# PUNTI DEBOLI E BIAS DELL'INTELLIGENZA ARTIFICIALE: SECONDA PARTE

**Data** 01 ottobre 2023
**Categoria** Medicinadigitale

In alcune precedenti pillole ( vedi ad esempio: pillole.org/public/aspnuke/print.asp?print=news&pID=7655 ) abbiamo esaminato le caratteristiche dei nostri ragionamenti, in particolare in ambito diagnostico, e ci siamo soffermati in particolare su alcune strategie, meglio definibili come "Euristiche", molto utilizzate in tutti i porocessi cognitivi perchè ci permettono di prendere decisioni in tempi brevi, ma che ci fanno tuttavia incorrere in errori inconsapevoli, definiti "Bias".
In questa nuova pillola elencheremo i più importanti errori cognitivi,del tutto involontari ed inconsapevoli, che anche i migliori informatici a volte"trasmettono" alle intelligenze artificiali. Come ci ha insegnato il premio Nobel Daniel Kahneman questi errori sono legati alle caratteristiche delle nostre menti e possono essere ridotti e corretti ma mai del tuttoeliminati...
Ecco l'elenco dei vizi di ragionamento ("bias") riscontrati nei programmi digitali...

**Bias di pregiudizi**
Il programmatore può inserire in maniera implicita e inconscia un determinato 'preconcetto' nel programma di IA: in tal modo opinioni, pregiudizi, aspettative culturali e sociali nei riguardi di razza, genere, sesso biologico, cultura, credenze, abitudini, atteggiamenti religiosi, ecc. (che preesistono in coloro che progettano il sistema) possono essere trasmessi inconsapevolmente alla tecnologia stessa. Quindi i sistemi di analisi dei dati delle Machine Learning potrebbero avere atteggiamenti discriminatori nei confronti di determinati gruppi di persone. Un esempio tipico è stato riscontrato nell'algoritmo OPTUM: utilizzato negli ospedali statunitensi per individuare quali pazienti avrebbero avuto necessità di cure mediche ulteriori, esso è stato fortemente criticato poiché avrebbe discriminato i pazienti neri rispetto a quelli bianchi.
**Bias di incertezza-campionamento**
Questo è un bias che altera le conclusioni dei processi algoritmici, poiché essi 'scotomizzano' i piccoli gruppi e le sotto-popolazioni statisticamente poco rilevanti. In pratica i risultati dell'elaborazione risultano validi solo per i campioni più grandi, ignorando i dati riguardanti popolazioni sottorappresentate nel dataset.
**Bias di 'appropriatezza'**
Viste le modalità operative dell'IA, di fronte a sintomi-segni aspecifici e poco definiti, l'algoritmo tende a proporre una grande quantità di esami ed accertamenti. Mancando una corretta profilazione (vedi più avanti) e la capacità di cogliere tutta la ricchezza del 'non-verbale', il sistema utilizza una 'rete a strascico' che, se utile ai fini di trovare la diagnosi corretta, non tiene conto dell'appropriatezza delle richieste, non avendo problemi di budget.
**Bias di 'profilazione'**
A tutt'oggi, gli algoritmi di solito registrano dati personali riguardanti la genetica, i gusti e gli stili di vita. Non riescono invece a cogliere tutta la ricchezza e le sfumature che servono a 'profilare' correttamente il paziente. Non possono 'catturare' il 'non verbale' con i suoi silenzi molto significativi, la postura e il tono della voce, le incertezze e i sentimenti espressi dalla mimica, ecc. Pertanto possono arrivare a conclusioni nettamente sbagliate.
**Bias di 'cornice'**
All'IA manca quello che M. Sclavi descrive come 'ascolto attivo': cioè la capacità di 'cambiare cornice-campo'. L'indagine 'variazionale' prevede che il soggetto, che osserva-conosce la realtà che lo circonda, usi la logica della scoperta fenomenologica per 'variare' il campo di osservazione: ciò gli permette di modificare il proprio punto di vista per comprendere il punto di vista di un'altra persona o di un'altra 'realtà'. Questa abilità è fondamentale nella comunicazione, in particolare nel colloquio tra medico e paziente, laddove le rispettive 'agende' cognitivo-comportamentali non si accordino. Tale situazione spesso genera conflitti che non è facile superare se non si pratica l'ascolto attivo. In queste situazioni affidare agli algoritmi le decisioni cliniche comporta errori di sistema che ricadono poi sull'assistito.
**Bias di 'ragionamento'**
L'IA non pensa, non ragiona; è una macchina capace di accumulare milioni di 'stringhe'-dati da cui estrarre quello che si avvicina di più al quesito posto dall'operatore (in questo caso il medico). Ciò, se da una parte accelera la capacità decisionale del clinico, può portare però a conclusioni false. Il ragionamento del medico, che tiene conto di tutto quel 'sapere di sfondo' di cui parla Popper, è ancora indispensabile nel processo diagnostico-terapeutico.
**Bias di 'condizionamento'**
L'IA rischia di condizionare pesantemente il mondo sanitario, cioè di controllarne le decisioni, con ricadute negative per la libertà di scelta dei pazienti. Y.N.Harari, nel suo libro "21 lezioni per il XXI secolo" afferma: "In passato ci si sentiva perfettamente in salute finché non si provava dolore o non si soffriva di un'evidente disabilità o malessere. Ma nel 2050, grazie ai sensori biometrici e agli algoritmi dei Big Data, le malattie potranno essere diagnosticate e trattate molto prima che portino al dolore o alla disabilità. Di conseguenza si soffrirà sempre di qualche disturbo e si seguirà questa o quella raccomandazione algoritmica. Se rifiuti di curarti, la tua assicurazione sanitaria non sarà più valida o il capo ti licenzierà: perché dovrebbero pagare il costo della tua ostinazione? Chi avrà il tempo e l'energia di gestire tutte queste malattie? Con ogni probabilità il nostro algoritmo." Il fatto che il paziente dipenda dalle decisioni dell'algoritmo, anche se magari non le condivide, impatta pesantemente e sistematicamente sulla libertà di ciascuno di noi.

**SOLUZIONIPOSSIBILI**

Vista l'importanza del problema, molti sono i tentativi di correggere questi gap e questi bias dell'IA. L' Algorithmic fairness è un campo di ricerca in grande sviluppo, che cerca di ridurre gli effetti di pregiudizi e discriminazioni ingiustificate dell'IA sugli individui. È un ambito di ricerca interdisciplinare che ha lo scopo di creare sistemi di auto apprendimento in grado di effettuare previsioni più eque e giuste.

Queste sono alcune delle soluzioni proposte:
 Rimozione di nomi e pronomi di genere (Anonymization and Direct Calibration)
 Utilizzo della stessa deep learning per trovare le distorsioni del sistema (Adversarial Learning)
 Riduzione dei pregiudizi (umani) che emergono dai set di dati su cui si basano queste applicazioni (Data cleaning): infatti se noi per primi correggiamo i nostri comportamenti discriminatori, anche i sistemi algoritmici che utilizziamo finiranno per averne meno
 Utilizzo di sistemi cosiddetti "ibridi", ossia che condividono sia il giudizio umano che quello dell'intelligenza artificiale.

Per finire ci sembra utile segnalare che anche la Unione Europea ritiene di dover aggiornare le proprie indicazioni operative, purtroppo superate dalla velocità e dalla potenza trasformativa di questi strumenti tecnologici: ecco il link....
altalex.com/documents/news/2023/06/23/ai-act-ue-traccia-futuro-intelligenza-artificiale

**Roberto Fassina, Riccardo De Gobbi, Giampaolo Collecchia**

**Bibliografia**

Daniel Kahneman :Pensieri lenti e veloci Mondadori Editore Milano 2020
G. Collecchia, R. De Gobbi. Intelligenza artificiale e medicina digitale. Il Pensiero Scientifico Editore, 2020
Y. N. Harar:21 lezioni per il XXI secolo. Giunti Editore SpA /Bompiani, 2018
M. Scav:. Arte di ascoltare e mondi possibili. Pearson, 2022
Afroditi KARIDOMATIS. https://www.i-com.it/2022/01/14/bias-intelligenza-artificiale/

G. Rizzi. https://ibicocca.unimib.it/bias-negli-algoritmi-come-le-macchine-apprendono-i-pregiudizi-dagli-esseri-umani/
M. T. Cimmino. https://ibicocca.unimib.it/bias-negli-algoritmi-come-le-macchine-apprendono-i-pregiudizi-dagli-esseri-umani/
S. Pescuma. https://ibicocca.unimib.it/bias-negli-algoritmi-come-le-macchine-apprendono-i-pregiudizi-dagli-esseri-umani/

L. Di Giacomo, https://www.diritto.it › regolamento-ue-intelligenza-artificiale/

**Per Approfondire la applicazioni della Intelligenza Artificiale** :
Collecchia G. De Gobbi R.: Intelligenza Artificiale e Medicina Digitale. Una guida critica. Il Pensiero Scientifico Ed. Roma 2020
pensiero.it/catalogo/libri/pubblico/intelligenza-artificiale-e-medicina-digitale