

GIUSTIZIA PREDITTIVA E TUTELA DEI DIRITTI FONDAMENTALI: UN DELICATO EQUILIBRIO TRA ESIGENZE CONTRAPPOSTE

di Letizia Mantovani

(*Dottore di ricerca in Diritto processuale penale,
Università degli studi di Milano*)

Il contributo analizza il potenziale e futuribile ruolo degli algoritmi predittivi nell'amministrazione della giustizia penale, focalizzandosi sulla tensione tra efficienza tecnologica e rispetto dei diritti fondamentali. Nell'ottica di una tutela "multilivello", i rischi di discriminazione derivanti dai *bias* algoritmici e l'opacità generata dall'effetto *black box* si presentano, infatti, come sfide inedite per la tenuta dell'impianto garantistico codificato dalle Carte internazionali. L'indagine si estende alla compatibilità della decisione penale "robotica" con il principio di legalità convenzionale *ex art. 7 CEDU*, mediante l'analisi delle nozioni di "prevedibilità" e "calcolabilità", anche alla luce di un recente studio sull'applicabilità di tecniche di *machine learning* alla giurisprudenza della Corte Edu. In conclusione, si evidenzia la necessità di un assetto normativo, in linea con il recente *AI Act*, che preservi la centralità dell'interprete umano e l'autonomia del giudizio, quali irrinunciabili presidi contro rischiose derive di standardizzazione statistica dell'*ars iudicandi*.

This paper examines the potential and future role of predictive algorithms in criminal justice administration, focusing on the tension between technological efficiency and respect for fundamental rights. Within a "multi-level" protection framework, the risks of discrimination stemming from algorithmic bias and the opacity generated by the "black box" effect represent unprecedented challenges to the integrity of the safeguard system codified by international Charters. The investigation extends to the compatibility of "robotic" criminal decisions with the principle of legality under Article 7 of the ECHR, through an analysis of the concepts of "predictability" and "calculability," also in light of a recent study on applying machine learning techniques to the European Court of Human Rights' case law. In conclusion, the study highlights the need for a regulatory framework - aligned with the recent *AI Act* - that preserves the centrality of the human interpreter and judicial autonomy as essential safeguards against risky trends toward the statistical standardization of the *ars iudicandi*.

Giustizia predittiva; Art. 7 CEDU; Giusto processo

Predictive justice; Art. 7 CEDU; Fair trial

Sommario: 1. Introduzione 2. Previsione algoritmica e giusto processo: profili di (in)compatibilità 3. Il ruolo dell'art. 7 CEDU, tra ragionevole prevedibilità e calcolabilità anticipata 4. La giurisprudenza della Corte europea dei diritti dell'uomo al vaglio della predizione decisoria automatizzata 5. Riflessioni conclusive.

1. Il ricorso agli algoritmi predittivi a supporto degli operatori del diritto, attualmente, è in fase di pieno e sempre più rapido sviluppo, sotto un profilo non esclusivamente tecnologico, ma anche di regolamentazione normativa. L'intelligenza artificiale rappresenta la punta di diamante dell'innovazione digitale, soprattutto alla luce dei considerevoli vantaggi conseguenti al suo impiego, in termini di efficienza, qualità e sicurezza. Sul piano giuridico, alla spinta evolutiva derivante da una rivoluzione di tale portata corrispondono, tuttavia, inevitabili rischi e problemi inediti, allo stato, non ancora del tutto risolti, «anche in ragione del consueto ritardo del diritto rispetto ai fenomeni della tecnologia»¹.

La tendenza umana a riporre inconsciamente fiducia «nelle tecnologie, ritenute oggettive e meritevoli di fiducia per il solo fatto di essere tecnologie»², occulta l'inevitabile influenza dei condizionamenti sociali sul funzionamento e i risultati dell'attività robotica. L'operatività di sistemi basati sull'apprendimento automatico è, infatti, influenzata da interventi esterni, che incidono nella formazione e nella scelta dei *cluster* ove sono raccolti e suddivisi i dati.

Se è vero che l'algoritmo «è ontologicamente condizionato dal sistema di valori e dalle intenzioni di chi ne commissiona la creazione e/o di chi lo crea»³, in presenza di vizi operativi, dovuti a difetti di programmazione del software o a un'errata configurazione del *dataset* su cui si basa l'apprendimento automatico della macchina, inevitabili appaiono gli effetti discriminatori conseguenti all'impiego di tali sistemi di IA in contesti che coinvolgono direttamente i diritti della persona. Ciononostante, è altresì vero che il rischio che il contenuto delle decisioni esclusivamente umane sia il prodotto, anche solo parziale, di *bias* cognitivi è parimenti elevato: sotto questo profilo, i sistemi di intelligenza artificiale non necessariamente offrono un grado di affidabilità inferiore a quello delle persone fisiche.

Posto, dunque, l'inevitabile ingresso di software algoritmici nel contesto giudiziario in un futuro sempre più prossimo, al netto dei potenziali benefici e delle inevitabili criticità, appare altrettanto necessaria la definizione di un assetto normativo idoneo a garantirne l'integrazione nell'ordinamento giuridico di riferimento, nel rispetto degli inviolabili principi che lo governano. Allo scopo, il compito di bilanciare le esigenze di

¹ F. Marasà, *Intelligenza artificiale e tutela dei dati personali. Quali riflessi sulla giustizia predittiva?*, in *ODCC* 2023, 74 ss.

² P. Comoglio, *Prefazione*, in J. Nieva - Fenoll, *Intelligenza artificiale e processo*, Torino 2019, X ss.

³ S. Signorato, *Giustizia penale e intelligenza artificiale. Considerazioni in tema di algoritmo predittivo*, in *RDP* 2020, 605 ss.

sviluppo economico e di efficienza amministrativa, da un lato, e la tutela dei diritti fondamentali, dall'altra, deve essere adempiuto dai competenti organi legislativi nel rispetto delle fonti sovranazionali. Nell'ottica di assicurare una tutela "multilivello"⁴ dei diritti inviolabili direttamente interessati dall'implementazione di tali tecnologie, occorre, infatti, avere contezza della permeabilità del tessuto normativo interno all'ingerenza dei principi fissati dalle Carte internazionali, aventi natura precettiva, e delle disposizioni comunitarie direttamente applicabili⁵.

Occorre preliminarmente soffermarsi sull'evoluzione normativa che ha definito le caratteristiche della *predictive justice* nel panorama internazionale e, ancor prima, sull'interazione delle disposizioni di diritto europeo e delle norme sovranazionali di natura pattizia con l'implementazione di sistemi di intelligenza artificiale nei diversi ambiti della società civile.

I due principali presidi posti dalla comunità internazionale a tutela dei diritti inviolabili della persona sono la Carta dei diritti fondamentali dell'Unione europea, dotata del medesimo valore giuridico dei Trattati istitutivi dell'Unione europea ai sensi dell'art. 6 TUE, e la Convenzione per la salvaguardia dei diritti dell'uomo e delle libertà fondamentali, che rappresenta, per l'ordinamento italiano, parametro interposto per il vaglio di costituzionalità delle norme interne, ai sensi dell'art. 117 Cost.

Il primo nodo da sciogliere è se e come le disposizioni normative contenute in tali Documenti siano idonee a contemperare gli effetti, potenzialmente dirompenti, della rivoluzione tecnologica attualmente in corso. Il preambolo della Carta dei diritti fondamentali dell'Unione europea, del resto, esplicita chiaramente l'obiettivo di «rafforzare la tutela dei diritti fondamentali, alla luce (...) degli sviluppi scientifici e tecnologici, rendendo tali diritti più visibili in una Carta». Analizzando, quindi, i principi enunciati e garantiti dalle Carte internazionali e operando un raffronto con le modalità di funzionamento che accomunano la maggior parte dei sistemi di IA, emerge, anzitutto, la necessità che il trattamento dei dati operato dagli algoritmi avvenga nel pieno rispetto della dignità umana, della libertà individuale, della vita privata e familiare, del diritto a un giusto processo e alla trasparenza dei processi decisionali, nonché del divieto di discriminazione.

⁴ R. G. Conti, *Il sistema di tutela multilivello e l'interazione tra ordinamento interno e fonti sovranazionali*, in *QuestG* 2016, 89 ss.; A. Ruggeri, *Itinerari di una ricerca sul sistema delle fonti*, vol. XIX., *Studi dell'anno 2015*, Torino 2016, 237 ss.

⁵ R. Mastroianni, *Supremazia del diritto dell'Unione e «controlimiti» costituzionali: alcune riflessioni a margine del caso Taricco*, www.penalecontemporaneo.it 7 novembre 2016.

Tale impostazione garantista, tuttavia, mal si concilia con la tendenza degli strumenti predittivi a restituire risultati poco accurati⁶, specialmente per quanto riguarda il supporto allo svolgimento di attività di *law enforcement* e di valutazione sia della tendenza a delinquere sia del rischio di recidiva per determinate categorie di soggetti⁷. La presenza di *bias* nella categorizzazione dei dati personali oggetto di trattamento ha, infatti, delle conseguenze di rilievo sulle previsioni elaborate dalla macchina, le quali, a loro volta, influenzano sensibilmente l'opinione del giudice umano. L'alterazione percettiva del software, determinata da simili errori di programmazione, comporta, pertanto, dei risvolti particolarmente sfavorevoli nei confronti dell'interessato al trattamento, il quale, entrato in contatto con il circuito della giustizia penale, vedrebbe leso il proprio diritto a essere considerato innocente sino a che non intervenga una sentenza irrevocabile che ne accerti l'effettiva responsabilità. Inoltre, a causa dell'effetto "*black box*", l'analisi dei dati e le modalità di individuazione delle correlazioni funzionali al raggiungimento dell'*output* finale non sono conoscibili e, pertanto, anche il percorso logico seguito dalla macchina, la struttura e persino il contenuto dell'attività computazionale rimangono, di fatto, ignoti. Segnatamente, le decisioni assunte all'esito delle valutazioni probabilistiche generate da sistemi algoritmici, sottraendosi anche solo parzialmente al controllo umano, rischiano di rivelarsi del tutto arbitrarie e illegittime⁸.

Il dibattito sull'implementazione di strumenti predittivi nell'amministrazione della giustizia e sulle conseguenze per la tutela dei diritti umani si concentra, in particolare, sul rischio che la macchina possa elaborare *output* ingiustificatamente discriminatori, legati alla razza, alla nazionalità, alle opinioni politiche, all'etnia, al genere, allo stato di salute, ai tratti somatici, o alle convinzioni personali e religiose di coloro che sono sottoposti alla profilazione algoritmica⁹. Un simile esito si verifica, ad esempio, laddove il *dataset* per la fase di *training* dell'algoritmo sia predisposto sulla base di dati, anche solo parzialmente, incompleti o errati: la macchina apprende informazioni non corrette e sviluppa una falsa rappresentazione della realtà che la conduce a scelte

⁶ A. Cowger, *The Threats of Algorithms and AI to Civil Rights, Legal Remedies, and American Jurisprudence: One Nation Under Algorithms*, Lanham 2020, 276 ss.

⁷ M. Caianello, *Potenzialità e rischi derivanti dall'interazione tra I.A. e giustizia penale preventiva*, in *XXVI Lezioni di Diritto dell'Intelligenza artificiale*, a cura di U. Ruffolo, Torino 2021, 206 ss.

⁸ S. Gless, *AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials*, in *GJIL* 2020, 195 ss.

⁹ C. Daelman, *AI through a human rights lens. The role of human rights in fulfilling AI's potential*, in *Artificial Intelligence and the Law*, a cura di J. De Bruyne, C. Vanleenhove, Cambridge 2022, 140 ss.

ingiuste e lesive per i suoi destinatari. Parimenti, effetti di carattere discriminatorio possono discendere da una scorretta attività di classificazione da parte dei programmatori: la selezione di alcune caratteristiche personologiche, a discapito di altre, influenza sensibilmente il processo di profilazione e, di conseguenza, anche le previsioni successivamente elaborate dal sistema di IA¹⁰.

Emblematica, sotto questo aspetto, è l'esperienza della giustizia penale statunitense con COMPAS, software predittivo in grado di valutare il rischio di recidiva dell'imputato sulla base di una scala di valori numerici, compresa tra 1 e 10¹¹. La soluzione prospettata dall'algoritmo, espressa in forma di percentuale, consegue all'elaborazione del contenuto del fascicolo istruttorio e alle risposte fornite dall'imputato in un colloquio valutativo, precipuamente teso all'acquisizione delle informazioni necessarie a implementare la base dati su cui opera il *tool*. Nella fase di processazione, a tali elementi è attribuito un punteggio, sulla base di precisi parametri, precedentemente delineati dagli sviluppatori del software mediante catalogazione dei fattori di rischio statisticamente più frequenti nella casistica dei fenomeni antisociali oggetto di indagine. L'algoritmo, pertanto, effettua una previsione comparando le informazioni ottenute dall'interessato al trattamento con quelle relative a un gruppo di soggetti dalle caratteristiche personologiche affini: oggetto della prognosi è, allora, la generale probabilità che individui con pregressi criminosi tra loro simili possano commettere nuovi reati della medesima specie¹².

Nell'ormai celebre "caso Loomis"¹³, l'imputato era stato sottoposto alla valutazione di COMPAS nel corso del processo a suo carico e, conseguentemente, ritenuto ad alto rischio di recidiva per almeno tre dei capi d'accusa contestatigli. La difesa ha opposto l'inaccessibilità del codice sorgente quale forma di violazione del diritto a un giudizio equo, posto che all'opacità del software corrisponde il concreto pericolo che il modello di apprendimento della macchina sia stato sviluppato sulla base di un *dataset* viziato da pregiudizi e *bias* di carattere discriminatorio¹⁴. La Corte Suprema del Wisconsin ha, tuttavia, respinto le argomentazioni sostenute nella tesi difensiva, ritenendo che la

¹⁰ C. Greco, *Intelligenza artificiale e diritti umani nel diritto internazionale e dell'Unione europea. Alla ricerca di un delicato equilibrio*, in *OIDU* 2022, 782 ss.

¹¹ M. Sciacca, *Algo-crazia e sistema democratico. Alla ricerca di una soluzione antropocentrica. Appunti sulla cd. giustizia predittiva*, in *LDE* 2024, 14 ss.

¹² F. Scamardella – M. Vestoso, *Modelli predittivi a supporto della decisione giudiziaria. Alcuni spunti di riflessione*, *RIFD* 2023, 135 ss.

¹³ Supreme Court of Wisconsin, *State v. Loomis* (2016, 881 N. W.2d749).

¹⁴ L. Lupària, *Prova giudiziaria e ragionamento artificiale: alcune possibili chiavi di lettura*, in J. SALLANTIN – J.J. SZCZECINIARZ (a cura di), *Il concetto di prova alla luce dell'intelligenza artificiale*, Giuffrè 2005, VII e ss.

manca di trasparenza non comporta la violazione del diritto dell'imputato a un giusto processo, a condizione che COMPAS non sia impiegato ai fini della determinazione della pena e che i punteggi di rischio indicati dal software non costituiscano fattori determinanti per il contenuto della decisione, dovendo i giudici fornire una motivazione che vada oltre la mera indicazione del risultato algoritmico¹⁵.

In questo senso, la copertura del brevetto e la tutela del segreto industriale non avrebbero impedito la possibilità di contestare gli esiti del calcolo compiuto dalla macchina: nonostante le modalità operative dello strumento predittivo risultino segrete, il manuale di funzionamento del software indica espressamente che i punteggi sono basati in gran parte su dati statistici, con uso limitato di variabili dinamiche¹⁶. Il problema dell'eventuale presenza di pregiudizi parrebbe, allora, arginato sia dalla coesistenza di un elevato numero di parametri di riferimento, quale garanzia di accuratezza dell'indagine condotta da COMPAS, sia dal divieto per i giudici, *expressis verbis*, di fondare le proprie valutazioni unicamente sui predetti risultati¹⁷. Tale spiegazione non appare, però, sufficientemente esaustiva, poiché non tiene conto dell'inevitabile incidenza di tali risultati, per quanto falsati, sulle sorti del giudizio, oltre che della difficoltà, per chi subisce il pregiudizio discriminatorio, di dimostrare l'esistenza di un "effetto ancoraggio" e le sue conseguenze sul contenuto del provvedimento decisorio.

2. Oltre all'implementazione di *risk assesment tools*, il processo penale si presta altresì al futuribile intervento, in funzione ausiliaria, dell'IA predittiva sia nella valutazione del quadro probatorio, consentendo la ragionevole esclusione di ricostruzioni alternative rispetto a quella che afferma la responsabilità dell'imputato¹⁸, sia nella formazione della decisione. In questo caso, il *dataset* non è creato a partire dalla profilazione di dati personali attinenti a determinate categorie di soggetti, ma sulla base del contenuto delle sentenze presenti nelle banche dati digitali.

La giurisprudenza è il risultato di un processo di elaborazione che si fonda sul ragionamento dei singoli giudici, sull'autorità dei principi normativi e persino su atti

¹⁵ S. Carrer, *Se l'amicus curiae è un algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin*, in *Giurisprudenza Penale Web*, 24 aprile 2019.

¹⁶ A. Simoncini, *Diritto costituzionale e decisioni algoritmiche*, in *Il ragionamento giuridico nell'era dell'intelligenza artificiale*, a cura di S. Dorigo, Pisa 2020, 46 ss.

¹⁷ L. Han-Wei, L. Ching-Fu, C. Yu-Jie, *Beyond State v Loomis: artificial intelligence, government algorithmization and accountability*, in *Int. J. Law Inf. Technol* 2019, 122 ss.

¹⁸ M. Caterini, *Il giudice penale robot*, in www.la-legislazionepenale.eu, 19 dicembre 2020.

di *soft law*, così qualificandosi come forma di “normatività istituzionalizzata” in senso lato¹⁹. L’esercizio del potere giurisdizionale modella e specifica il significato, altrimenti generale e astratto, delle disposizioni di legge, contestualmente arricchendo il patrimonio giurisprudenziale già esistente. Ci si domanda, pertanto, se l’utilizzo di dati d’ingresso apparentemente più neutrali, in quanto espressione della volontà di un soggetto terzo e imparziale, minimizzi il rischio che la macchina elabori prognosi viziate da pregiudizi o, comunque, inadatte a supportare efficacemente le determinazioni del decisore umano.

In primo luogo, bisogna considerare che ogni provvedimento, oltre a riferirsi a un preciso fatto storico, unico e irripetibile, rappresenta il prodotto finale di un ragionamento giuridico inevitabilmente influenzato dalle convinzioni personali e dal contesto di appartenenza del giudice. La sentenza, infatti, si concretizza in una triplice dimensione: istituzionale, ermeneutica e sociale. Questi ultimi due aspetti, in particolare, sono imprescindibilmente legati all’esperienza individuale connaturata all’esercizio dell’*ars iudicandi*, con ripercussioni anche sul contenuto della motivazione, quale momento finale di un articolato *iter* della ragione umana²⁰. L’algoritmo predittivo, elaborando massime e argomentazioni correlate al quesito giuridico oggetto d’indagine, realizza la propria previsione anche sulla base delle opinioni personali espresse dai giudici che hanno deciso in merito a un caso simile, ma non identico, a quello da risolvere. Nonostante l’implementazione della giustizia predittiva sia tesa alla massima formalizzazione del processo interpretativo e decisionale²¹, non si può infatti negare che il soggettivismo sia ontologicamente intrinseco a tutti i provvedimenti giudiziari, risultando, per l’effetto, ineliminabile, posto che «*objectivity is not the goal of the judicial function and subjectivity is not the price of judicial disfunction*»²². La neutralità di tali applicativi, quindi, è soltanto apparente.

Si ricorda, inoltre, come anche l’opacità della processazione algoritmica contribuisca a minare l’affidabilità degli *output*, poiché non consente di ricostruire la logica sottesa alla rilevanza attribuita dalla macchina a determinati dati e alle relative correlazioni. Infine, se il sistema di I.A. individua nessi causali mediante l’analisi di

¹⁹ F. Scamardella – M. Vestoso, op. cit., 148.

²⁰ G. Canzio, *La ‘dike’ degli antichi e la ‘giustizia’ dei moderni: ‘Edipo re’ e ‘Antigone’*, in www.penalecontemporaneo.it, 23 aprile 2018.

²¹ O. Di Giovane, *Il Judge bot e le sequenze giuridiche in materia penale (intelligenza artificiale e stabilizzazione giurisprudenziale)*, in *CP 2020*, 262 ss.

²² D. E. Edlin, *Common Law Judging: Subjectivity, Impartiality, and the Making of Law*, Michigan 2016, 5 ss.

precedenti che presentano caratteristiche comuni, la decisione del software si fonda su mere correlazioni statistiche, ritenute rilevanti poiché ripetibili e, dunque, generalizzabili sulla base del principio “*id quod plerumque accidit*”, spesso a discapito della peculiarità del caso concreto e della situazione personale dell'imputato²³.

Se da un lato, allora, tale tecnologia consente di ottimizzare, in termini quantitativi, l'utilizzo delle risorse contenute nelle banche dati giuridiche, dall'altra rischia, sotto il profilo qualitativo, di assegnare importanza a meri schemi decisionali ripetuti nel tempo e, quindi, di attribuire rilievo a un criterio di stampo statico – seriale, a discapito di uno di tipo dinamico – evolutivo²⁴: gli algoritmi predittivi, infatti, si basano su *patterns* di dati che fotografano la realtà in un dato momento storico, operando un'interpretazione prettamente logico – letterale del dato normativo, senza considerare il contesto in cui lo stesso si inserisce²⁵. L'impiego di tali *tools*, «proiettando la valutazione del singolo caso sullo sfondo di generalizzazioni statistiche, [...] allontana la valutazione dal fatto, e mette in primo piano l'autore, secondo processi di standardizzazione che prospettano una riedizione postmoderna del “tipo criminologico di autore”»²⁶: il loro inserimento nella sequenza procedimentale in cui si articola il giudizio penale, con particolare attenzione alla fase decisoria, deve essere attentamente ponderato e, conseguentemente, regolamentato, pena l'inevitabile violazione delle garanzie a tutela del giusto processo.

Alla luce delle precedenti considerazioni, la futuribile prospettiva di migliorare la qualità del sistema giudiziario mediante l'impiego di «*big data for automated decision – making*»²⁷, si scontra, inevitabilmente, con la necessità di meglio definire il rapporto di tale innovazione tecnologica con la tutela dei diritti fondamentali. Il ricorso, sempre più frequente, all'automazione e al processo decisionale algoritmico potrebbe, infatti, condurre a un'alterazione del concetto stesso di “diritti umani”²⁸ come codificati, in ambito europeo, dalla Carta di Nizza, dai Trattati istitutivi, unitamente ai principi generali non scritti, dal diritto derivato, e, sul versante del diritto internazionale, dalla

²³ S. Quattrocolo, *Artificial Intelligence, Computational Modelling and Criminal Proceedings*, Londra 2020, 192 ss.

²⁴ M. F. Carriero, *Crisi della legalità e intelligenza artificiale: uno sguardo (comparato) d'insieme*, in *AP* 2024, 15 ss.

²⁵ C. Cavaceppi, *L'intelligenza artificiale applicata al diritto penale*, in *Intelligenza artificiale - Algoritmi giuridici: Ius condendum o fantadiritto?*, a cura di G. Taddei Elmi, A. Contaldo, Pisa 2020, 113 ss.

²⁶ V. Manes, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in *Intelligenza artificiale. Il diritto, i diritti, l'etica*, a cura di U. Ruffolo, Milano 2020, 547 ss.

²⁷ European Union Agency For Fundamental Rights (FRA), *#BigData: Discrimination in data – supported decision making* 2018, 3.

²⁸ C. Greco, op. cit., 806.

CEDU.

Tale prospettiva appare tutt'altro che utopistica, soprattutto se si considera che i diritti tutelati, tanto dalla Carta di Nizza quanto dalla CEDU, non rivestono carattere assoluto. Secondo l'articolo 52 paragrafo 1 CDFUE, infatti, una limitazione dei diritti fondamentali è possibile se prevista dalla legge, purché effettivamente rispondente a obiettivi di interesse generale riconosciuti dall'Unione europea e sempre nel rispetto dei criteri di necessità e proporzionalità: la compressione dei diritti, oltre a dover risultare giustificata da esigenze imperative, dovrà anche essere subordinata alla possibilità di poter soddisfare tali istanze mediante soluzioni alternative e meno invasive. Nondimeno, la Corte di giustizia dell'UE ha ribadito che garanzie fondamentali come quelle sancite dagli artt. 7 e 8 CDFUE possono essere contemperate dalla tutela di interessi diversi ma altrettanto rilevanti, a condizione che tale compressione sia tesa a perseguire finalità di interesse generale, riconosciute dall'UE o dall'esigenza di proteggere i diritti e le libertà altrui.

Parimenti, anche la Corte EDU, proprio con riferimento all'impiego di nuove tecnologie, ha avuto modo di chiarire che gli Stati dovrebbero trovare un giusto equilibrio tra la protezione dei diritti fondamentali e la promozione del progresso tecnico – scientifico. In questa prospettiva, l'applicazione su vasta scala di sistemi di intelligenza artificiale implica un'inevitabile connessione, e una possibile reciproca interferenza, con un'ampia gamma di diritti fondamentali, dall'intensità variabile, a seconda delle specifiche funzionalità di tali strumenti, della loro complessità e del settore in cui vengono impiegati²⁹. Peraltro, il continuo mutamento e la conseguente imprevedibilità di questi sistemi rappresentano degli ostacoli concreti³⁰ alla preventiva individuazione dei possibili profili di ingerenza che potrebbero concretamente realizzarsi con l'assetto di garanzie previsto dalle Carte internazionali³¹. Con particolare riguardo ai software di giustizia predittiva, corrispondendo all'implementazione di modelli basati sull'autoapprendimento una sensibile riduzione

²⁹ Si veda la posizione del Consiglio d'Europa espressa nella *Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems*, Appendix, par. A.8.

³⁰ A. Adinolfi, *L'unione europea dinanzi allo sviluppo dell'intelligenza artificiale: la costruzione di uno schema di regolamentazione europeo*, in *Il ragionamento giuridico nell'era dell'intelligenza artificiale*, a cura di S. Dorigo, Pisa 2020, 13 ss.

³¹ MSI-AUT, *A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework*, 9 novembre 2018. Lo studio si riferisce alla rivoluzione digitale in atto come alla "New" Industrial Revolution ed evidenzia che «*The 'New' Industrial Revolution now dawning is also likely to bring myriad benefits to individuals and societies yet, like the original Industrial Revolution, might also generate unintended adverse effects that were not recognised at the time of the revolution's unfolding*».

della supervisione umana, tutt'altro che marginale è il rischio che possano verificarsi, nel corso dell'*iter* decisorio, violazioni della dignità umana o della privacy, ovvero esiti di natura discriminatoria, quali dirette conseguenze dell'intervento della macchina a supporto dell'attività del giudice. In assenza di presidi adeguati potrebbero, inoltre, essere compromessi i diritti a un rimedio giurisdizionale efficace, alla buona amministrazione, a un ricorso effettivo e ad un processo equo³².

Attestata la necessità di interventi di carattere strettamente normativo tesi alla definizione e alla regolamentazione, nel tempo, dei principali punti di contatto tra IA e amministrazione della giustizia, si sottolinea l'opportunità che tale base giuridica, frutto degli sforzi congiunti delle istituzioni sovranazionali e dei singoli ordinamenti statali, valorizzi l'importanza di principi quali la trasparenza e il giusto processo, la certezza del diritto e, nello specifico, predisponga misure concrete, volte a tutelare le libertà fondamentali e le garanzie associate al ricorso, al trattamento e all'analisi dei dati da parte del settore pubblico e di quello privato.

D'altro canto, più che concetti giuridici inediti, dall'interazione tra innovazione tecnologica e diritti fondamentali sembrano emergere nuove declinazioni di significato di diritti già esistenti. Ne rappresenta un chiaro esempio il concetto di "dignità umana", di cui all'art. 1 CDFUE, il quale potrebbe, *ex se*, ben ricomprendere anche il diritto della persona a non dover dipendere esclusivamente dalle valutazioni e dalle prognosi operate della macchina. La pretesa dei singoli individui a non essere «*measured, analysed or coached*»³³, sintetizzabile, secondo parte della dottrina, nel "*right to cognitive liberty*"³⁴, rientrerebbe, allora, a pieno titolo tra i diritti già sanciti e garantiti dalle Carte internazionali³⁵. Controversa appare, pertanto, la possibilità che alla formalizzazione di nuove prerogative individuali, discendenti direttamente dall'ingresso dell'IA nelle dinamiche su cui si fonda uno Stato di diritto, corrisponda una loro più intensa ed efficace protezione. Ampliando, sotto un profilo puramente nominale, il catalogo delle situazioni giuridiche soggettive tutelabili si prospetta, al

³² C. Greco, *op. cit.*, 807.

³³ Rathenau Instituut, *Human rights in the robot age Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality* 2017, 6 ss.

³⁴ J. Bublitz, *My mind is mine!? Cognitive liberty as a legal concept*, in *Cognitive Enhancement: An Interdisciplinary Perspective*, a cura di E. Hildt - A. Franke, Londra 2013, 233 ss.

³⁵ F. Siracusano, *Tutela multilivello dei diritti fondamentali e diritto penale. Sviluppi e questioni aperte nel "dialogo" tra le Corti*, in *Tra diritti fondamentali e principi generali della materia penale. La crescente influenza della giurisprudenza delle Corti europee sull'ordinamento penale italiano*, a cura di G. Grasso, A. M. Maugeri, R. Sicurella, Pisa 2020, 3 ss.

contrario, il rischio di estendere ulteriormente un già ricco catalogo di diritti riconosciuti solo in potenza, senza riscontri nella prassi³⁶.

Occorre, allora, rispondere all'iniziale quesito sull'idoneità delle disposizioni normative contenute nelle Carte internazionali a far fronte a tali cambiamenti. La risposta, alla luce di quanto finora osservato, appare affermativa, sebbene non preclusiva della necessità di individuare specifici accorgimenti, quali: l'introduzione di meccanismi di controllo rigorosi ed effettivi che impongano ai produttori dei sistemi di IA di conformarsi ad una serie di requisiti finalizzati a rafforzare la tutela nei confronti degli utilizzatori e, in particolare, dei loro diritti personalissimi, nonché la previsione di efficaci meccanismi sanzionatori atti a contrastare un uso distorto di tali tecnologie.

Sul punto, di particolare interesse è il rapporto annuale dell'Alto Commissario delle Nazioni Unite per i Diritti Umani del 2021, nel quale viene chiaramente evidenziato l'impatto dei sistemi di IA sulla tutela della privacy e sugli altri diritti umani strettamente connessi a settori di interesse pubblico, compresa la giustizia penale. Nel documento sono formulate, a beneficio sia dei legislatori nazionali sia delle imprese produttrici, una serie di raccomandazioni finalizzate ad affrontare le sfide del processo tecnologico, così da *«ensure that the use of AI is in compliance with all human rights and that any interference with the right to privacy and other human rights through the use of AI is provided for by law, pursues a legitimate aim, complies with the principles of necessity and proportionality and does not impair the essence of the rights in question»*³⁷.

In questo senso, si è orientato anche il Legislatore europeo che, nel corso dell'ultimo decennio, ha tentato a più riprese di definire i confini applicativi di un modello di intelligenza artificiale giuridicamente ed eticamente accettabile, in ultimo consacrato dall'adozione del Regolamento (UE) del Parlamento europeo e del Consiglio 2024/1689 (AI Act).

3. L'articolo 7 CEDU sancisce una garanzia irrinunciabile di ogni Stato di diritto, mai derogabile, nemmeno nelle ipotesi di eccezionale gravità individuate dall'art. 15 della medesima Convenzione. Il divieto assoluto di applicazione retroattiva *in peius*

³⁶ A. Santosuosso, *Intelligenza artificiale e diritto, Perché le tecnologie di IA sono una grande opportunità per il diritto*, Milano 2021, 413 ss.

³⁷ A/HRC/48/31, *The right to privacy in the digital age – Report of the United Nations High Commissioner for Human Rights*, 13 settembre 2021, 15.

della norma penale, a prescindere dalla natura interna o sovranazionale della fonte incriminatrice, costituisce, infatti, il fulcro del principio di legalità convenzionale³⁸. Diversamente dall'impostazione tradizionalmente riconducibile al sistema penale italiano³⁹, tuttavia, la CEDU non prioritizza la dimensione istituzionale della riserva di legge, quale granitico presidio della rappresentatività parlamentare, bensì, riconosce primaria importanza alla libertà di autodeterminazione del singolo. Riconducendo direttamente l'irretroattività della norma penale sfavorevole al principio di legalità, l'art. 7 CEDU, infatti, garantisce all'agente libere scelte di azione sulla base della preconnoscenza di ciò che è lecito e di ciò che non lo è⁴⁰. Si configura, così, un vero e proprio diritto alla "calcolabilità anticipata"⁴¹, da intendersi come garanzia fondamentale di ogni individuo, riconosciuta e tutelata tanto dalle Carte internazionali quanto dalla giurisprudenza europea⁴².

In veste di principali corollari del divieto assoluto di retroattività *in peius*, i criteri di ragionevole prevedibilità e di accessibilità costituiscono condizioni qualitative della norma incriminatrice, relative agli elementi essenziali della fattispecie antiggiuridica, al contenuto e alla portata del trattamento sanzionatorio, nonché alla loro effettiva conoscibilità da parte dei consociati. Mediante l'adozione di tali canoni ermeneutici è, dunque, demandato alla Corte EDU il delicato, quanto cruciale, compito di perseguire e tutelare la certezza del diritto penale.

Un orientamento giurisprudenziale consolidato⁴³ ha postulato il soddisfacimento del principio di prevedibilità in presenza di una base legale il più possibile chiara e completa, atta a formalizzare la pretesa punitiva nei confronti di una determinata condotta illecita, alla quale deve far seguito una prassi applicativa costante e coerente. Tali fattori, solitamente, operano in un rapporto di proporzionalità inversa, poiché, a una maggiore indeterminatezza della norma incriminatrice corrisponde una più pregnante indagine sulla sua idoneità a disciplinare fattispecie concrete, al fine di sanare, in sede giudiziale, eventuali vizi di imprecisione presenti nel testo della

³⁸ C. Sotis, "Ragionevoli prevedibilità" e giurisprudenza della Corte Edu, in *QuestG* 2018, 68 ss.

³⁹ G. Marinucci, E. Dolcini, G. L. Gatta, *Manuale di diritto penale. Parte generale*, Milano 2024, 43.

⁴⁰ C. Sotis, op. cit., 69.

⁴¹ M. Donini, *Prescrizione e irretroattività tra diritto e procedura penale*. "D", in *FI* 1998, 324.

⁴² V. Manes, *Art. 7*, in *Commentario breve alla Convenzione europea dei diritti dell'uomo*, a cura di S. Bartole, De Sena, V. Zagrebelsky, Padova 2012, 279 ss.

⁴³ *Ex multis*, C. eur., 24 maggio 1988, *Muller e altri c. Svizzera*; C. eur., 25 maggio 1993, *Kokkinakis c. Grecia*.

disposizione⁴⁴. Secondo la Corte EDU, tuttavia, la precisione assoluta del dettato normativo non solo non è possibile, ma nemmeno auspicabile, vista la necessità per il diritto penale di adeguarsi tempestivamente ai mutamenti sociali. Preminente rilevanza assume, allora, l'esercizio dell'*ars iudicandi*, che, specificando la portata del precetto, eleva l'attività giurisprudenziale al rango di fonte di produzione. In questo modo «la norma incriminatrice diviene una fattispecie a formazione progressiva, in cui al testo di legge si chiede fundamentalmente di esprimere delle linee di indirizzo in grado di orientare l'interprete in modo chiaro e coerente, e a quest'ultimo di applicarle in modo ragionevole. Non solo testo, quindi, ma nemmeno solo interpretazione»⁴⁵.

Qualora la giurisprudenza preesistente non appaia coerente, ovvero risulti del tutto insussistente, il vaglio di prevedibilità della norma si plasma a partire dai suoi destinatari, nonché dalla percezione sociale del suo disvalore, fino ad arrivare, in alcuni casi, a fondare una «misura ipersoggettiva della prevedibilità»⁴⁶ modellata sulle caratteristiche personali del singolo ricorrente. Da ultimo, se della disposizione oggetto del controllo non risulta un'applicazione costante e omogenea, la Corte ha talvolta fatto ricorso al criterio della “conformità alla sostanza del reato”⁴⁷, mediante il quale la prevedibilità viene valutata a seguito di un raffronto proporzionale tra il disvalore espresso dal nucleo essenziale della fattispecie delittuosa e quello ravvisabile nella situazione concreta, oggetto di giudizio⁴⁸: la norma penale è da considerarsi ragionevolmente prevedibile solo se l'effetto criminogeno attribuito al tipo criminoso legale si rispecchia omogeneamente nel fatto storico *de qua*.

Al riconoscimento, da parte della Corte EDU, dell'idoneità del formante giurisprudenziale ad assicurare la ragionevole prevedibilità convenzionale, corrisponde l'assegnazione al giudice del compito di codeterminare il contenuto della norma penale, con la conseguenza che il principio di tipicità ricomprenda al suo interno non solo l'evidenza del dato normativo, ma anche il risultato dell'attività interpretativa⁴⁹. Nell'ottica, dunque, di una tecnica di penalizzazione “mista”, nella

⁴⁴ O. Di Giovine, *L'interpretazione in diritto penale*, Milano 2006; M. Donini, *Europeismo giudiziario e scienza penale. Dalla dogmatica classica alla giurisprudenza – fonte*, Milano 2011.

⁴⁵ C. Sotis, op. cit., 69.

⁴⁶ *Ibidem*.

⁴⁷ M. Pifferi, *Generalia Delictorum. Il Tractatus criminalis di Tiberio Deciani e la “parte generale” del diritto penale*, Milano 2006, 128 ss.

⁴⁸ F. Palazzo, *Legalità e determinatezza della legge penale: significato linguistico, interpretazione e conoscibilità della regula iuris*, in *Diritto penale e giurisprudenza costituzionale*, a cura di G. Vassalli, Napoli 2006, 75.

⁴⁹ W. Hassemer, *Fattispecie e tipo*, Napoli 2007 (1968), 183.

quale la legge esprime concetti, che, per essere pienamente compresi dai consociati, necessitano di essere sottoposti al vaglio ermeneutico dei giudicanti, il valore del precedente giurisprudenziale assume un significato decisamente più pregnante ed evocativo.

In tale contesto, l'avvento dell'intelligenza artificiale rappresenta un interessante banco di prova per la tenuta del principio di prevedibilità delle decisioni giudiziarie ex art. 7 CEDU, inteso come «diritto alla possibilità di prevedere, in capo ai consociati, le decisioni giurisdizionali in materia penale, e cioè gli esiti della vicenda processuale quale immediata e concreta applicazione della legge»⁵⁰. Offrendo al giudice la possibilità di disporre, in tempi brevi, di una circostanziata e accurata casistica relativa alla norma generale, il ricorso ad algoritmi predittivi nella fase decisoria semplificherebbe sensibilmente l'attività di sussunzione della *res iudicanda* nel perimetro della disposizione interpretata⁵¹. A conclusione di tale procedimento, che combina l'azione umana e le capacità di calcolo di un agente robotico, si otterrebbe, quindi, «una sorta di massima individualizzata, risultato della sintesi tra il precedente e il caso cogente»⁵².

4. Allo scopo di verificare la realizzabilità dello scenario appena descritto, un gruppo di ricercatori ha sviluppato un applicativo di IA in grado di prevedere, o per meglio dire, giungere alle stesse conclusioni della Corte europea dei diritti dell'uomo, con un grado di precisione prossimo all'80%. Muovendo dal presupposto che «*recent advances in Natural Language Processing and Machine Learning provide us with the tools to build predictive models that can be used to unveil patterns driving judicial decisions*»⁵³, il software è stato programmato per individuare e catalogare i modelli decisionali seguiti dai giudici della Corte EDU. All'attività di classificazione è seguita l'adozione di uno schema operativo binario, nel quale il contenuto dei documenti processuali costituiva la base dati elaborata dall'algoritmo, mentre l'*output* corrispondeva, alternativamente, alla verifica o alla confutazione dell'inosservanza di una norma CEDU nel caso oggetto di giudizio. Tale analisi empirica, tesa a dimostrare come le informazioni cristallizzate negli atti di un singolo processo rappresentino indicatori significativi per una

⁵⁰ L. Rocca – A. Gaito, *Il "controlimite" della tutela dei diritti processuali dell'imputato: visioni evolutive dalle Corti europee tra legalità e prevedibilità*, in AP 2019, 8 ss.

⁵¹ E. Vincenti, *Massimazione e conoscenza della giurisprudenza nell'era digitale*, in QuestG 2018, 147 ss.

⁵² F. Falato, *L'inferenza generata dai sistemi esperti e dalle reti neurali nella logica giudiziale*, in AP 2020, 29.

⁵³ N. Aletras, D. Tsarapatsanis, D. Preotiuc-Pietro, V. Lampos, *Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective*, in PeerJ Comput. Sci. 2016, 1.

valutazione in chiave prognostica, si è allineata all'impostazione "realista", particolarmente cara ai sistemi di *common law*, che considera la tipizzazione interpretativa dei fatti eziologicamente rilevanti un presupposto necessario per la corretta applicazione della norma incriminatrice⁵⁴.

Sempre secondo gli Autori, anche il contesto in cui si realizza la fattispecie concreta, incidendo sensibilmente sull'esito della controversia, influenza la previsione algoritmica e consente, al contempo, una disamina qualitativa delle possibili relazioni che intercorrono tra «*stimulus of the facts*» e «*judicial decision – making*»⁵⁵. Allo scopo, occorre presupporre aprioristicamente la concreta attuabilità di un "approccio testuale", con la conseguenza che «*the text extracted from published judgments of the Court bears a sufficient number of similarities with, and can therefore stand as a (crude) proxy for, applications lodged with the Court as well as for briefs submitted by parties in pending cases*».

Il modello utilizzato dagli sviluppatori si basa su una funzione lineare *kernel* che, facilitando l'interpretazione dei dati mediante la loro catalogazione in categorie "multidimensionali"⁵⁶, consente all'algoritmo di dedurre, con un elevato grado di precisione, il probabile esito di una questione giuridica a partire dall'analisi testuale degli atti processuali⁵⁷. Coerentemente con le premesse metodologiche appena illustrate, la ricostruzione del fatto storico si è, dunque, dimostrata cruciale per l'attività predittiva realizzata dal software, in quanto «*the topical content of a case is an important indicator whether there is a violation of a given Article of the Convention or not*»⁵⁸.

Il potenziale innovativo di un simile approccio sistematico appare di tutta evidenza

⁵⁴ R. BLAIOTTA, *Il realismo di Karl Popper: un ideale di conoscenza oggettiva per il giudizio penale*, in CP, 1997, 3689 ss.

⁵⁵ N. Aletras, D. Tsarapatsanis, D. Preotiuc-Pietro, V. Lampos, op. cit., 1.

⁵⁶ I "metodi kernel" costituiscono una classe di algoritmi per l'analisi di schemi e si basano su funzioni lineari operanti in uno spazio multidimensionale, in cui a ogni coordinata corrisponde una specifica caratteristica dei dati analizzati: l'attività di mappatura avviene mediante l'individuazione di correlazioni generali tra i singoli dati, senza che siano calcolate le coordinate di questi ultimi nello spazio - funzione. M. S. Gerber, *Predicting crime using Twitter and kernel density estimation*, in DSS 2014, 115 ss.

⁵⁷ Nel caso di specie è stato impiegato un algoritmo SVM (*Support Vector Machine*), le cui capacità di apprendimento automatico si sono dimostrate particolarmente efficaci in attività di classificazione testuale, soprattutto se impiegato su *database* di piccole dimensioni. S. Wang, C.D. Mannin, *Baselines and bigrams: simple, good sentiment and topic classification*, in *Proceedings of the 50th annual meeting of the Association for Computational Linguistics: short papers*, a cura di L. Haizhou, L. Chin-Yew, M. Osborne, G. Geunbae Lee, J. C. Park, Jeju Island 2012, 90 ss.

⁵⁸ N. Aletras, D. Tsarapatsanis, D. Preotiuc-Pietro, V. Lampos, op. cit., 2.

allorché se ne vagliano alcune implicazioni pratiche. L'analisi testuale orientata alla previsione decisoria consentirebbe non solo di valutare la probabile incidenza dei singoli fattori causali, individuati e descritti nelle carte processuali, sull'esito delle singole controversie, ma anche di dedurre dai precedenti giurisprudenziali una regola di giudizio potenzialmente applicabile al caso concreto, a condizione che ci sia «*enough similarity between (at least) certain chunks of the text of published judgments and applications lodged with the Court and/or briefs submitted by parties with respect to pending cases*». La formulazione di previsioni accurate mediante processazione algoritmica del linguaggio naturale non è, tuttavia, un'operazione priva di complicazioni e richiede, pertanto il ricorso a particolari cautele.

Un primo *caveat* degli Autori riguarda la ricostruzione del fatto storico e delle relative dinamiche processuali realizzata dalla Corte EDU, non sempre qualificabile «*as a neutral mirroring of the factual background of the case*» e, pertanto, potenzialmente viziata da pregiudizi o errori di valutazione. Ciononostante, il rischio di un'eventuale alterazione della realtà fattuale si ritiene mitigato, in primo luogo, dal limitato potere di accertamento fattuale riservato alla CEDU: la Corte, infatti, riassume il contesto in cui si è verificata la presunta violazione basandosi sul contenuto dell'istanza presentata dal ricorrente, nonché sulla ricostruzione emersa dal procedimento promosso avanti al giudice nazionale, specie se confermata in più gradi di giudizio⁵⁹. Si consideri, inoltre, che oggetto di contestazione delle parti non sono quasi mai i fatti, quanto piuttosto la loro scorretta qualificazione giuridica. Secondo gli sviluppatori del progetto, pertanto, tali considerazioni suggeriscono che, in assenza di altri dati testuali, la sezione della sentenza dedicata al fatto può ragionevolmente svolgere la funzione di «*proxy for a textual representation of the factual background of a case*»⁶⁰.

Un secondo accorgimento riguarda, poi, la base giurisprudenziale su cui opera il modello predittivo. Il raggiungimento di un risultato accurato è, infatti, subordinato alla preesistenza di una nutrita casistica di riferimento: maggiore è il numero di sentenze relative a un determinato articolo della Convenzione, più accurata risulta la previsione algoritmica in merito alla probabile violazione di detta norma nel caso concreto. A dimostrazione di tale assunto, il contenuto del *dataset* creato dagli

⁵⁹ P. Leach, C. Paraskeva, G. Uelac, *Human rights fact-finding. The European Court of Human Rights at a crossroads*, in *NQHR* 2010, 41 ss.

⁶⁰ N. Aletras, D. Tsarapatsanis, D. Preotiuc-Pietro, V. Lampos, op. cit., 5.

sviluppatori è stato limitato alle decisioni della Corte di Strasburgo relative agli artt. 3, 6, e 8 CEDU, per un totale di 584 sentenze: in questo modo, l'efficacia dell'algoritmo è stata testata con riferimento a una base informativa specifica, idonea a fornire alla macchina gli strumenti necessari per rispondere con precisione al quesito sottoposto.

Segnatamente, per ognuno degli articoli della Convenzione ivi considerati è stata, in primo luogo, individuata la relativa casistica nella banca – dati della Corte europea dei diritti dell'uomo, limitatamente alle controversie trattate in lingua inglese. Dopodiché, sono stati selezionati, in egual misura, casi di violazione e di non violazione delle norme di riferimento: al fine di garantire una base giurisprudenziale sufficientemente bilanciata, la norma meno commentata dalla Corte ha proporzionalmente orientato, sotto il profilo quantitativo, la scelta delle cause aventi ad oggetto le altre due disposizioni CEDU. Dalle singole sezioni che compongono ogni sentenza è stato, poi, estratto il testo rilevante per l'elaborazione del linguaggio naturale⁶¹, escludendo le statuizioni processuali strettamente connesse all'esito del singolo giudizio, così da non influenzare eccessivamente la valutazione prognostica successivamente operata dall'algoritmo. Alle caratteristiche più rilevanti e ricorrenti, individuate nel corso dell'analisi del dato testuale precedentemente acquisito, sono stati, infine, assegnati dei valori: positivo qualora risultasse evidente il rispetto della normativa CEDU (+1), negativo nel caso di violazione (- 1).

I risultati ottenuti a seguito dell'elaborazione algoritmica del linguaggio naturale hanno, infine, evidenziato la capacità del software di pervenire a conclusioni molto simili a quelle indicate nei dispositivi delle sentenze analizzate. In particolare, i dati contenuti nella sezione dei provvedimenti denominata "*Circumstances*" si sono dimostrati determinanti per l'accuratezza delle previsioni aventi ad oggetto decisioni sugli articoli 6 e 8 CEDU, con punteggi rispettivamente di 0,82 e 0,77. Riguardo all'articolo 3 CEDU, invece, la massima accuratezza predittiva (0,70) è stata ottenuta utilizzando il testo estratto dall'intera sentenza.

All'analisi della sezione "*Law*", relativa alle argomentazioni sostenute dalle parti e alle motivazioni fornite dalla Corte in punto di diritto, sono conseguite performance predittive caratterizzate da un grado di precisione nettamente inferiore (0,62): ragione

⁶¹ Si è fatto ricorso, in particolare, a modelli statistici per la linguistica e la modellazione del linguaggio naturale. D. Turney-Pantel, *From frequency to meaning: vector space models of semantics*, in *JAIR* 2010, 141 ss.

principale di tale risultato è che un numero elevato di casi non include un paragrafo esclusivamente dedicato alle questioni giuridiche ad essi sottese, specie con riferimento alle decisioni della Corte che dichiarano l'inammissibilità dei ricorsi presentategli.

In conclusione, gli sviluppatori hanno osservato che sia il tipo di linguaggio utilizzato nei documenti elaborati dal software sia il contesto fattuale ivi descritto risultano indicatori di interesse, idonei a costituire *proxy* affidabili per le future decisioni giudiziarie aventi ad oggetto casi simili. Se da un lato, infatti, «*the consistently more robust predictive accuracy of the 'Circumstances' subsection suggests a strong correlation between the facts of a case, as these are formulated by the Court in this subsection, and the decisions made by judges*», dall'altro, le considerazioni espresse dalla Corte in punto di diritto paiono incidere in modo meno significativo sull'esito della prognosi algoritmica. Tale risultato trova, allora, giustificazione se si considera che «*the relatively lower predictive accuracy of the 'Law' subsection could also be an indicator of the fact that legal reasons and arguments of a case have a weaker correlation with decisions made by the Court. However, this last remark should be seriously mitigated since, as we have already observed, many inadmissibility cases do not contain a separate 'Law' subsection*»⁶².

A parere degli Autori, il risultato ottenuto induce una riflessione su come i giudici, nel processo decisionale, appaiano «*primarily responsive to non – legal, rather than to legal, reasons when they decide appellate cases*»⁶³. Siffatta considerazione sottende una critica, sebbene appena accennata, all'impostazione formalista, ove il diritto è inteso come un sistema coerente e completo di norme precostituite rispetto all'attività degli interpreti e della giurisdizione⁶⁴. Preferibile appare, dunque, il modello prospettato dal realismo giuridico americano, che sposta il baricentro dell'analisi teorica dalla formulazione delle norme all'attività interpretativa, nonché alle ideologie espresse dalle operazioni ermeneutiche⁶⁵.

La ricerca empirica condotta in anni recenti sull'*iter* decisorio seguito nelle Corti

⁶² N. Aletras, D. Tsarapatsanis, D. Preotiuc-Pietro, V. Lampos, op. cit., 11.

⁶³ *Ibidem*.

⁶⁴ R. Guastini, *Giovanni Tarello giurista. Introduzione*, in, L'opera di Giovanni Tarello nella cultura giuridica contemporanea, a cura di S. Castignone, Bologna 1989, 123; R. Guastini, *Saggi scettici sull'interpretazione*, Torino 2017, 1 ss.

⁶⁵ L. Malagoli, *Una costruttiva sfiducia. La disciplina costituzionale della proprietà e la lezione metodologica del realismo giuridico americano*, in *MSCG* 2024, 31 ss.

internazionali e nella *US Supreme Court*, infatti, ha evidenziato la frequenza con cui «*pure legal models, especially deductive ones, are false as an empirical matter when it comes to cases decided by courts further up the hierarchy. As a result, it is suggested that the best way to explain past decisions of such courts and to predict future ones is by placing emphasis on other kinds of empirical variables that affect judges*»⁶⁶. Coerentemente, anche nei giudizi prognostici sulle decisioni della CEDU l'influenza esercitata da fattori extra – giuridici si è dimostrata prevalente rispetto al contenuto, *strictu sensu*, precettivo del dato normativo. In particolare, qualora si assuma che il paragrafo delle sentenze analizzate dall'algoritmo denominato “*Circumstances*” costituisca un *proxy*, seppur approssimativo, delle caratteristiche essenziali del fatto storico e che la sottosezione “*Law*” rappresenti un archetipo delle argomentazioni giuridiche più rilevanti, è allora di tutta evidenza la maggiore efficacia predittiva della prima rispetto alla seconda. Occorre però precisare che tali dati dovrebbero essere letti anche alla luce del cosiddetto “effetto di selezione”⁶⁷, in base al quale i ricorsi presentati alla Corte, spesso privi dei requisiti di ricevibilità necessari per essere decisi da Camera e Grande Camera, non costituiscono necessariamente un campione rappresentativo di tutte le potenziali violazioni della Convenzione. Secondo gli Autori sarebbero, quindi, necessarie ulteriori analisi testuali al fine di verificare se i risultati ottenuti siano generalizzabili e, conseguentemente, replicabili con riferimento all'intero articolato della Carta internazionale. Solo successivamente a tale verifica si potrebbe valutare l'estensione del modello predittivo ad altri contesti giurisdizionali, in veste di possibile supporto per i giudici nazionali chiamati a decidere su questioni analoghe a quelle fondanti i ricorsi decisi dalla stessa CEDU.

5. Al netto delle preannunciate criticità e in attesa dei relativi correttivi, specie sotto il profilo dell'accessibilità informativa, le prospettive di una giustizia attuariale e automatizzata appaiono, allo stato, sempre meno utopistiche. In particolare, se contestualizzato nell'applicazione del diritto penale, il ricorso, per quanto non esclusivo, a uno strumento di predizione decisoria dall'elevato grado di affidabilità consentirebbe una rapida individuazione dei *pattern* decisori più frequenti, in relazione a una specifica norma, efficacemente garantendo il diritto alla prevedibilità,

⁶⁶ N. Aletras, D. Tsarapatsanis, D. Preotiuc-Pietro, V. Lampos, op. cit., 12; L. Baum, *The puzzle of judicial behavior*, Michigan 2009; F. Schauer, *Prediction and particularity*, Boston 1998, 773 ss.; J.A. Segal, H.J. Spaeth, *The Supreme Court and the attitudinal model revisited*, Cambridge 2002.

⁶⁷ G.L. Priest, B. Klein, *The selection of disputes for litigation*, in *JLS* 1984.

intesa come calcolabilità della portata del precetto sanzionatorio, sancito dall'art. 7 CEDU.

Sotto un altro profilo occorre, tuttavia, sottolineare come, ancora una volta, l'applicativo di intelligenza artificiale non sia stato in grado di replicare il ragionamento giuridico seguito dai giudici europei o di prevedere l'esito di una causa muovendo da una base informativa limitata, quale, ad esempio, il contenuto del solo ricorso promosso dalla parte che lamenta la lesione di un suo diritto fondamentale. L'algoritmo ha, infatti, realizzato un'inferenza statistica, basata sull'individuazione di gruppi lessicali ricorrenti, preventivamente identificati allo scopo di accertare la presenza o meno di una violazione della normativa CEDU in termini generali. Limite intrinseco a questo tipo di operazione è, come già rilevato, l'impossibilità di pervenire a una prognosi che contempra anche la valutazione soggettiva, e segnatamente discrezionale, propria del singolo decisore umano.

A tale assunto consegue, quale corollario, che il ricorso a strumenti di giustizia predittiva così configurati dissuade il giudice dal ricercare la regola di diritto applicabile alla fattispecie concreta, inducendolo, di converso, a verificare che la soluzione suggerita come più probabile dall'algoritmo corrisponda a quella oggetto del corrente giudizio: in questo modo l'approccio alla *res iudicanda* si allinea, inevitabilmente, al sistema di *common law*, rimarcando l'autorità del precedente giurisprudenziale, nonché la sua vincolatività. Se, dunque, l'automatizzazione del principio dello *stare decisis* appare confacente all'impianto sistematico del diritto anglosassone – cui si riferiscono le brevi considerazioni esposte dagli sviluppatori del software – e, pur nella sfumatura meno intensa del “precedente persuasivo”⁶⁸, applicabile anche all'assetto normativo e giurisprudenziale delineato dalla CEDU, le stesse considerazioni non valgono necessariamente per gli ordinamenti di *civil law*⁶⁹. Per certi versi, l'avvicinamento al modello angloamericano, promuovendo l'uniformità interpretativa della legge, appare propedeutico alla formazione di un più democratico “diritto isometrico”⁷⁰, in cui «*l'intégralité des décisions de justice est mesurée de façon égale par un programme informatique, dont la synthèse finit par devenir la norme*»⁷¹, con

⁶⁸ M.G. Civinini, *Il valore del precedente nella giurisprudenza della Corte europea dei diritti dell'uomo*, in *QuestG* 2018, 102 ss.

⁶⁹ S. Chiarloni, *Ruolo della giurisprudenza e attività creative di nuovo diritto*, in *RTrimDPC* 2002,1 ss.

⁷⁰ P. Comoglio, *Nuove tecnologie e disponibilità della prova*, Torino 2018, 359.

⁷¹ D. Cholet, *La justice prédictive et les principes fondamentaux du procès civil*, in *La justice prédictive*, AA.VV., Parigi 2018, 234.

il conseguente, per quanto implicito, riconoscimento alla programmazione della qualifica di nuova fonte di produzione normativa.

D'altro canto, i rischi paventati da tale approccio appaiono di non scarso rilievo. In particolare, occorre precisare che l'utilizzo dell'intelligenza artificiale predittiva consente di codificare il passato sulla base di schemi e *set* valoriali predefiniti, cristallizzando i trascorsi storici nel processo prognostico e influenzando, conseguentemente, i risultati delle decisioni prese a valle della valutazione automatizzata⁷². Si consideri, inoltre, che la scelta espressa in sentenza dal giudice avvalso del supporto algoritmico si conforma, inevitabilmente, a una corrente giurisprudenziale maggioritaria: quest'ultima, da espressione di un orientamento consolidato, ma pur sempre relativo, assume le vesti di dato incontrovertibile e assoluto per effetto del processo di standardizzazione operato dal *tool* predittivo, così incidendo negativamente sull'autonomia e sull'indipendenza dell'organo giudicante.

A causa dell'insidioso «effetto performativo esponenziale dell'algoritmo»⁷³, l'adesione alla previsione riduce la disponibilità del decisore umano a rivedere la propria posizione, irrigidendo il meccanismo decisionale e privandolo della sua dimensione umana ed emotiva, essenziale per l'evoluzione dell'ordinamento giuridico. Riconoscere autorità di «precedente vincolante» a tutte le decisioni comporta, infine, l'ulteriore accorgimento – che rischia di trasformarsi in onere inadempito – di prevedere, in sede di programmazione, l'istruzione del software sul rapporto tra giurisdizioni inferiori e superiori, affinché l'algoritmo calibri qualitativamente i dati d'ingresso, attribuendo a ogni sentenza un diverso valore, a seconda che si tratti o meno di decisioni rese da giudici di ultima istanza⁷⁴.

Tra le numerose potenziali soluzioni ipotizzate per mitigare il più possibile le suddette criticità, di particolare interesse appare la proposta di ricorrere all'intelligenza artificiale per monitorare e guidare il percorso decisionale e motivazionale del giudice, mediante l'individuazione automatica di lacune o incongruenze nel ragionamento giuridico rispetto a decisioni già prese in casi simili,

⁷² S. Quattrocchio, *Intelligenza artificiale e giustizia: nella cornice della Carta etica europea, gli spunti per un'urgente discussione tra scienze penali e informatiche*, in www.la legislazione penale.eu, 1 gennaio 2018.

⁷³ G. Resta, *Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza*, in *PD* 2019, 214.

⁷⁴ S.M. Ferrié, *Les algorithmes à l'épreuve du droit au procès équitable*, in *La Semaine Juridique* 2018,6; N. Lettieri, *Contro la previsione. Tre argomenti per una critica del calcolo predittivo e del suo uso in ambito giuridico*, in *Ars int.* 2021, 91.

in un rapporto di determinazione reciproca tra giudice e macchina⁷⁵.

Affinché all'implementazione tecnologica del sistema giudiziario consegua un miglioramento qualitativo delle decisioni penali, è necessario limitare le conseguenze derivanti dall'opacità algoritmica e dal già citato "effetto ancoraggio", nonché garantire la centralità della dimensione antropocentrica dell'*iter* decisorio, che deve rimanere prerogativa assoluta del giudice, il quale si assume la piena responsabilità delle proprie scelte, anche nel caso in cui siano state assistite dall'attività di calcolo di una macchina. A tal proposito, l'*acquis* comunitario si è arricchito, in particolare nell'ultimo decennio, di linee guida e disposizioni normative significative, finalizzate all'inquadramento di un uso consapevole e bilanciato dell'intelligenza artificiale, nel rispetto dei diritti fondamentali e dei principi che governano il giusto processo.

⁷⁵ A. Punzi, *Judge in the machine: e se fossero le macchine a restituirci l'umanità del giudicare?*, in *Decisione robotica*, a cura di A. Carleo, Bologna 2019, 329.