





Machine Learning Fairness Analisi empirica dello stato della pratica

Prof. Fabio Palomba Dott.ssa. Giulia Sellitto Carmine Ferrara Mat: 0522500990







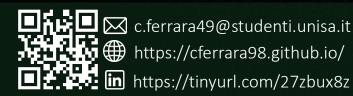
When you think of A.I., it's forward-looking.
But A.I. is based on data, and data is a
reflection of our history

Joy Buolamwini



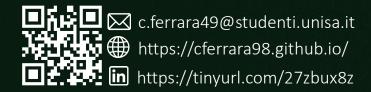
Le soluzioni di Machine Learning sono progettate e sviluppate da umani, e naturalmente gli uomini riproducono il loro giudizio all'interno delle soluzioni di Machine Learning





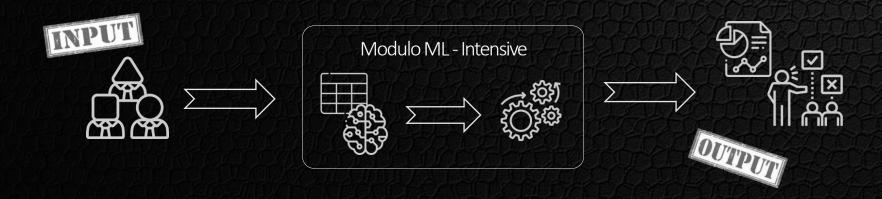


Nuove tipologie di vulnerabilità nello sviluppo ML-Intensive



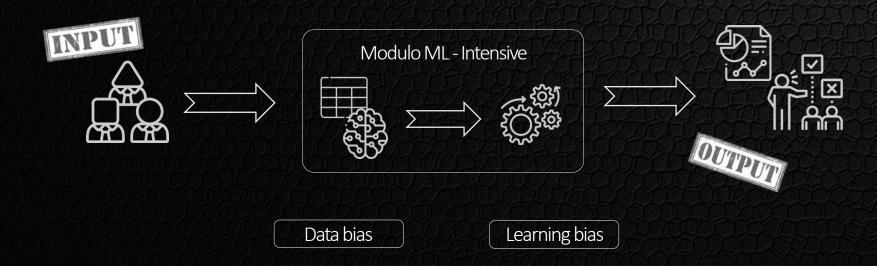


Nuove tipologie di vulnerabilità nello sviluppo ML-Intensive



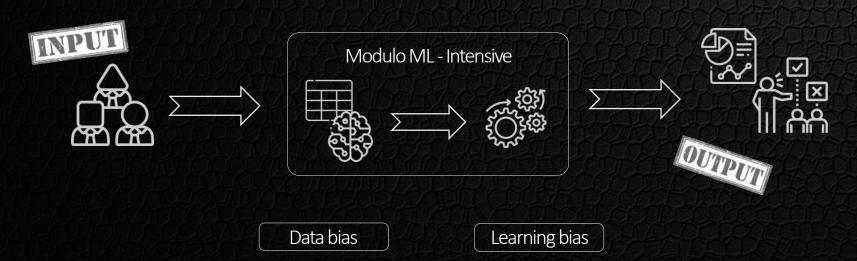


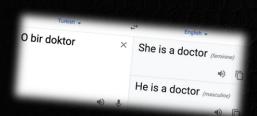
Nuove tipologie di vulnerabilità nello sviluppo ML-Intensive





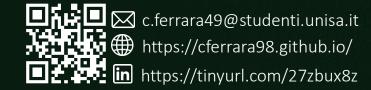
Nuove tipologie di vulnerabilità nello sviluppo ML-Intensive







Per alcune professioni quali "Dottore o Dottoressa", la traduzione dal turco di Google Translate non tengono conto del Gender del soggetto.





Machine Learning Fairness





Machine Learning Fairness

Un vero e proprio Hot Topic di ricerca

Con più di 20 differenti definizioni e metriche



Machine Learning Fairness

Un vero e proprio Hot Topic di ricerca

Con più di 20 differenti definizioni e metriche

Fairness: Set di requisiti, metodi e tecniche che rendono una soluzione di intelligenza artificiale eticamente corretta.



Machine Learning Fairness

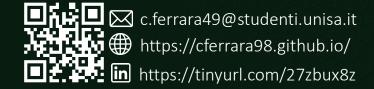
Un vero e proprio Hot Topic di ricerca

Con più di 20 differenti definizioni e metriche

Fairness: Set di requisiti, metodi e tecniche che rendono una soluzione di intelligenza artificiale eticamente corretta.

Approcci Probabilistici







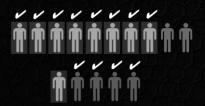
Machine Learning Fairness

Un vero e proprio Hot Topic di ricerca

Con più di 20 differenti definizioni e metriche

Fairness: Set di requisiti, metodi e tecniche che rendono una soluzione di intelligenza artificiale eticamente corretta.

Approcci Probabilistici





Approcci di Similarità



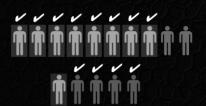
Machine Learning Fairness

Un vero e proprio Hot Topic di ricerca

Con più di 20 differenti definizioni e metriche

Fairness: Set di requisiti, metodi e tecniche che rendono una soluzione di intelligenza artificiale eticamente corretta.

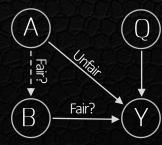
Approcci Probabilistici

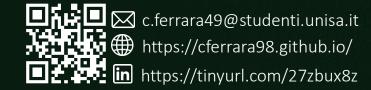




Approcci di Similarità

Approcci Causali



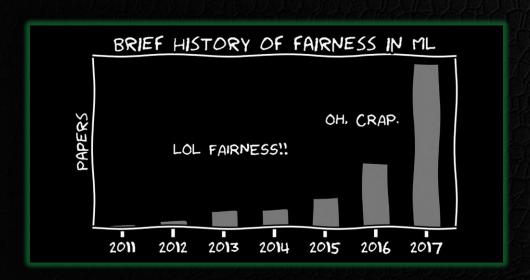




È osservabile come negli ultimi anni Fairness sia cresciuta considerevolmente come argomento di ricerca...



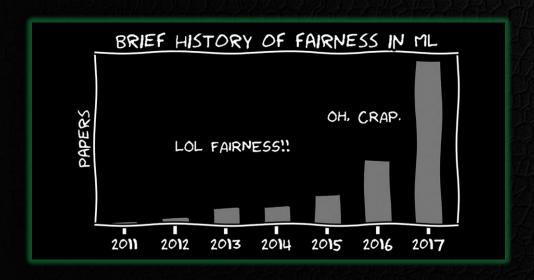
È osservabile come negli ultimi anni Fairness sia cresciuta considerevolmente come argomento di ricerca...







È osservabile come negli ultimi anni Fairness sia cresciuta considerevolmente come argomento di ricerca...



Dai numerosi studi di intelligenza artificiale atti a migliorare i livelli di fairness di un sistema ML-Intensive...





... Fino a quelli ingegneristici che trattano fairness come un vero aspetto di qualità nello sviluppo ML-Intensive



Ma come applicare fairness in casi pratici di sviluppo?





Ma come applicare fairness in casi pratici di sviluppo?

Secondo i più recenti studi di ricerca, Fairness è un'aspetto di qualità...





Ma come applicare fairness in casi pratici di sviluppo?

Secondo i più recenti studi di ricerca, Fairness è un'aspetto di qualità...

Astratto





Ma come applicare fairness in casi pratici di sviluppo?

Secondo i più recenti studi di ricerca, Fairness è un'aspetto di qualità...

Astratto

Context-Specific





Ma come applicare fairness in casi pratici di sviluppo?

Secondo i più recenti studi di ricerca, Fairness è un'aspetto di qualità...

Astratto

Context-Specific

Difficile da trattare praticamente

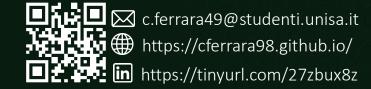




... si rende quindi necessario indagare empiricamente circa le pratiche lavorative nel trattamento di Fairness, cercando di capire...



Quali sono le definizioni o gli approcci maggiormente utilizzati?





... si rende quindi necessario indagare empiricamente circa le pratiche lavorative nel trattamento di Fairness, cercando di capire...

Quali sono le figure necessarie in un Team Fair-Oriented?





Quali sono le definizioni o gli approcci maggiormente utilizzati?



... si rende quindi necessario indagare empiricamente circa le pratiche lavorative nel trattamento di Fairness, cercando di capire...

Quali sono le figure necessarie in un Team Fair-Oriented?



Quali sono le definizioni o gli approcci maggiormente utilizzati?



Quanto Fairness è rilevante rispetto altre specifiche?



... si rende quindi necessario indagare empiricamente circa le pratiche lavorative nel trattamento di Fairness, cercando di capire...

Quali sono le figure necessarie in un Team Fair-Oriented?

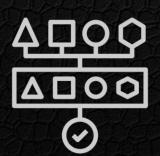


Quali sono le definizioni o gli approcci maggiormente utilizzati?



Quanto Fairness è rilevante rispetto altre specifiche?

In che fasi di sviluppo ML viene trattata Fairness?





... si rende quindi necessario indagare empiricamente circa le pratiche lavorative nel trattamento di Fairness, cercando di capire...

Quali sono le figure necessarie in un Team Fair-Oriented?

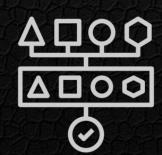


Quali sono le definizioni o gli approcci maggiormente utilizzati?



Quanto Fairness è rilevante rispetto altre specifiche?

In che fasi di sviluppo ML viene trattata Fairness?





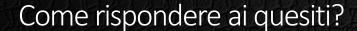
Quanto mature sono le compagnie Fair-Oriented?



Come rispondere ai quesiti?





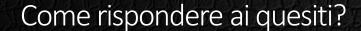




Redazione di un survey atto a coinvolgere esperti del settore









Redazione di un survey atto a coinvolgere esperti del settore

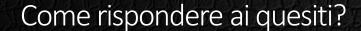


Manager
Data Scientists
Ingegneri del Software

• • • •









Redazione di un **survey** atto a coinvolgere esperti del settore



Manager
Data Scientists
Ingegneri del Software

....

Perché un **survey**?

Coinvolgere in modo mirato uno specifico gruppo di candidati

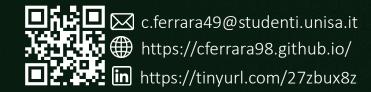
Integrare differenti culture e background

Ottenere rapidamente dati omogenei facili da analizzare



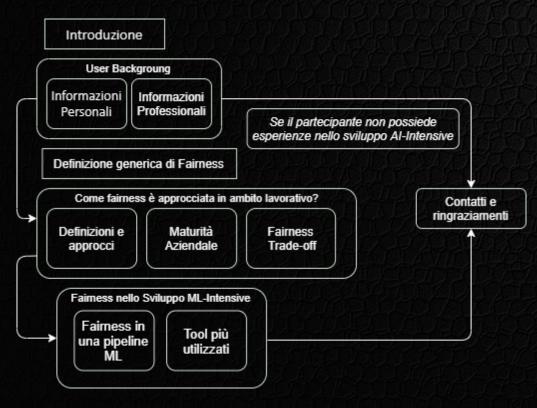


Dettagli sulla fase di studio Empirico





Dettagli sulla fase di studio Empirico



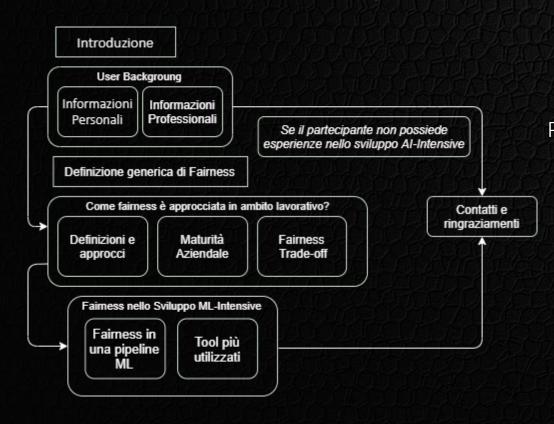
Test pilota con studenti di SE4Al



Piattaforma nota nel mondo della ricerca per questa tipologia di studi



Dettagli sulla fase di studio Empirico

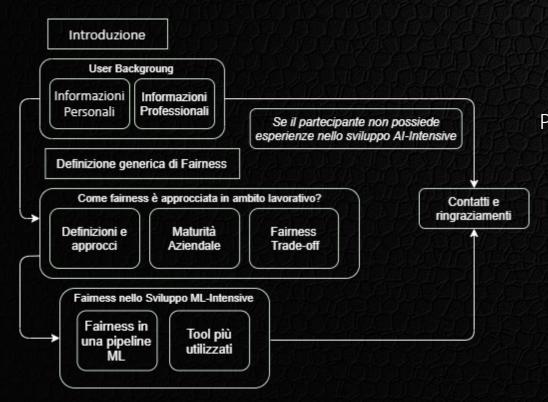


Test pilota con studenti di SE4AI



Piattaforma nota nel mondo della ricerca per questa tipologia di studi la quale offre sistemi di filtering automatici per i partecipanti

Dettagli sulla fase di studio Empirico



Test pilota con studenti di SE4AI

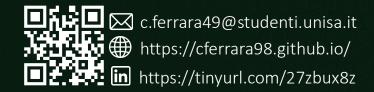


Piattaforma nota nel mondo della ricerca per questa tipologia di studi la quale offre sistemi di filtering automatici per i partecipanti

Pulizia dei dati

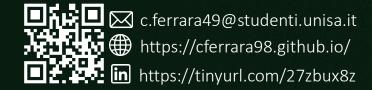
È stata condotta un'attività di data quality prescreening per rimuovere risposte non rilevanti e/o ottenute da figure non adeguate ai nostri obiettivi

203 risposte collezionate e 116 usate in analisi





Analisi dei Dati





Analisi dei Dati



Statistica descrittiva e Grafici



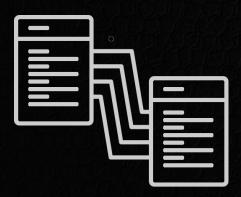






Statistica descrittiva e Grafici



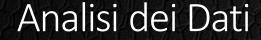


In fase di analisi è stato ricostruita la mappatura tra quesiti del Survey e gli obiettivi di ricerca



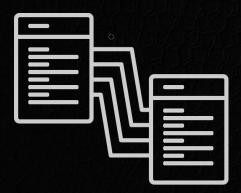


Metodologia di Ricerca



Statistica descrittiva e Grafici



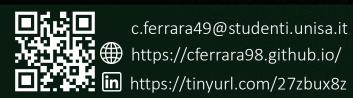


In fase di analisi è stato ricostruita la mappatura tra quesiti del Survey e gli obiettivi di ricerca

Attività di pre-processing come trasformazioni di scala o uso di acronimi utili all'elaborazione dei dati



Composizione del Campione





Composizione del Campione



Circa il 90% dei partecipanti all'indagine possiede un titolo di studi triennale o superiori



Tutti i 116 partecipanti hanno esperienza lavorativa con Machine Learning e Fairness



Composizione del Campione



Circa il 90% dei partecipanti all'indagine possiede un titolo di studi triennale o superiori



Tutti i 116 partecipanti hanno esperienza lavorativa con Machine Learning e Fairness



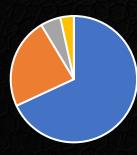
Provenienza dei partecipanti





■ Nord America









Composizione del Campione



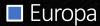
Circa il 90% dei partecipanti all'indagine possiede un titolo di studi triennale o superiori



Tutti i 116 partecipanti hanno esperienza lavorativa con Machine Learning e Fairness



Provenienza dei partecipanti



Africa

■ Nord America

Altro









Composizione del Campione



Circa il 90% dei partecipanti all'indagine possiede un titolo di studi triennale o superiori



Tutti i 116 partecipanti hanno esperienza lavorativa con Machine Learning e Fairness



Provenienza dei partecipanti



Africa

■ Nord America

Altro









■ Ingegneri del Software

Data Scientist

■ Manager

Più posizioni

Affini







Trattare Fairness in ambito aziendale









Trattare Fairness in ambito aziendale

Definizioni e Metriche Formali

Similarità degli Individui



Probabilità e Casual Reasoning







Trattare Fairness in ambito aziendale

Definizioni e Metriche Formali

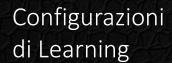
Similarità degli Individui

Approcci Pratici

Analisi delle dipendenze



Probabilità e Casual Reasoning











Trattare Fairness in ambito aziendale

Definizioni e Metriche Formali

Approcci Pratici

Similarità degli Individui

Analisi delle dipendenze



Probabilità e Casual Reasoning

Configurazioni di Learning





Altrimenti

Metodologie empiriche o Correlazione statistica









Trattare Fairness in ambito aziendale



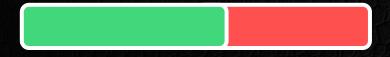




Trattare Fairness in ambito aziendale

Usability & Scalability

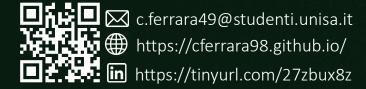
Accuracy & Sicurezza



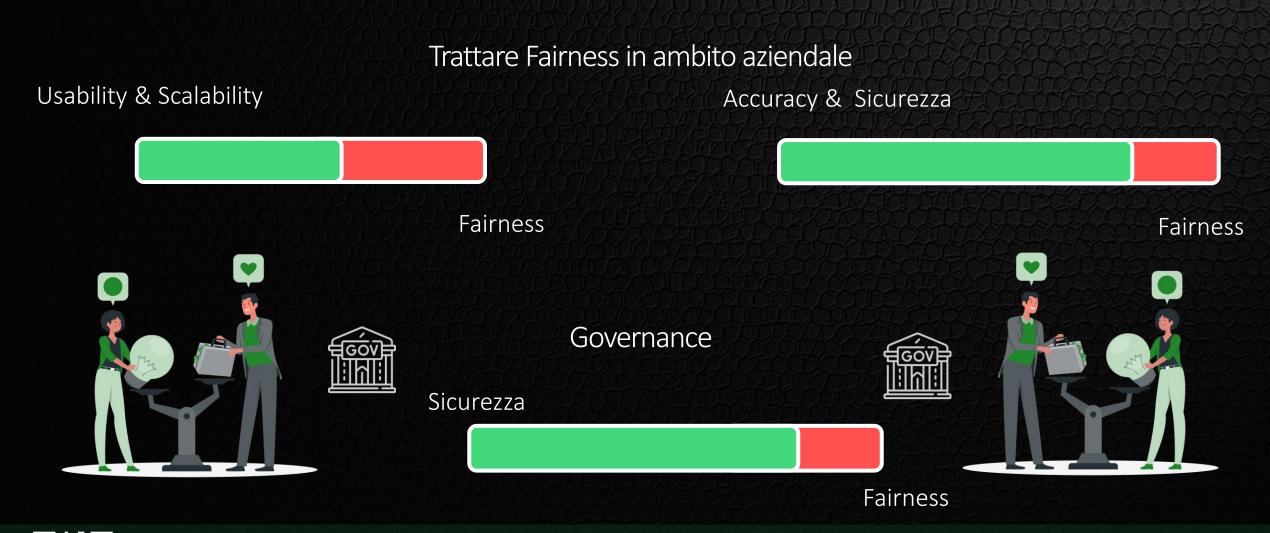
Fairness





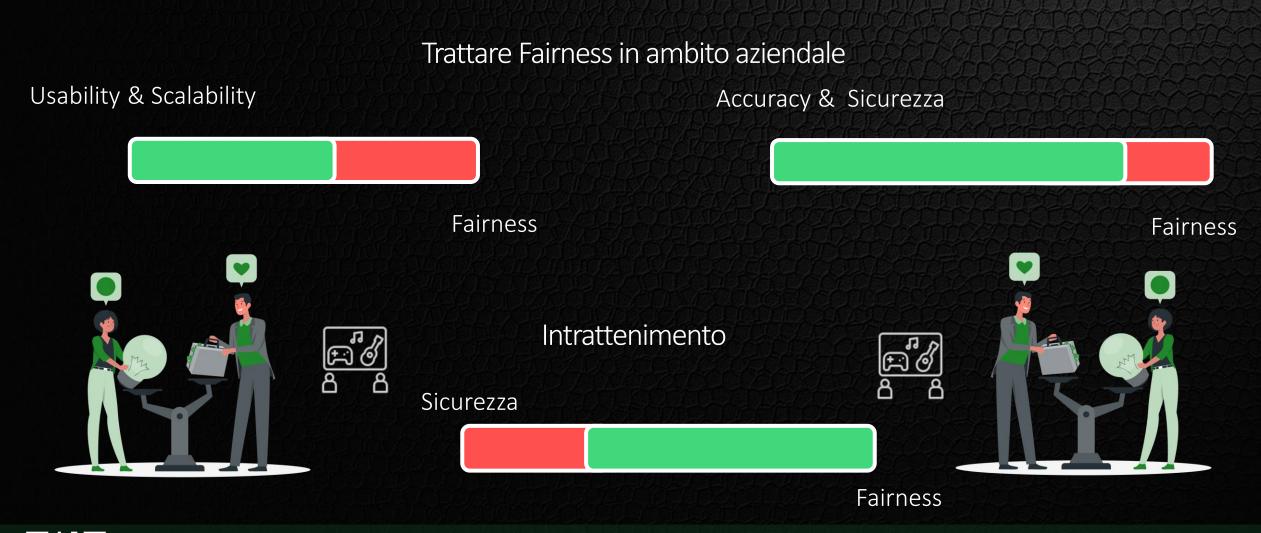
















Trattare Fairness in ambito aziendale









Trattare Fairness in ambito aziendale

Il trattamento di fairness è ancora poco maturo e standardizzato nelle pratiche aziendali ma questo aspetto è da considerarsi in costante miglioramento









Trattare Fairness in ambito aziendale

Il trattamento di fairness è ancora poco maturo e standardizzato nelle pratiche aziendali ma questo aspetto è da considerarsi in costante miglioramento



Il 26% degli intervistati tratta Fairness in maniera sporadica senza standard precisi







Trattare Fairness in ambito aziendale

Il trattamento di fairness è ancora poco maturo e standardizzato nelle pratiche aziendali ma questo aspetto è da considerarsi in costante miglioramento



Il 26% degli intervistati tratta Fairness in maniera sporadica senza standard precisi

Un ulteriore 24% afferma che la propria compagnia adotti standard di sviluppo Fair Oriented







Trattare Fairness in ambito aziendale

Il trattamento di fairness è ancora poco maturo e standardizzato nelle pratiche aziendali ma questo aspetto è da considerarsi in costante miglioramento



Il 26% degli intervistati tratta Fairness in maniera sporadica senza standard precisi

Un ulteriore 24% afferma che la propria compagnia adotti standard di sviluppo Fair Oriented





Della restante parte oltre il 30% dichiara di trattare le specifiche etiche con standard definiti e volti all'ottimizzazione

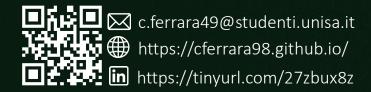








Implicazioni e Studi Futuri





Implicazioni e Studi Futuri

... L'immaturità di Fairness fa sì che studi di ricerca mirati siano necessari



Implicazioni e Studi Futuri

... L'immaturità di Fairness fa sì che studi di ricerca mirati siano necessari



Sistematizzazione di Definizioni e Approcci di Sviluppo per il trattamento di Fairness secondo le esigenze professionali e di dominio

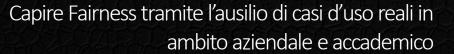


Implicazioni e Studi Futuri

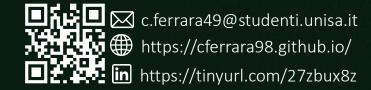
... L'immaturità di Fairness fa sì che studi di ricerca mirati siano necessari



Sistematizzazione di Definizioni e Approcci di Sviluppo per il trattamento di Fairness secondo le esigenze professionali e di dominio



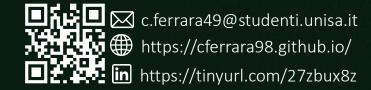






Team Ottimale per il trattamento di Fairness







Team Ottimale per il trattamento di Fairness





Figure come **Manager e Esperti Specifici** sono di estrema rilevanza





Team Ottimale per il trattamento di Fairness





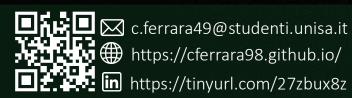
Figure come **Manager e Esperti Specifici** sono di estrema rilevanza





Data Scientists e Ingegneri del Softaware Figure trasversali con responsabilità equiparabili







Team Ottimale per il trattamento di Fairness





Figure come **Manager e Esperti Specifici** sono di estrema rilevanza





Data Scientists e Ingegneri del Softaware



Figure trasversali con responsabilità equiparabili

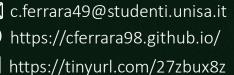


Implicazioni e Studi Futuri



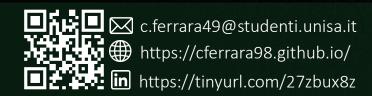
Standard di Responsabilità in un Team Fair Oriented







Fairness nello Sviluppo di Soluzioni ML-Intensive







Fairness nello Sviluppo di Soluzioni ML-Intensive

Fairness matura man mano che il modello evolve e matura



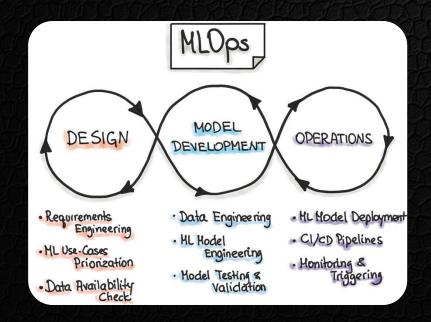




Fairness nello Sviluppo di Soluzioni ML-Intensive

Fairness matura man mano che il modello evolve e matura





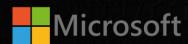




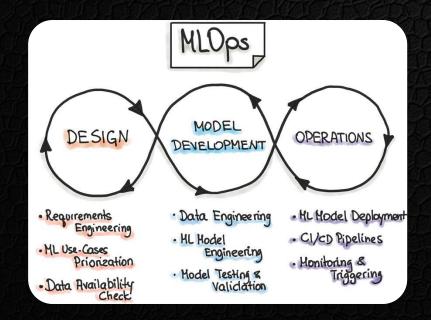
Fairness nello Sviluppo di Soluzioni ML-Intensive

Fairness matura man mano che il modello evolve e matura











What If...
you could inspect a machine learning model,
with minimal coding required?















Analisi di Bad & Best Practices Specifiche Implicazioni e Studi Futuri

Preparazione Dati di Training

Protocolli Gestionali Evolutivi Fair Oriented





c.ferrara49@studenti.unisa.it
https://cferrara98.github.io/
in https://tinyurl.com/27zbux8z





Grazie per l'attenzione!

Crediti: Flaticon (icone), Storyset (illustrazioni)

Machine Learning Fairness Analisi empirica dello stato della pratica







c.ferrara49@studenti.unisa.it



https://cferrara98.github.io/



https://tinyurl.com/27zbux8z



Backup

Introduzione e Background

Alcuni Esempi Famosi



Nel 2015, un tool di Recruting di Amazon, era solito giudicare effettuare discriminazioni sul Gender dei candidati.

Nel 2014 Compass, un tool americano AI - Intensive per l'analisi di colpevolezza, era solito giudicare facilmente, come colpevoli, le persone di colore.







Per alcune professioni quali "Dottore o Dottoressa", la traduzione dal turco di Google Translate non tengono conto del Gender del soggetto.



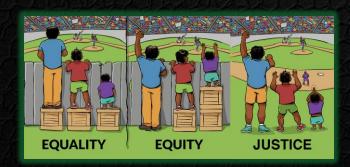


Introduzione e Background

Machine Learning Fairness



... che a seconda dei domini applicative può essere trattata in svariati modi Una caratteristica complessa che coinvolge innumerevoli fattori e contesti sociali...



Fairness: Set di requisiti, metodi e tecniche che rendono una soluzione di intelligenza artificiale eticamente corretta.

