



Corso di Laurea Magistrale in Informatica

Machine Learning Fairness

Analisi empirica dello stato della pratica

Prof. Fabio Palomba
Dott.ssa. Giulia Sellitto

Carmine Ferrara
Mat: 0522500990



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

sesa^{lab}
SOFTWARE ENGINEERING
SALERNO

*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 gli umani naturalmente riproducono il loro giudizio all'interno delle Soluzioni di Machine Learning*



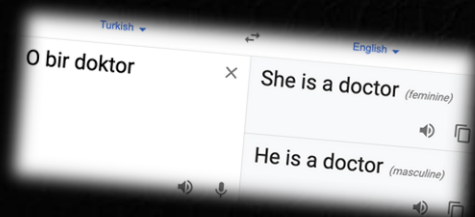
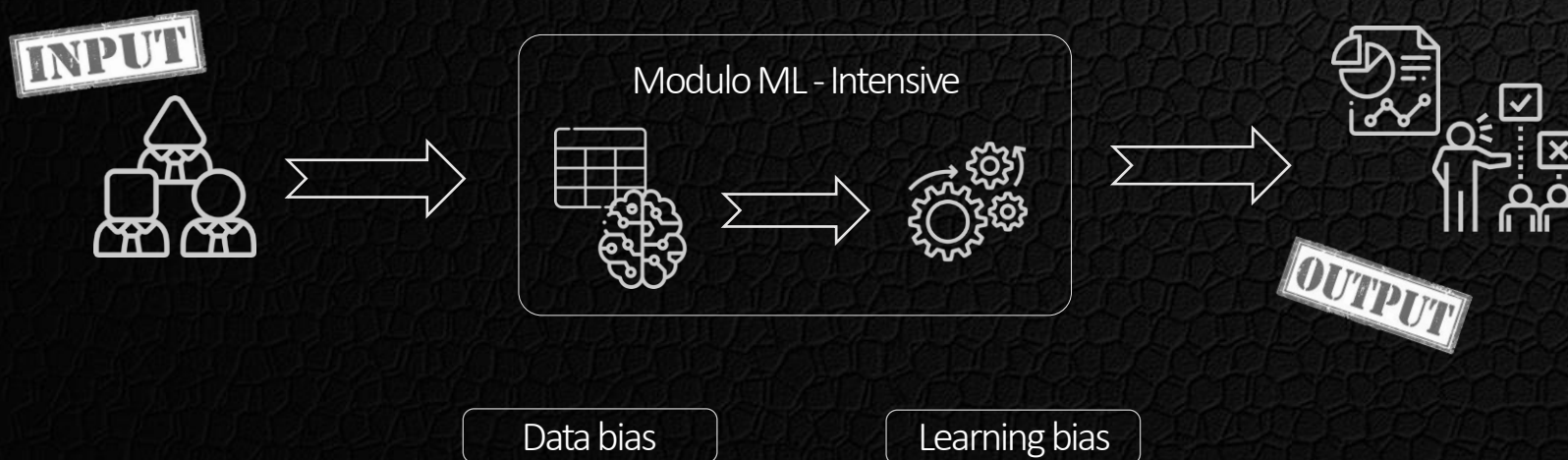
✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Introduzione e Background

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.



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Introduzione e Background

Machine Learning Fairness

Un vero e proprio *Hot Topic* di ricerca

Con più di 20 differenti *definizioni e metriche*

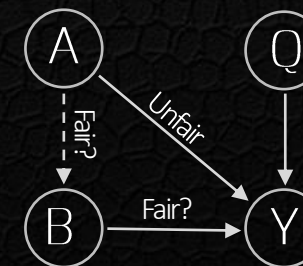
Fairness: Set di requisiti, metodi e tecniche che rendono una soluzione di intelligenza artificiale eticamente corretta.

Approcci Probabilistici



Approcci di Similarità

Approcci Causali



✉ c.ferrara49@studenti.unisa.it

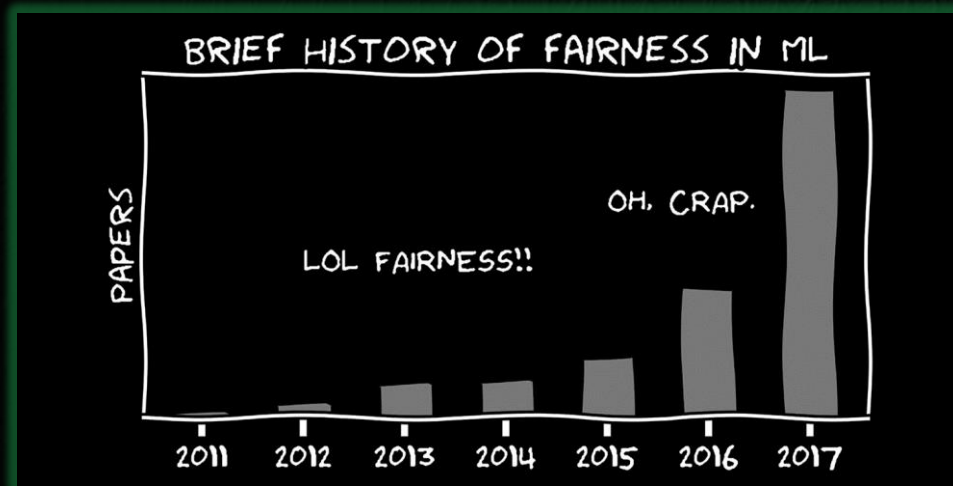
🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Problema e Obiettivi

Osservare la pratica lavorativa nel trattamento di Fairness?

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



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Problema e Obiettivi

*Fairness è un
concetto astratto*

*Fairness è difficile da trattare
praticamente*

*Fairness è
context-specific*

Osservare la pratica lavorativa nel trattamento di Fairness?

Obiettivo 1



Quali sono le definizioni o gli approcci
maggiormente utilizzati?



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Problema e Obiettivi

*Fairness è un
concetto astratto*

*Fairness è difficile da trattare
praticamente*

*Fairness è
context-specific*

Osservare la pratica lavorativa nel trattamento di Fairness?

Obiettivo 2



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



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Problema e Obiettivi

*Fairness è un
concetto astratto*

*Fairness è difficile da trattare
praticamente*

*Fairness è
context-specific*

Osservare la pratica lavorativa nel trattamento di Fairness?

Obiettivo 3



Quanto Fairness è rilevante
rispetto altre specifiche?



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Problema e Obiettivi

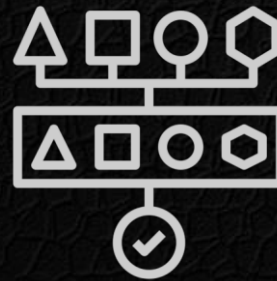
*Fairness è un
concetto astratto*

*Fairness è difficile da trattare
praticamente*

*Fairness è
context-specific*

Osservare la pratica lavorativa nel trattamento di Fairness?

Obiettivo 4



In che fasi di sviluppo ML
viene trattata Fairness?



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Problema e Obiettivi

*Fairness è un
concetto astratto*

*Fairness è difficile da trattare
praticamente*

*Fairness è
context-specific*

Osservare la pratica lavorativa nel trattamento di Fairness?

Obiettivo 5



Quanto mature sono le
compagnie Fair-Oriented?



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Metodologia di Ricerca

Dettagli sulla fase di studio Empirico



Test pilota con studenti di SE4AI



Piattaforma nota nel mondo della ricerca per questa tipologia di studi

Pulizia dei dati

Abbiamo condotto 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



✉ c.ferrara49@studenti.unisa.it

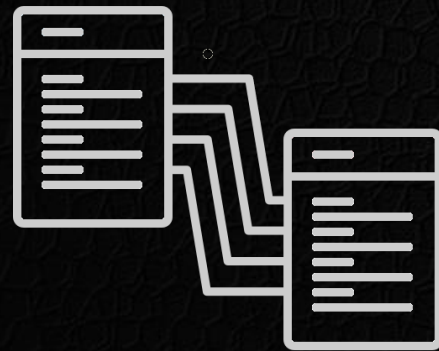
🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Analisi dei Dati



Statistica descrittiva e Grafici



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



✉ c.ferrara49@studenti.unisa.it

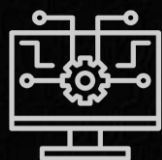
🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Risultati di Ricerca

Composizione del Campione

68% dei partecipanti dichiara di essere di provenienza europea



Più della metà dei partecipanti dichiara di avere anni di esperienza in Ingegneria del software, ma anche ruoli come Data Scientists o Manager sono ampiamente rappresentati



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



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Risultati di Ricerca

Trattare Fairness in ambito aziendale

Definizioni e Metriche Formali

Approcci Pratici

Similarità degli Individui

Over

Metriche di predizione o Casual Reasoning

Analisi di dipendenze dei Dati

Over

Configurazioni di Learning

I professionisti suggeriscono altre tecniche specifiche per il trattamento di Fairness



Metodologie empiriche o Correlazione statistica.



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Risultati di Ricerca

Trattare Fairness in ambito aziendale

Usability & Scalability A little Bit Over **Fairness**

Accuracy & Sicurezza Over **Fairness**

Sicurezza Over **Fairness** in ambito governativo



Fairness Over **Sicurezza** nell'Intrattenimento



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Risultati di Ricerca

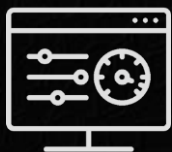
Fairness e Maturità Aziendale

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



Il 26% degli intervistati dichiara che la propria compagnia tratta Fairness in maniera sporadica e 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



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Risultati di ricerca

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

Capire Fairness tramite l'ausilio di casi d'uso reali in
ambito aziendale e accademico



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

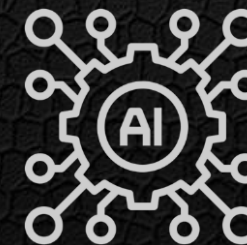
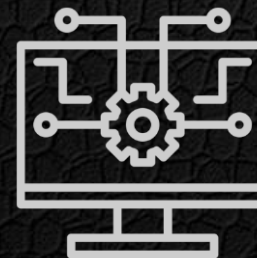
🌐 <https://tinyurl.com/27zbux8z>

Risultati di Ricerca

Team Ottimale per il trattamento di Fairness



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



Data Scientists e Ingegneri del Software

Figure trasversali con responsabilità equiparabili

Implicazioni e Studi Futuri



Standard di Responsabilità in un Team Fair Oriented



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

Risultati di Ricerca

Fairness nello Sviluppo di Soluzioni ML-Intensive



Fairness matura man mano che il modello evolve e matura

Fasi di una pipeline di Machine Learning per il trattamento di Fairness



Implicazioni e Studi Futuri



Protocolli Gestionali
Evolutivi Fair Oriented



Analisi di Bad & Best
Practices Specifiche



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

*Grazie
Per l'Attenzione!*

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 Recruiting 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.



✉ c.ferrara49@studenti.unisa.it

🌐 <https://cferrara98.github.io/>

🌐 <https://tinyurl.com/27zbux8z>

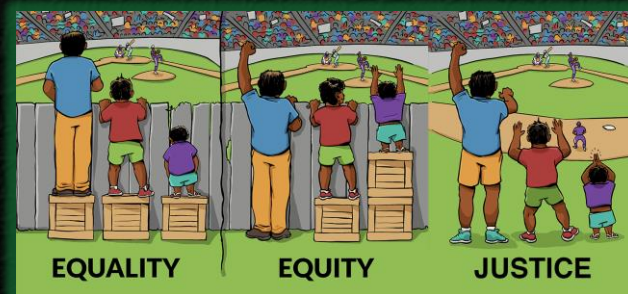
Introduzione e Background

Machine Learning Fairness




Una caratteristica complessa che coinvolge innumerevoli fattori e contesti sociali...


... che a seconda dei domini applicative
può essere trattata in svariati modi




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



 c.ferrara49@studenti.unisa.it

 <https://cferrara98.github.io/>

 <https://tinyurl.com/27zbux8z>