

Titolo Assegno Fairness and Bias in AI

Progetto di Ricerca

L'avvento dell'intelligenza artificiale ha portato notevoli progressi, ma la presenza di pregiudizi e disuguaglianze nei sistemi AI è diventata una preoccupazione critica. Questo progetto di ricerca mira a esplorare approcci innovativi per promuovere la giustizia e ridurre i bias nei sistemi di intelligenza artificiale.

Obiettivi Principali:

- **Analisi Critica:** Condurre una revisione approfondita delle attuali pratiche e metodologie di sviluppo di algoritmi AI per identificare le fonti di unfairness e bias.
- **Sviluppo di Metriche Fair:** Definire e implementare metriche robuste per valutare la fairness in diverse applicazioni AI, considerando le variabili demografiche e culturali.
- **Tecniche di Mitigazione:** Investigare tecniche avanzate di mitigazione dei bias, inclusi approcci basati su machine learning, regolamenti algoritmici e miglioramento delle dataset.
- **Validazione Pratica:** Applicare le soluzioni sviluppate a casi studio reali per valutare l'efficacia delle nuove metriche e delle tecniche di mitigazione.

Metodologia:

La ricerca utilizzerà un approccio multidisciplinare, combinando analisi etiche, metodologie di machine learning, e test empirici su dataset specifici. L'obiettivo è creare un quadro comprensivo e pratico per affrontare la fairness e il bias nei sistemi AI.

Risultati Attesi:

- **Framework Innovativo:** Sviluppo di un framework integrato per la valutazione e il miglioramento della fairness nei sistemi AI.
- **Metriche Standard:** Introduzione di nuove metriche standardizzate per valutare la fairness, contribuendo a norme e best practices nel settore.
- **Sistemi AI Fairer:** Implementazione di soluzioni pratiche per mitigare il bias in applicazioni reali, dimostrando l'efficacia del nostro approccio.

Piano di Attività

Analisi Preliminare (Mesi 1-3):

- Revisione della letteratura sull'attuale stato di fairness e bias in sistemi AI.
- Identificazione delle fonti di unfairness e bias nei contesti applicativi specifici.

Sviluppo delle Metriche (Mesi 3-6):

- Mese 5-6: Definizione delle nuove metriche per valutare la fairness in algoritmi AI.
- Mese 7-8: Implementazione delle metriche su dataset di test, ottimizzando e validando le performance.

Tecniche di Mitigazione e Applicazione Pratica (Mesi 6-9):

- Mese 9-10: Esplorazione e implementazione di tecniche avanzate di mitigazione dei bias.
- Mese 11-12: Applicazione pratica delle soluzioni sviluppate su casi studio reali, valutando l'efficacia delle tecniche di mitigazione.

Valutazione e Pubblicazioni (Mesi 9-12):

- Mese 13-14: Analisi dei risultati ottenuti nelle fasi precedenti, identificando punti di forza e possibili miglioramenti.
- Mese 15-16: Pubblicazioni accademiche e partecipazione a conferenze per condividere i risultati e contribuire alla comunità scientifica.

Questo piano spalmato su 24 mesi offre un approccio strutturato che copre ogni fase del progetto, dalla fase iniziale di analisi alla diffusione dei risultati. Ad ogni passo, l'obiettivo è garantire un contributo significativo alla comprensione e all'implementazione della fairness nei sistemi AI.