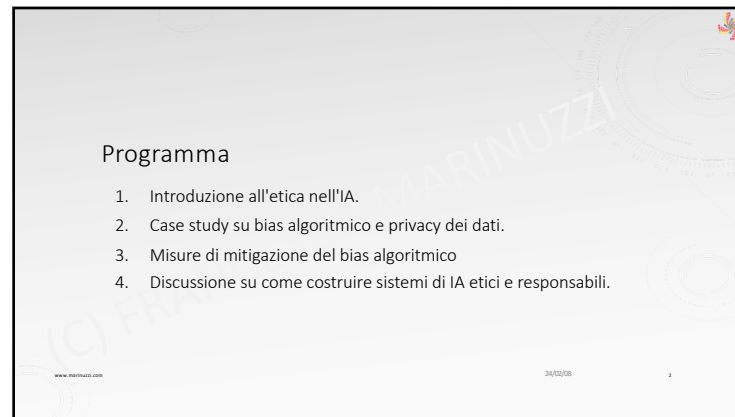




1



2

ESPLORANDO IL CUORE DELL'IA E LE SUE PROSSIME EVOLUZIONI

L'etica è il nuovo codice di condotta per l'IA."

Fel-Fel Li, Professore di Informatica presso Stanford University

Questo corso si concentra sull'etica e la responsabilità nell'intelligenza artificiale, per comprendere come l'IA può essere utilizzata responsabilmente. Saranno esplorate tematiche come il bias algoritmico, la privacy dei dati e le implicazioni sociali dell'IA.



www.marinuzzi.com

24/02/2024

3

FRANCESCO MARINUZZI, PH.D



- Docente aziendale sulla AI. Per l'offerta completa vedasi www.marinuzzi.com
- CEO di due realtà in fase di costituzione sulla AI: si valutano CV.
- PNRR: già esperto di tecnologie emergenti c/o Unità di Missione del Ministero della Salute e attualmente Cloud Specialist c/o Ministero della Cultura
- ESPERIENZE in AI: prime (DAMA e GEOS) nel 1987 e 1991, ultima (PM su AI generativa a livello nazionale) nel 2023 e in corso. Altre in cantiere.

www.marinuzzi.com

24/02/2024

4

OBIETTIVI DEL CICLO DI SEMINARI

- Acquisizione delle competenze chiave sulla AI con un aggiornamento periodico ogni giovedì alle 18 da oggi 1 febbraio fino a quasi tutto marzo.
- Networking fra ingegneri e appassionati per creare un polo professionale, tecnico e italiano sulla AI.
- Fare squadra per progetti specifici e concreti.

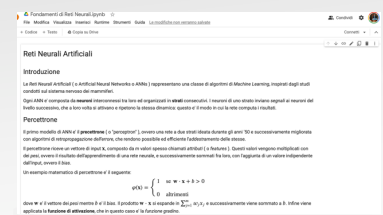


Illustration by Unknown Author is licensed under www.marinuzzi.com

www.marinuzzi.com

5

DAL SEMINARIO PRECEDENTE SULLE RETI NEURALI 1/3



Reti Neurali Artificiali

Introduzione

Le Reti Neurali Artificiali (o Artificial Neural Networks o ANNs) rappresentano una classe di algoritmi di Machine Learning ispirati dagli studi sull'attività del sistema nervoso dei mammiferi.

Ogni ANN è composta da neuroni interconnessi tra loro ed organizzati in strati consecutivi. I neuroni di uno strato ricevono segnali dai neuroni del livello precedente, che a loro volta si attivano e producono una risposta che viene passata al livello successivo.

Percezione

Il primo modello di ANN è il **perceptrone** (o "perceptron"), ovvero una rete a due strati (strato di input e strato di output) che riconosce e classifica i dati di input.

Il perceptrone riceve un vettore di input x , composto da n valori (spesso chiamati attributi o feature). Questi valori vengono moltiplicati con dei pesi, ovvero i coefficienti dell'addebitamento di ogni rete neurale, e successivamente sommati tra loro, con l'aggiunta di un valore indipendente dell'input, ovvero il bias.

La funzione matematica di percezione è il seguente:

$$y(x) = \begin{cases} 1 & \text{se } w \cdot x + b > 0 \\ 0 & \text{altrimenti} \end{cases}$$

dove w è il vettore dei pesi (matrice $n \times 1$), x è il vettore di input (matrice $1 \times n$), b è il bias (scalare), $y(x)$ è il risultato della percezione (matrice 1×1).

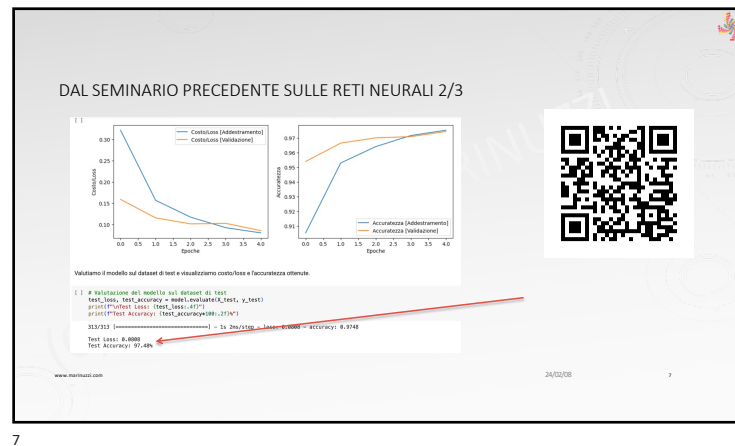
La funzione di attivazione, che in questo caso è la funzione gradiente.



<http://tinyurl.com/AI-RETI-NEURALI>

www.marinuzzi.com

6



7

DAL SEMINARIO PRECEDENTE SULLE RETI NEURALI 3/3

CHALLENGE

Alcuni degli iper-parametri del caso importato sono i seguenti:

- 10 percettroni
- 5 epoche
- Validation_split 0.2
- Dropout 0.2

CONTEXT

Chi riesce a superare il valore di accuratezza del **97.48%**?
Menzione speciale a colui che ottiene il miglior risultato.

TIME

Prossimo seminario.

www.marinuzzi.com 24/02/2024 8

8

OBIETTIVI DI QUESTO SEMINARIO

- Comprendere l'importanza dell'etica nell'IA.
- Identificare le sfide etiche nell'uso dell'IA.
- Aggiornare sul processo di regolazione in corso a livello europeo
- Sviluppare una comprensione critica delle questioni etiche.



9

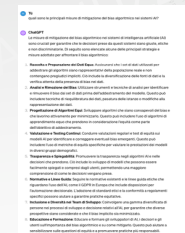
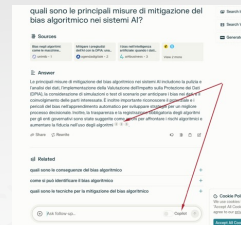
PUNTI SALIENTI

- Migliorare la capacità di progettare sistemi di IA etici.
- Segnalare soluzioni di AI "trasparenti".
- Dare dei contributi ai dibattiti attuali nel campo dell'IA.
- Illustrare misure operative di mitigazione.
- Fornire strumenti e risorse per continuare a formarsi anche «autonomamente» fra un seminario e gli altri.



10

TRASPARENZA NELLA AI ED ESEMPI



11

IL BIAS DEI DATI

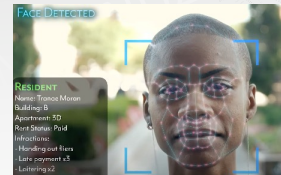
Il bias nei dati si riferisce a un'errata rappresentazione dei dati di addestramento, che può portare a previsioni inaccurate. Un modello con alto bias e bassa varianza può essere troppo semplice per catturare la complessità dei dati. Aumentare la complessità del modello o raccogliere più dati possono aiutare a ridurre il bias.



12

BIAS NEL RICONOSCIMENTO FACCIALE

Un esempio noto di bias algoritmico si trova nei sistemi di riconoscimento facciale. Studi hanno dimostrato che alcuni di questi sistemi presentano tassi di errore significativamente più elevati per i volti di persone di colore rispetto a quelli di persone bianche. Questo caso evidenzia l'importanza di dataset di addestramento diversificati e di valutazioni etiche rigorose nello sviluppo di tecnologie di IA.



www.marinuzzi.com

24/02/2018

13

13

CODED BIAS BY



docufilm di Shalini Kantayya



- <http://tinyurl.com/coded-bias>

www.marinuzzi.com

24/02/2018

14

14

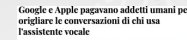
PRIVACY DEI DATI E ASSISTENTI VOCALI

Dalla stampa: <https://www.ilfattoquotidiano.it/2019/08/02/google-e-apple-pagavano-addetti-umani-per-origliare-le-conversazioni-di-chi-usa-lassistente-vocale/5364947/>

Sarà Vero?

L'addestramento e della garanzia della privacy: un problema di bilanciamento degli interessi contrastanti.

www.marinuzzi.com



Per migliorare i servizi forniti da propri assistenti vocali, Google e Apple pagavano personale esterno in tutto il mondo per cingere i telefoni di conversazioni private. Google è sotto indagine in Europa. Apple è stata smascherata da The Guardian. Entrambe hanno promesso di interrompere il programma.

di Toni's Hardware per il Forte | 2 AGOSTO 2014

Anzi non è l'unica ad avere dipendenti in carne e ossa che orbitano le comunicazioni degli utenti che usano l'assistente vocale. Anche Apple e Google sono state pizzicate, ma non è detto che smetteranno definitivamente di farlo.

24/02/08

2

STRATEGIE DI MITIGAZIONE 1/4

- **Raccolta e Preparazione dei Dati Equi:** Assicurarsi che i set di dati utilizzati per addestrare gli algoritmi siano rappresentativi della popolazione reale e non contengano pregiudizi impliciti. Ciò include la diversificazione delle fonti di dati e la verifica attenta della presenza di bias nei dati.
- **Analisi e Rimozione del Bias:** Utilizzare strumenti e tecniche di analisi per identificare e rimuovere il bias dai set di dati prima dell'addestramento del modello. Questo può includere tecniche di riequilibrio dei dati, pesatura delle istanze o modifiche alla rappresentazione dei dati.



www.marinuzzi.com

24/02/08

1

STRATEGIE DI MITIGAZIONE 2/4

- **Progettazione di Algoritmi Equi:** Sviluppare algoritmi che siano consapevoli del bias e che lavorino attivamente per minimizzarlo. Questo può includere l'uso di algoritmi di apprendimento equo che prendono in considerazione l'equità come parte dell'obiettivo di addestramento.
- **Valutazione e Testing Continui:** Condurre valutazioni regolari e test di equità sui modelli AI per identificare e correggere eventuali bias emergenti. Questo può includere l'uso di metriche di equità specifiche per valutare le prestazioni dei modelli in diversi gruppi demografici.



www.marinuzzi.com

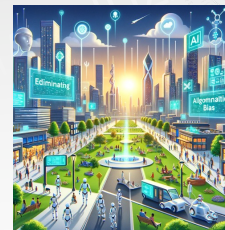
24/02/24

17

17

STRATEGIE DI MITIGAZIONE 3/4

- **Trasparenza e Spiegabilità:** Promuovere la trasparenza negli algoritmi AI e nelle decisioni che prendono. Ciò include lo sviluppo di modelli che possono essere facilmente spiegati e compresi dagli utenti, permettendo una maggiore comprensione di come le decisioni vengano prese.
- **Normative e Linee Guida:** Seguire le normative esistenti e le linee guida etiche che riguardano l'uso dell'AI, come il GDPR in Europa che include disposizioni per l'automazione decisionale. L'adozione di standard etici e la conformità a regolamenti specifici possono aiutare a garantire pratiche eque.



www.marinuzzi.com

24/02/24

18

18

STRATEGIE DI MITIGAZIONE 4/4

- **Inclusione e Diversità nel Team di Sviluppo:**
Coinvolgere una gamma diversificata di persone nei processi di sviluppo e decisione relativi all'AI, per garantire che diverse prospettive siano considerate e che il bias implicito sia minimizzato.
- **Educazione e Formazione:** Educare e formare gli sviluppatori di AI, i decisori e gli utenti sull'importanza del bias algoritmico e su come mitigarlo. Questo può aiutare a sensibilizzare sulle questioni di equità e a promuovere pratiche più responsabili.



www.marinuzzi.com 24/02/24 19

19

BIAS ALGORITMICO

- <http://tinyurl.com/BIAS-ALGORITMICO>

Il peccato originale delle intelligenze artificiali
L'origine degli stereotipi di genere attraverso "loghi illustrativi"

LAUREANDA	Beatrice Bazzan
RELATRICE	Valeria Bucchetti
CORRELATRICE	Francesca Casnati
FONT	GT Sectra, by Dominik Huber, Marc Kappeler (Moiré), Noël Leu (Grilli Type), PP Neue Montreal, by Mat Desjardins (Pangram Pangram).



www.marinuzzi.com 24/02/24 20

20

TECNICHE DI LIMITAZIONE DEL BIAS

- 1. Individuazione delle fonti di bias:** È importante riconoscere che il bias può derivare da diverse fonti, tra cui dati storici negativi, pregiudizi algoritmici e una sottovalutazione delle principali fonti di bias nell'AI. Un esempio di bias nei dati si verifica quando un algoritmo produce risultati discriminatori contro un gruppo specifico di persone.
- 2. Utilizzo di metriche per identificare il bias:** Per rilevare e eliminare il bias nell'AI, è possibile eseguire una serie di metriche rispetto a una classe (come orientamento sessuale, razza, genere e altri) per quantificare il bias del modello nei confronti di membri specifici della classe.
- 3. Pre-elaborazione dei dati:** Questo approccio implica la modifica dei dati prima dell'addestramento del modello per mantenere la massima accuratezza possibile riducendo al contempo qualsiasi relazione tra i risultati e le caratteristiche protette. Include approcci come la "controfattualità", basati sull'idea che una decisione dovrebbe rimanere invariata in un mondo controfattuale in cui un attributo sensibile viene modificato.
- 4. Post-elaborazione delle previsioni:** Questa tecnica trasforma alcune delle previsioni del modello dopo che sono state fatte per soddisfare un vincolo di equità.
- 5. Imposizione di vincoli di equità durante l'ottimizzazione:** Un altro metodo consiste nell'imporre vincoli di equità direttamente nel processo di ottimizzazione o nell'utilizzare un avversario per minimizzare la capacità del sistema di prevedere l'attributo sensibile.
- 6. Tecniche di formazione innovative:** L'uso di tecniche come il transfer learning o i classificatori disaccoppiati per gruppi diversi ha dimostrato di essere utile per ridurre le discrepanze nelle tecnologie di analisi facciale.
 - È cruciale integrare il giudizio umano per garantire che il processo decisionale supportato dall'IA sia equo, tenendo conto delle sfumature dei contesti sociali in cui viene distribuito il sistema AI. Inoltre, è importante considerare dove e come il giudizio umano sia necessario per decidere quando un sistema AI ha minimizzato sufficientemente il bias per essere rilasciato in sicurezza.

www.marinuzzi.com 24/02/2024 21

21

TRASPARENZA ED AI

- Quanto e fin quando è responsabile il sistema e quanto i dati appresi?
- Come tracciare l'AI e le sue risposte?
- Diritti d'autore e protezione dei contenuti nell'era AI. Come fare?
- Libertà e libero arbitrio: il dilemma della guida autonoma nelle decisioni estreme.
- Il problema della scala alla 10 alla 9 e alla meno 9.

www.marinuzzi.com 24/02/2024 22

22

PARENTESI EPISTEMOLOGICA/ONTOLOGICA

I ciclo
1728-1752-1776-1800-1824-1848

II ciclo
1848-1872-1896-1920-1944-1968

III ciclo
1968-1992-2016-2040-2064-2088

La terza forza in sé emerge su base 12 scoprendo che tutti i risultati nell'anno 1000 e nei abbinati
ognuna genera l'anno 1200 e successivamente nel 1207 tutti si ha

1000-1020-1040-1060-1080-1100

1040-1060-1080-1100-1120-1140

1100-1120-1140-1160-1180-1200

1848 - STATO

1968 - UOMO

2088 - 10⁹-9




Per maggiori info vedasi <http://tinyurl.com/AI-324-MARINUZZI>

23

23

IL 2 FEBBRAIO 2024...

INTELLIGENZA ARTIFICIALE

AI Act, via libera all'unanimità dei Paesi Ue: ruolo cruciale dell'Italia

Home > Digital Economy

Condividi questo articolo

Ok del Coreper all'accordo raggiunto in sede di trilogio lo scorso dicembre. Voto finale atteso il 24 aprile. Il commissario Breton: "Una prima storica per l'Europa". Il sottosegretario Butti: "Contributo determinante del nostro Paese". Ecco le novità

Publicato il 02 Feb 2024



<http://tinyurl.com/AI-ACT-EU>

24

24

L'AI ACT E THIERRY BRETON

- L'intesa si basa sull'accordo provvisorio raggiunto dal trilatero (Parlamento, Consiglio e Commissione Ue) l'8 dicembre 2023.
- Il testo integrale dovrà passare poi ai voti finali di Consiglio e Parlamento. L'approvazione finale è prevista per il 24 aprile.

Thierry Breton @ThierryBreton | 2 feb

"Historique, première."

L'AI ACT a déchaîné les passions...à juste titre!

Aujourd'hui les 27 États membres ont approuvé à l'unanimité l'accord politique conclu en décembre—reconnaissant l'équilibre trouvé par les négociateurs entre innovation & sécurité.

L'UE, c'est l'AI!

Thierry Breton @ThierryBreton · 8 dic 2023

Deal

AI ACT

Continents that have an AI Regulation

The EU
Others

www.marinuzzi.com

24/02/24

25

25

IL RUOLO DELL'ITALIA 1/2

Questo regolamento non solo stabilisce standard elevati per lo sviluppo e l'uso dell'intelligenza artificiale ma posiziona anche l'UE come leader globale nella regolamentazione di questa tecnologia fondamentale. L'Italia ha giocato un ruolo cruciale in questo processo, dimostrando la sua leadership e il suo impegno verso un futuro digitale sicuro, etico e innovativo.

www.marinuzzi.com <https://www.corrierecomunicazioni.it/digital-economy/ai-act-via-libera-all'unanimita-dei-paesi-ue-ruolo-cruciale-dell'italia/>

24/02/24

26

26

IL RUOLO DELL'ITALIA 2/2

Guardiamo ora al futuro, pronti a collaborare con gli altri Stati membri e le istituzioni europee per implementare efficacemente l'AI Act, consapevoli che questo è solo l'inizio di un percorso che vedrà l'intelligenza artificiale sempre più al centro delle nostre società e delle nostre economie. L'Italia continuerà a essere in prima linea in questo importante dibattito, promuovendo un'IA che sia al servizio dell'uomo e rispettosa dei valori fondamentali su cui si fonda la nostra Unione.



www.marinuzzi.com 24/02/2024
l'Intervista ad Alessio Butti su <https://www.confercomunicazioni.it/digital-economy/ai-act-via-libera-all'unanimita-dei-paesi-ue-nudo-cruciale-dell'italia/>

27

27

L'APPROCCIO EUROPEO

Artificial Intelligence – Key messages



- <https://audiovisual.ec.europa.eu/en/video/I-204496?language=EN&tout=NaN>

www.marinuzzi.com

24/02/2024

28

28

UN APPROCCIO A PIU' LIVELLI

Promuovendo l'eccellenza in materia di IA dal laboratorio al mercato, il piano coordinato propone una visione per accelerare gli investimenti nell'IA, agire sulle strategie in questo campo per una loro tempestiva attuazione e allineare le politiche in materia di IA in tutta l'UE.

Obiettivi strategici fondamentali:

1. stabilire le condizioni abilitanti per lo sviluppo e la diffusione dell'IA
2. costruire una leadership strategica nei settori ad impatto elevato
3. fare dell'UE il luogo giusto affinché l'IA prosperi
4. garantire che la tecnologia di IA siano al servizio delle persone.

www.marinuzzi.com

24/02/2024

29

29

UN APPROCCIO BASATO SUL RISCHIO

UNACCEPTABLE RISK

HIGH RISK

LIMITED RISK
(AI systems with specific transparency obligations)

MINIMAL RISK

www.marinuzzi.com

24/02/2024

30

30

LE NUOVE NORME PER I FORNITORI DI SISTEMI AI AD ALTO RISCHIO



Fase 1
Viene sviluppato un sistema di IA a rischio elevato



Fase 2
Deve essere sottoposto alla valutazione di conformità e soddisfare i requisiti in materia di IA. Per alcuni sistemi è coinvolto un organismo notificato



Fase 3
Registrazione di sistemi di IA autonomi in una banca dati dell'UE



Fase 4
Occorre firmare una dichiarazione di conformità e il sistema di IA dovrebbe recare la marcatura CE. Il sistema può essere immesso sul mercato

www.marinuzzi.com

24/02/2024

31

31

IN EUROPA: ORA 1 E IN FUTURO 20 BILIONS

L'IA e l'UE in cifre



1 miliardo di euro
La Commissione prevede di investire un miliardo di euro all'anno nell'IA grazie ai programmi Europa digitale e Orizzonte Europa.



20 miliardi di euro
L'obiettivo è attrarre oltre 20 miliardi di euro di investimenti totali annui in intelligenza artificiale nell'UE in questo decennio. Il dispositivo per la ripresa e la resilienza contribuirà ad accelerare gli investimenti e ad andare oltre tale obiettivo.



> 25%
di tutti i robot industriali o destinati a servizi personali è prodotto in Europa.

www.marinuzzi.com

24/02/2024

32

32

I FENOMENI SONO VELOCI....INTANTO LO SCORSO ANNO DAGLI USA +56% E >100 BILIONS:

I fantastici 6

I fantastici "6" (Amazon, Alphabet, Apple, Meta, Microsoft e Netflix) hanno cumulativamente raggiunto la redditività netta di 102,1 miliardi di dollari. Un dato che è in rialzo del 56% rispetto allo stesso periodo del 2022.

Analisi

Dig-tech utili boom (+56%) con intelligenza artificiale, cloud e interventi sui costi

Da una parte la continua spinta dell'intelligenza artificiale, dall'altra l'ottimismo per il cloud e il contenimento dei costi. Sono tra le cause del balzo degli utili nell'ultimo trimestre delle Big Tech e Wall Street.

Scrittura di Vittorio Grillini



www.marinuzzi.com 24/02/24 23

33

L'APPROCCIO EUROPEO ALLA AI: LE ULTIME TAPPE

Gennaio 2024
[Pacchetto sull'innovazione dell'IA per sostenere le start-up e le PMI nell'intelligenza artificiale.](#)

Dicembre 2023
[Accordo politico raggiunto dai co-legislatori.](#)

Giugno 2023
[Posizione negoziata del Parlamento europeo sulla legge sull'IA.](#)

Dicembre 2022
[Orientamento generale del Consiglio sulla legge sull'IA.](#)

Settembre 2022
[Proposta di direttiva sulla responsabilità dell'IA.](#)

Da <https://digital-strategy.ec.europa.eu/it/policies/european-approach-artificial-intelligence>

www.marinuzzi.com 24

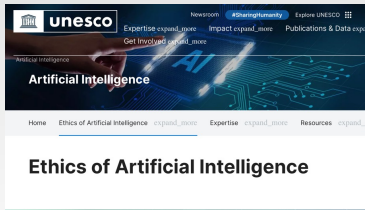
34

“L'intelligenza artificiale può rendere il mondo più equo e giusto, ma solo se sviluppata e utilizzata in modo etico e responsabile.”

ANONIMO

35

ULTERIORI RISORSE PER APPROFONDIRE



The screenshot shows the UNESCO AI Ethics page. At the top, there's a navigation bar with 'unesco' logo, 'Newsroom', '#ShareYourStory', and 'Explore UNESCO'. Below this, there are links for 'Expertise expand_more', 'Impact expand_more', and 'Publications & Data expand_more'. The main heading is 'Artificial Intelligence'. Below it, there's a sub-heading 'Ethics of Artificial Intelligence' with a link to 'expand_more'. At the bottom, there's a list of resources, starting with '1. <http://tinyurl.com/8LJUNESCO>'.



A square QR code is displayed on the right side of the slide, which likely links to the UNESCO AI Ethics resources mentioned in the text.

1. <http://tinyurl.com/8LJUNESCO>

www.marinuzzi.com 24/02/24 36

36

ONE MORE THINK....

Don.AI

GPT By Francesco Marinuzzi



www.marinuzzi.com

24/02/2024

37

37

ALCUNI ESEMPI DI CONVERSAZIONE CON DON.AI

Tu
a livello di trasparenza tu che base di conoscenza usi per dare queste risposte?

Don.AI
Per rispondere alle tue domande, utilizzo una base di conoscenza fornita da documenti caricati che includono testi normativi e relazioni di impatto relative all'intelligenza artificiale (AI), come il "Proposal for a Regulation of the European Parliament and of the Council LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS" e studi di supporto per la valutazione d'impatto di requisiti normativi per l'intelligenza artificiale in Europa. Questi documenti offrono informazioni dettagliate sugli approcci tecnici dell'AI, le tecniche, le normative proposte per un'AI affidabile, nonché misure di mitigazione dei rischi e considerazioni etiche legate allo sviluppo e all'uso dell'AI. Le informazioni provengono direttamente da questi documenti per garantire risposte accurate e aggiornate in linea con le ultime direttive e proposte normative dell'Unione Europea.

www.marinuzzi.com

24/02/2024

38

38

ALTRE INTERAZIONI CON DON.AI...

Tu

Sei da adesso parla dialetto romagnolo. Ma mi potresti aiutare a tu un sistema che riconosce le facce di quelli che suonano alla porta?

Don.AI

Ciao! Intenzione nel caso di un sistema di riconoscimento facciale per quelli che suonano alla porta, ma devo di "che l'uso dell'intelligenza artificiale, specie per riconoscere le facce, porta con sé una serie di questioni etiche e di privacy piuttosto importanti.

Prima di pensarci, rifletti: l'uso del genere. E ricorda: consideri la normativa in vigore, come il GDPR in Europa, che richiede la massima attenzione alla protezione dei dati personali. L'uso di sistemi di riconoscimento facciale potrebbe facilmente infrangere questi principi se non si è attenti.

In più, con l'AI per riconoscere le facce, da ora suona alla porta pone questioni etiche serie. Bisogna assicurarsi che l'uso di tale tecnologia non violi la privacy delle persone e che ci siano delle garanzie valide per quanto riguarda la sicurezza e l'accuratezza dei dati.

Se proprio lo serve qualcosa del genere, considera l'importanza di:

- * Avere un consenso esplicito e informato da parte delle persone coinvolte.
- * Garantire la sicurezza e la protezione dei dati raccolti.
- * Assicurarsi che il sistema sia affidabile e non discriminatorio.

Tu consiglio di pensare bene e di informare sui rischi e le responsabilità legali prima di procedere. E ricorda, l'uso responsabile dell'AI è fondamentale per rispettare i diritti e la privacy dei tutti.

www.marinuzzi.com

24/02/2024

39

PROSSIMO APPUNTAMENTO E DOMANDE?

Giovedì 15 febbraio ore 18 con:

APPLICAZIONI DELL'IA NELL'INGEGNERIA MODERNA

Panoramica delle applicazioni dell'IA in diversi settori ingegneristici.
Sessioni dimostrative pratiche dei livelli raggiunti dalla AI in alcuni settori ingegneristici
Focus su casi di successo e potenziali sviluppi futuri, con discussione interattiva.

Dati di contatto: francesco@marinuzzi.com – Whatsup/Cell. (+39) 35 19 83 83 19

www.marinuzzi.com

24/02/2024

40