Decisioni automatizzate su algoritmi: il quadro giuridico

Estratto dalla relazione di Enrico Pelino tenutasi al 10° Cyber Crime Conference 2019

Entriamo subito nel vivo della materia: si può essere condannati da un software?
La risposta è sì e non stiamo parlando del futuro, ma di qualcosa che già esiste.
Negli Stati Uniti c'è stato un caso eclatante: Eric Loomis, fermato dalla polizia del Wisconsin per due imputazioni minori (che, a quanto apprendiamo, nel sistema statunitense non avrebbero determinato la pena della reclusione) è stato poi condannato a 6 anni di carcere.

Com'è potuto accadere?

Perché è stato utilizzato un software predittivo chiamato COMPAS, che, analizzando una serie di input riguardanti la storia penale di Loomis ma anche vari altri aspetti della sua vita, l'ha indicato come un soggetto ad alto rischio di commettere, in futuro, ulteriori reati; quindi parliamo di un soggetto che non è stato condannato soltanto per cose fatte in passato – recidiva "classica" – ma anche per quello che potrebbe fare nel suo futuro. Una cosa sconvolgente, che ci porta alla mente film distopici come Minority Report.

Come può succedere tutto questo?

Per contestualizzare è necessario fare un passo indietro. Negli Stati Uniti, fin dagli anni Venti vengono svolte analisi

predittive soprattutto per quanto riguarda la liberazione condizionale, quindi allo scopo di capire se un soggetto rappresenti un pericolo una volta rientrato nella società. Poi, negli anni ’70, questi studi vengono condotti in maniera più sofisticata: la ricerca, frequentemente citata, di Wolfgang, Figlio e Sellin è considerata un lavoro seminale in questa materia. Questi tre ricercatori scoprono, esaminando un campione di reati, che oltre il 50% di essi viene commesso da un piccolo numero di persone; il 6-7% commette più del 50% dei reati.

Di conseguenza, ipotizzano che se si riuscisse a “bloccare” questi soggetti si avrebbe un’incidenza di reati molto minore in quel determinato contesto. Invece di avere carceri sovraffollate, teorizzano di tenervi questi soggetti per più tempo e detenere, invece, per meno tempo (o non detenere affatto) chi presenti un profilo di rischio molto basso, in quanto è improbabile che commetta altri reati. Le risorse risparmiate potrebbero essere utilizzate, ad esempio, per riabilitare i soggetti ad alto rischio.

Quindi c’è qui un brillante obiettivo di politica criminale: abbattere della metà l’incidenza dei reati in un certo settore. Il problema, a parte le agghiaccianti implicazioni – come condannare qualcuno o fargli scontare un periodo di detenzione maggiore perché si “prevede” che possa commettere nell’immediato futuro altri reati – è: come facciamo a profilare questi soggetti, a capire chi sono? La teoria della selective incapacitation è infatti completamente fallita negli anni ‘80 perché dava luogo a molti falsi positivi e molti falsi negativi.

Oggi abbiamo, però, strumenti di gran lunga più sofisticati: possiamo accedere a set statistici molto più ricchi (che implicano una maggiore qualità del risultato) e abbiamo strumenti di machine learning. Ma quindi possiamo (davvero) premere un bottone e sapere se il signor Loomis presenti un alto rischio di commettere reati in futuro? Questo,

ovviamente, ci spaventa.

Un altro punto correlato è che, se io faccio degli investimenti per realizzare algoritmi predittivi, voglio che questi investimenti siano protetti. Nel caso Loomis la difesa ha cercato in tutti i modi – anche rendendosi disponibile a sottoscrivere un non disclosure agreement – di avere accesso alla logica del sistema COMPAS per capire quale peso venisse dato a certi input rispetto ad altri: non è stato possibile perché si tratta di un sistema proprietario, di una black box.

Questo è un primo tema. Il secondo è più generale: quando costruiamo un processo facciamo una scelta, cioè selezioniamo alcune cose che riteniamo importanti e ne scartiamo altre. Chiaramente, questa scelta ha delle conseguenze nel processo. E inoltre, spesso, questa scelta riguarda dei proxy: uso cioè degli elementi che ne rappresentano altri, però, in questo modo, mi sto allontanando dalla realtà. Sto compiendo un'astrazione. Applico una logica, incrocio dei dati; quindi posso essere anche molto sofisticato, ma lavoro comunque su un piano astratto e lo scarto tra questo piano e la realtà può determinare un bias, ossia un pregiudizio. Paradossalmente, cercando di correggere il pregiudizio umano – lo human bias – potrei determinare il pregiudizio chiamato machine bias, cioè quello prodotto dall'utilizzo di un algoritmo.

Questa è una slide che a cui tengo tengo molto

racconta di una ricerca fatta dall'associazione ProPublica proprio sul sistema COMPAS, quello utilizzato nella condanna Loomis. Dato che il sistema è una black box, ProPublica ha cercato di lavorare sugli output, mettendo a confronto una serie di casi con delle caratteristiche comuni e paragonando un soggetto di pelle bianca a un soggetto di pelle nera. Curiosamente, ne è emerso che il sistema COMPAS attribuiva quasi costantemente al soggetto di pelle nera un rischio di

commissione di nuovi reati particolarmente più alto rispetto al soggetto di pelle bianca. Qui c'è uno schema che riassume il dossier.

Questo indica un Machine Bias su base razziale: non stiamo dicendo che chi ha programmato quel sistema lo abbia fatto con questa intenzione – molto probabilmente è stato utilizzato un set di dati che già incorporava un bias – però la questione, dal punto di vista giuridico, diventa interessantissima. Possiamo aprire questa black box, sì o no? E un ipotetico "caso Loomis" potrebbe verificarsi anche da noi? Abbiamo strumenti che ci mettono al riparo? Questa slide è interessante perché affronta proprio il punto chiave, che è il seguente: la Corte Suprema del Wisconsin – tra le altre cose – ha rilevato come il sistema COMPAS non sia stato decisivo ai fini della condanna. Ha cioè fornito delle indicazioni al giudice, ma poi il giudice ha fatto le sue valutazioni. Ora, il fatto che non sia stato decisivo ci porta esattamente nel nostro perimetro, perché – vedete – noi pensiamo al GDPR come a uno "scudo" universale.

Continua a leggere…

Scarica gratuitamente gli atti del 10° Cyber Crime Conference 2019

https://www.ictsecuritymagazine.com/articoli/decisioni-automatizzate-su-algoritmi-il-quadro-giuridico/