof. Stefano Melacci, Dipartimento di Ingegneria dell'Informazione e Scienze Matematiche
<i>Durata dell'incarico</i>	4 mesi



<p>Requisiti/competenze richieste</p>	<p>Laurea magistrale (o specialistica o Laurea vecchio ordinamento) nell'area dell'Ingegneria dell'Informazione; possesso del titolo di dottore di ricerca su temi ad AI o iscrizione al terzo anno di un programma di dottorato con attività orientate a tematiche di AI; adeguato curriculum professionale idoneo allo svolgimento di attività di ricerca con modelli di generative Artificial Intelligence (AI) ed esperienza con architetture per AI distribuita/decentralizzata.</p>
<p>Descrizione dell'attività complessiva di progetto</p>	<p>L'attività oggetto del presente incarico si inserisce nel contesto dell'Intelligenza Artificiale distribuita e decentralizzata, con un focus specifico sull'integrazione di Small Language Model (SLM) in architetture decentralizzate. L'obiettivo è superare i limiti dei modelli centralizzati in termini di privacy, latenza e consumo di risorse, implementando meccanismi di adattamento progressivo del modello basati sui dati locali dei nodi della rete.</p> <p>L'attività mira a realizzare un framework software capace di gestire il ciclo di vita di modelli linguistici leggeri, garantendo l'interoperabilità tra nodi eterogenei e l'aggiornamento continuo delle conoscenze senza la necessità di un repository di dati centrale.</p> <p>Articolazione dell'attività in fasi</p> <p>L'attività, della durata complessiva di 4 mesi, sarà suddivisa nelle seguenti tre fasi funzionali:</p> <p>Fase 1: Analisi dei requisiti e progettazione dell'architettura decentralizzata [1 mese]</p> <p>In questa fase iniziale, il collaboratore si occuperà della definizione dei protocolli di comunicazione tra i nodi della rete distribuita.</p>



	<ul style="list-style-type: none">• Selezione degli SLM: Individuazione dei modelli di linguaggio ottimizzati per l'esecuzione su dispositivi con risorse computazionali limitate.• Definizione del protocollo di orchestrazione: Progettazione della logica di distribuzione del carico e della gestione dei messaggi in ambiente decentralizzato, tramite piattaforma stabilita dal responsabile scientifico all'avvio delle attività.• Analisi della sicurezza: Definizione dei criteri di integrità per lo scambio dei pesi o dei gradienti tra i nodi. <p>Fase 2: Sviluppo del framework di integrazione e adattamento progressivo [2 mesi]</p> <p>La seconda fase riguarderà l'implementazione core del software e dei meccanismi di apprendimento.</p> <ul style="list-style-type: none">• Sviluppo software: Realizzazione dei moduli per l'integrazione dei modelli di linguaggio all'interno dell'ambiente distribuito selezionato.• Implementazione dell'adattamento progressivo: Sviluppo di algoritmi di Fine-Tuning efficiente (es. LoRA o tecniche di apprendimento federato) che permettano al modello di adattarsi localmente ai dati specifici di ogni nodo.• Gestione della decentralizzazione: Implementazione di meccanismi di consenso o aggregazione per la sincronizzazione della conoscenza tra i nodi senza compromettere la privacy dei dati originali. <p>Fase 3: Testing, validazione e ottimizzazione delle</p>
--	---



UNIVERSITÀ
DI SIENA
1240

DIPARTIMENTO
INGEGNERIA DELL'INFORMAZIONE
E SCIENZE MATEMATICHE

performance [1 mese]

L'ultima fase sarà dedicata alla verifica sperimentale della soluzione sviluppata.

- Valutazione del comportamento del sistema in rete distribuita eterogenea per misurare latenza e accuratezza dell'adattamento.
- Benchmarking: Valutazione delle prestazioni del sistema decentralizzato.
- Documentazione tecnica: Redazione del report finale sulle attività svolte e sui risultati ottenuti e consegna del software prodotto.

In entrambe le fasi:

L'utilizzo di ogni tipo di stack software o dipendenza deve essere preventivamente concordato con il responsabile dell'attività.

Il Proponente

Il Responsabile del Progetto