

Progetto di ricerca e piano delle attività

1. Descrizione del progetto di ricerca

Il progetto si colloca nell'ambito di IT4LIA AI Factory, infrastruttura nazionale per lo sviluppo e la diffusione dell'intelligenza artificiale in Italia. L'attività di ricerca riguarda la progettazione e l'implementazione di ambienti di testing per verificare compliance normativa e fairness etica dei modelli di AI integrati nella piattaforma. Obiettivo del progetto è definire un framework sperimentale che consenta di valutare in modo sistematico correttezza, non discriminazione e conformità dei sistemi di AI rispetto ai principi di equità, trasparenza e accountability, in linea con AI Act, GDPR e normative settoriali. Il ricercatore contribuirà alla definizione di protocolli e metriche di fairness, all'analisi dei trade-off tra accuratezza e equità e allo sviluppo di tecniche di mitigazione dei bias. I risultati confluiranno in un servizio affidabile e responsabile per la validazione etica, legale e tecnica dei sistemi di intelligenza artificiale di IT4LIA.

2. Piano delle attività del/della contrattista di ricerca

Fase 1 – Analisi requisiti e stato dell'arte

- Analisi della letteratura scientifica su fairness, bias mitigation e auditing dei sistemi di AI.
- Studio del quadro regolatorio (AI Act, GDPR, linee guida EDPB e standard emergenti) e delle policy interne di IT4LIA.
- Raccolta e formalizzazione dei requisiti funzionali e non funzionali dell'ambiente di testing, in collaborazione con i partner tecnici e legali della piattaforma.

Fase 2 – Progettazione del framework di testing

- Definizione dell'architettura logica dell'ambiente di sperimentazione (moduli, interfacce, flussi di dati).
- Progettazione di protocolli di valutazione per diverse tipologie di modelli (modelli predittivi, generativi, sistemi ibridi).
- Selezione e/o definizione di metriche di fairness, trasparenza e accountability, comprensive di indicatori di trade-off con accuratezza e performance.

Fase 3 – Sviluppo e integrazione degli strumenti

- Implementazione dei moduli software per il calcolo delle metriche di fairness e per l'analisi dei trade-off con le metriche di accuratezza.
- Sviluppo di componenti per l'applicazione e il confronto di diverse tecniche di mitigazione dei bias (pre-processing, in-processing, post-processing).
- Integrazione dell'ambiente di testing con la piattaforma IT4LIA AI Factory (API, pipeline di MLOps, servizi di logging e monitoraggio).