

Intelligenze artificiali o artifici intelligenti? Cosa c'è dietro questo tipo di algoritmi? Una breve rassegna di interessanti risultati recenti

di *Gennaro Tedesco e Angelo Limone*

Introduzione

In un noto racconto di Isaac Asimov, l'umanità decide di costruire un calcolatore mastodontico per porgli la domanda fondamentale: “esiste Dio?”. L'enorme cervello artificiale viene costruito e finalmente acceso, il quesito viene posto e la risposta arriva subito, fredda e laconica: “Ora sì.”¹.

L'intelligenza artificiale sta vivendo negli ultimi (pochi) anni un'epoca d'oro, ricca di sviluppi. Nuovi algoritmi e ingegnose metodologie stanno prepotentemente acquisendo grande visibilità e i più recenti risultati possono evocare scenari inquietanti ma anche grandiose opportunità. Persino nelle attività considerate da sempre area di competenza del “flessibile” cervello umano e non del “lineare” e “rigido” cervello digitale, le macchine stanno rivelando prestazioni super-umane, ovvero commettendo meno errori degli umani stessi. Esempi recenti sono il riconoscimento facciale o delle immagini, territorio considerato esclusivamente umano fino a pochi anni fa.

Un ulteriore aspetto dove le intelligenze artificiali si stanno diventando un valido aiuto è l'analisi della mole gargantuesca di dati prodotti dall'umanità al giorno d'oggi. Si stima infatti che l'uomo produca l'equivalente di 5 milioni di computer portatili di dati al giorno, e questo numero è destinato molto probabilmente a crescere. È evidente che l'analisi – soprattutto se in tempo reale – dei suddetti dati da parte delle persone fisiche è opera tediosa, imperfetta e lenta – se non ai fini pratici addirittura impossibile.

Il presente breve saggio vuole spiegare in maniera semplificata in cosa consiste l'intelligenza artificiale divulgandone alcuni aspetti essenziali che la differenziano dalla programmazione classica. Gli autori – non essendo scienziati cognitivi – vogliono proporre materiale spunti di ri-

¹ I. Asimov, *The Last Answer*, 1980.

flessione di natura essenzialmente tecnica, senza porsi quesiti riguardanti l'auto-coscienza delle macchine (l'intelligenza artificiale sa di sapere? Capisce ciò che dichiara? Se sì, sa di capire? Ha propriocezione?) o scomodare la definizione del concetto di "intelligenza" o la dicotomia tra "umano" e "artificiale". L'intento è anche quello di demistificare alcuni concetti che possono apparire spaventosi ai non addetti ai lavori. In altre parole, l'enfasi è posta più sulla parola "artificiale" che "intelligenza".

1. Come funziona il *machine learning*

Supponiamo un programmatore abbia l'obiettivo apparentemente banale di insegnare a un computer come riconoscere con una certa accuratezza una mela analizzando solo l'immagine della stessa. Il computer è – si dice sorridendo – un "sapiente stupido" e non ha alcuna idea innata di cosa sia una mela. Il programmatore deve quindi lavorare selezionando alcune caratteristiche essenziali della mela: il suo colore rosso (ma alcune mele sono verdi o gialle), la sua forma essenzialmente tondeggiante ma rastremata al contempo, il caratteristico picciolo o una fogliolina sulla parte superiore (la cui presenza però a onor del vero non rende "mela" una mela), et cetera. Una volta scelte queste caratteristiche peculiari, implementarne l'essenza è cosa tutt'altroche banale. Il programmatore potrebbe ad esempio fare un'analisi statistica delle possibili forme di una mela e programmare una serie di *if-conditions* (letteralmente "condizioni se", ad esempio "se il rapporto tra il girovita della mela e la base è compreso tra 1.17 e 1.89, e se il rapporto tra altezza e base è...") ma cosa risponderà la macchina quando il programmatore le proporrà l'immagine di una mela vista dall'alto (la cui forma è più tondeggiante del profilo tipico a cui siamo abituati)?

O – aumentando ulteriormente la difficoltà – come reagirà di fronte a una mela morsicata? Dopotutto una mela morsicata è pur sempre una mela, ma spiegarlo a un computer è impresa ardua.

Di fronte a queste grandi difficoltà della programmazione classica, le tecniche dell'intelligenza artificiale propongono uno slittamento del paradigma che si concretizza nel cosiddetto *machine learning* (apprendimento macchina). L'idea è sottile ed è l'analogo informatico della transizione da una visione à la Leibniz basata sulle idee innate verso una visione à la Locke e Berkeley basata sull'esperienza. Qualcuno afferma che il nuovo metodo si basi apparentemente sul modo in cui il cervello umano risolve il problema. Dopotutto probabilmente nessun cervello umano calcola il rap-

porto tra girovita e base di un frutto per stabilire che si tratti di una mela e non di una pesca, ma molto più probabilmente ricorre all'“esperienza” che ha accumulato in passato e che è in qualche modo registrata nello *hard disk* neuronico. Ma come si concede alla macchina il prezioso dono dell'esperienza? La si espone a un bombardamento di immagini, alcune di mele, altre di non-mele. Senza entrare nei dettagli computazionali, alla macchina viene comunicato quali di queste immagini rappresentino effettivamente mele, e quali no. Questa fase – che viene chiamata training o “addestramento” – è molto delicata, perché la macchina deve essere esposta a un campione molto variegato di mele, pena il fenomeno dell'overtraining, ovvero una bulimica e negativa iper-specializzazione: se mostro al computer solo mele della tipologia Granny Smith, il computer classificherà una mela Golden come non-mela, perché non corrispondente ai canoni e ai “pregiudizi” a cui il calcolatore è stato abituato durante la fase di addestramento.

Il fenomeno dell'overtraining è assolutamente “umano”: chiunque abbia provato a imparare una lingua straniera basata su un alfabeto diverso da quello della propria lingua madre avrà sperimentato una inaspettata difficoltà a decifrare la scrittura a mano nell'alfabeto straniero (quello che per noi è il cosiddetto corsivo, per intenderci), in quanto abituato ad essere addestrato tramite lettura di soli testi stampati.

Il suddetto esempio della mela non è una metafora, bensì la descrizione di un problema reale quotidiano affrontato dalle compagnie e dagli istituti di ricerca che applicano l'intelligenza artificiale al riconoscimento e classificazione automatico delle immagini o alla generazione automatica di didascalie. L'esempio ovviamente può essere generalizzato a situazioni molto più complicate. Si pensi ad esempio al riconoscimento dell'immagine di un gatto, a prescindere dalla razza, dall'angolazione della foto, dalla variabilissima posizione dell'animale ritratto, dal livello di illuminazione e contrasto, dalla presenza di altri elementi disturbatori nell'immagine. E così via: insegnare al computer tramite la forza bruta dell'esperienza il concetto di “tavolinità” e “cavallinità” di platonica memoria esponendolo a un numero adeguato di esempi. *Caveat*: la fase di training è ulteriormente complicata dal fatto che un'esposizione eccessivamente variegata a tutti gli esemplari di mela (o di gatto) è parimenti pericolosa: il computer potrebbe non essere in grado di astrarre le caratteristiche essenziali dell'oggetto da riconoscere se – a causa dell'eccesso di varietà – non riesce a notare una sorta di regolarità e di *pattern*, ovvero schemi e strutture tipiche che si ripetono nelle varie figure usate durante l'addestramento.

A questo punto il gioco è fatto: dopo un addestramento idoneo, il computer “saprà” (virgolette d'obbligo) cosa è una mela e riconoscerà

nuovi esemplari di mele, e – qui sta la potenza del metodo – esso riconoscerà i nuovi esemplari anche se questi non sono mai apparsi tra gli esempi dell’addestramento. Da questo punto di vista si evince che la terminologia utilizzata in letterature di intelligenza artificiale è quantomeno sfortunata: negli esempi di cui sopra (e negli altri a seguire) le macchine non presentano alcuna caratteristica intelligente di per sé, bensì riconoscono strutture e *pattern* ai quali esse sono state precedentemente esposte, giudicando e creando in base alle sole informazioni ricevute; pertanto, esse non possiedono alcuna caratteristica di “animale pensante”, almeno fino ad oggi.

2. Analisi del testo

Un ulteriore campo di ricerca dove il *machine learning* sta portando contributi sostanziali è il *Natural Language Processing*, ovvero la comprensione automatica di linguaggio di testo scritto. Dal punto di vista concettuale, umano, un linguaggio si riduce (chiaramente tenendo conto delle possibili complicazioni) ad una successione di termini ai quali è associato un concetto (o significato) predefinito a seconda dalla maniera e della frequenza con la quale tali termini entrano all’interno del testo; una stessa parola può portare a diversi significati a seconda se essa sia seguita (o preceduta, rispettivamente) da un altro gruppo di parole determinate. È facile derivare, a questo punto, come lo scheletro della comprensione concettuale delle lingue possa “ridursi” (chiaramente in maniera più o meno complicata) in prima istanza all’analisi di gruppi o collezioni di termini, presi in successione. La decifrazione del linguaggio di testo si basa infatti proprio sul principio sopraenunciato: per facilitare la comprensione si prendano ad esempio gli enunciati seguenti

Il gatto è seduto sul tavolo
 Il gatto è seduto sulla sedia
 Il cane rincorre la palla

dove è facile riconoscere intuitivamente che le prime due frasi sono simili tra loro mentre invece la terza è diversa da ciascuna delle precedenti. Procedendo ad una analisi del testo si nota infatti che nelle prime due frasi il soggetto ed il verbo coincidono, il complemento essendo l’unica distinzione: in particolare i gruppi di parole (ai quali cisi riferisce, nella letteratura, come N-grammi)

(Il, gatto) (gatto, è) (è, seduto)
 (Il gatto, è) (gatto è, seduto)

si ripetono identicamente, con la stessa frequenza; tale caratteristica manca invece nell'ultimo enunciato, dove

(Il, cane) (cane, rincorre) (rincorre, la)

(Il cane, rincorre) (cane rincorre, la)

non occorrono in alcuna delle prime due frasi. Si può a questo punto costruire una rappresentazione di una collezione di frasi assegnando a ciascuna di esse un insieme di numeri che rappresentano la frequenza con la quale le parole vengono appaiate consecutivamente (per essere più precisi, invece che singole parole bisogna costruire collezioni di N-grammi, come negli esempi di cui sopra); l'insieme di componenti numeriche così calcolate può essere collezionato e trattato come una unica entità, la quale può essere sottoposta ad operazioni matematiche che andranno a sostituire i legami logici tra le proposizioni letterali.²

Tale approccio, investigato dagli autori T. Mikolov, K. Chen, G. Corrado e J. Dean³, fornisce lo strumento principale di analisi automatica del testo; una volta fornita la rappresentazione vettoriale è facile calcolare la similitudine (o la differenza, rispettivamente) e la "distanza" di due porzioni di testo semplicemente⁴ come distanza geometrica tra due punti descritti dalle sopra citate rappresentazioni numeriche.

Questa tecnica permette persino di poter effettuare "equazioni di testo" dove è possibile indovinare porzioni di informazione nascoste utilizzando il contesto precedente e successivo: ad esempio, nella sottostante

re + regina = uomo + ?

la parola mancante può essere letteralmente "calcolata" sommando le varie rappresentazioni

donna = re + regina - uomo.

Sebbene la descrizione fornita di sopra sia elementare, il lettore può immaginare come sia possibile raffinare ulteriormente tali algoritmi utilizzando il contesto, i tempi verbali e qualsiasi altra struttura linguistica per una migliore rappresentazione delle frasi, in modo tale da ottenere una precisione avanzata nel calcolo della componente testuale: i cosiddetti

² Si dice che al testo viene associata una rappresentazione vettoriale.

³ T. Mikolov, K. Chen, G. Corrado, and J. Dean, *Efficient Estimation of Word Representations in Vector Space*, in *ArXiv e-prints*, January 2013.

⁴ Si legga, invece, in maniera più o meno complicata.

embedding permettono una vettorizzazione basata sulla grammatica (tempi verbali, presenza di verbi ausiliari, declinazione di genere, numero e caso) e sul contesto lato in cui una parola appare nel linguaggio; la precisione di tali tecniche sta velocemente raggiungendo livelli in cui informazioni dettagliate pertinenti a testo scritto possono essere automaticamente estratte partendo dal contesto, e testi possono essere costruiti dagli algoritmi partendo semplicemente da modelli probabilistici di catene di parole ottenute utilizzando le tecniche descritte nel paragrafo precedente. Le applicazioni di tale studio sono diverse, e trovano spazio nell'analisi e gestione di contratti legali (dove l'algoritmo si sostituisce al lettore umano ed estrae le porzioni di testo interessanti), la classificazione di documenti ed articoli di ricerca, la gestione di CV nei dipartimenti di risorse umane et cetera.

3. Applicazioni a astronomia e scienza

L'astronomo osservativo moderno ha il lusso di poter accedere a una significativa quantità di dati. Questo grazie alla costruzione di grandi telescopi che stanno esplorando e digitalizzano alcune porzioni di cielo con risoluzione molto raffinata, per di più a diverse lunghezze d'onda della radiazione (dalle onde radio all'infrarosso, dal visibile all'ultravioletto, fino ai raggi X) e soprattutto nel tempo (la variabilità di molte sorgenti luminose nei cieli può essere misurata ed è informazione molto preziosa, perché può rivelare la natura della sorgente). Lamole di dati è cospicua. Non solo, è composita: nello stesso grado quadrato di cielo, un telescopio può registrare la presenza di oggetti della natura più diversa e a distanze completamente scorrelate (una galassia a spirale a fianco a un lontanissimo quasar primordiale di sottofondo o una stella della Via Lattea relativamente vicina, o un asteroide roccioso). L'esplorazione certosina e profonda del cielo è ancora agli inizi e i cataloghi moderni possono già contenere svariate decine di migliaia di galassie, ognuna delle quali è caratterizzata da decine di misure (ad esempio la luminosità in differenti bande o filtri ottici). Il problema della classificazione degli oggetti diventa quindi stringente e la classificazione manuale di tutte queste osservazioni è un lavoro molto lento e difficile (molti oggetti si somigliano in differenti caratteristiche). L'intelligenza artificiale può venire incontro all'astronomo nei problemi di classificazione, e non solo per dipanare la massa enorme di dati, ma per un ulteriore motivo, ovvero per trovare nuove correlazioni tra dati, che spesso possono essere la manifestazione di leggi fisiche non del tutto chiarite.

Per verificare che esista una correlazione tra due variabili misurate,

spesso basta porre in grafico i dati su un piano cartesiano i cui assi rappresentano le due variabili. Se le due variabili sono del tutto scorrelate, i punti sperimentali si spalmano sul piano in maniera casuale, senza aggregarsi in forme particolari. Ad esempio, supponiamo di avere a disposizione dati di varia natura riguardanti un campione statistico di gattini. Come primo esercizio, potremmo mettere in grafico l'età e il peso di questi gattini. Molto probabilmente noteremo che gatti più giovani tendono a pesare di meno e viceversa, quindi i punti si disporranno prevalentemente su una linea che indica una certa proporzionalità tra le due variabili. Ovviamente questa non è una legge esatta, ma una tendenza, in quanto esisteranno gatti più pesanti e robusti della media sebbene più giovani. Quando analizza uno di questi grafici, il ricercatore è spesso alla ricerca dei cosiddetti *cluster*, ovvero punti sperimentali che si ammassano sospettosamente nella stessa regione del piano cartesiano invece di spaziare nell'intero piano quando c'è assenza di correlazioni o lungo curve che denotano una tendenza già nota (come la curva età-peso di cui sopra). Questi ammassi di dati possono essere spia di una peculiarità sottostante inaspettata che necessitano approfondimento.

Nell'esempio dei gattini, i dati corrispondenti a gattini soggetti a malnutrizione o affetti da malattie della crescita potrebbero discostarsi dall'andamento lineare e andare ad ammassarsi altrove, mostrando che la patologia rompe la tendenza e crea un riconoscibile *cluster* di dati. È chiaro che il problema diventa molto difficile da affrontare quando le variabili in gioco sono tante. La grafica moderna permette di graficare – usando trasparenze e prospettiva – anche tre assi, quindi – a vista – possono essere rivelate correlazioni tra anche tre variabili. Aggiungendo l'ausilio del colore come “quarto asse”, si può provare ad arrivare forse fino a quattro. Ma come fare quando le variabili sono diciassette o di più? Bisognerebbe studiare i grafici delle varie variabili a due a due (o a tre a tre), in tutte le possibili combinazioni, e sperare di vedere una aggregazione di dati o una correlazione.

È chiaro che il compito diventa arduo e inefficiente per un essere umano, ma non per la macchina, la quale cerca correlazioni senza l'ausilio visivo e soggettivo dei grafici, ma esplorando lo spazio dei parametri senza restringerlo ogni volta a sole due dimensioni.

Come già accennato, la correlazione non è una legge fisica, ma un andamento statistico e potrebbe non essere il frutto di alcuna causalità ma semmai di una casualità. Per esempio, la quantità media di mozzarella consumata negli Stati Uniti tra il 2000 e il 2009 correla molto bene con il numero di dottorati assegnati agli ingegneri civili negli stessi anni; il nu-

mero di film in cui appare Nicolas Cage tra il 1999 e il 2009 correla con il numero di persone morte annegate in piscina, ma gli autori di questo breve saggio dubitano fortemente si tratti di causalità genuine. Allo stesso tempo però, le correlazioni possono segnalare un indizio di una legge o di qualcosa di inaspettato (come appunto la presenza nel campione di gattini malnutriti). Di questi indizi si ciba volentieri l'astronomo moderno, che cerca di trovare regolarità e interessanti eccezioni nell'apparente caotica zoologia astronomica.

Ma torniamo ai problemi di classificazione. Come già spiegato nell'esempio della mela, la macchina va addestrata. In altre parole, l'astronomo deve mostrare all'intelligenza artificiale una serie di galassie di cui conosce già le caratteristiche o che sono già state classificate in

precedenza. Un esempio ormai routinario riguarda il calcolo della distanza di una galassia osservata. Questo calcolo si basa spesso sull'osservazione accurata della radiazione/luce che questa galassia emette.

Senza entrare nei dettagli tecnici, classicamente l'astronomo riesce a calcolare la suddetta distanza in base a quanto lo spettro elettromagnetico di questa galassia è "arrossato" o, meglio, "spostato verso il rosso". Questo calcolo potrebbe tanto essere effettuato manualmente quanto in maniera automatizzata. In entrambi i casi tale calcolo può risultare relativamente lungo e computazionalmente costoso se il catalogo di galassie in studio è molto vasto. A complicare le cose, non sempre si possiede la misura dettagliata e con buona risoluzione di tutto lo spettro elettromagnetico⁵ della radiazione della galassia. Effettuare tale misura, infatti, è un lavoro relativamente lungo, relativamente costoso e necessita degli strumenti adeguati. Spesso quindi l'astronomo deve rinunciare allo spettro e accontentarsi invece di poche misure in pochi filtri ottici, di vario colore, che perdono buona parte dell'informazione dettagliata che lo spettro avrebbe invece provveduto, rendendo il calcolo della distanza veramente poco accurato. Se però si possiede un sottoinsieme di galassie di cui è già nota la distanza e di cui sono note anche le meno nobili misure nei vari filtri ottici, l'astronomo può provare a rivolgersi all'intelligenza artificiale, sperando che questa trovi una relazione empirica tra i dati (filtri ottici) e la quantità che si vuole calcolare (la distanza). All'intelligenza artificiale viene insegnato quale distanza essa debba aspettarsi da una galassia avendo a disposizione le misure nei pochi filtri ottici della stessa, ma non dello spettro.

⁵ Lo spettro di una sorgente elettromagnetica riporta quanta energia è irradiata per ogni frequenza o – equivalentemente – lunghezza d'onda della radiazione. L'immagine del prisma di vetro che scompone la luce bianca nelle diverse componenti di diverso colore (si legga frequenza) può dare al lettore un'idea grossolana del concetto.

La macchina impara quindi da una serie di dati di cui è già noto il risultato (come nel caso della mela) ed è chiamata poi ad astrarre questa idea che essa si è fatta della relazione – qualsiasi essa sia – tra distanza e misure ottiche, e applicare tale relazione a quelle misure ottiche di cui invece non si conosce la distanza. Si noti che la relazione empirica trovata dall'intelligenza artificiale è spesso di difficile interpretazione da parte di un umano. I risultati trovati con questo metodo saranno meno precisi di un esame completo spettroscopico, ma saranno veloci e meno costosi, quindi applicabili ad un ampio catalogo in modo da fare – ad esempio – una prima scrematura (l'astronomo, ai fini della propria ricerca specifica, potrebbe ad esempio essere interessato a studiare solo galassie molto lontane o molto vicine). Il caso del catalogo di galassie e quello della mela sono entrambi esempi di *supervised learning* (apprendimento supervisionato), dove cioè la macchina impara grazie al tramite umano, che funge da supervisore, spiegandole gli esempi noti da cui la macchina deve imparare. L'astronomia è un campo molto prone all'applicazione delle intelligenze artificiali, in quanto al momento è ricco di dati e si presta bene ai tipici problemi che l'intelligenza artificiale sa risolvere con ottima affidabilità (classificazione, riconoscimento, analisi delle immagini e di serie temporali). Ci sarebbero altri possibili esempi interessanti da proporre, come la scoperta molto recente – da parte della NASA – di un esopianeta⁶ probabilmente roccioso e non gassoso, Kepler-90i, ottenuta analizzando la mole di dati tramite tecniche di *machine learning*⁷.

4. Teoria dei giochi

Parte della storia dell'informatica è strettamente legata all'abilità di risolvere giochi deterministici ed in particolare gli scacchi sono da sempre stati utilizzati per testare calcolatori e le loro abilità a trovare soluzioni di tipo combinatorio. Il diagramma sottostante mostra il primo problema scacchistico risolto da un computer: il bianco muove e vince in due mosse.

⁶ Un pianeta esterno al Sistema Solare.

⁷ C. J. Shallue and A. Vanderburg, *Identifying exoplanets with deep learning: A five planet resonant chain around kepler-80 and an eighth planet around kepler-90*, accepted in *The Astronomical Journal*, 2018.



Posizione analizzata dal primo programma di scacchi. *Mate-in-two* di Dietrich Prinz utilizzava ricerca iterativa di tutte le mosse legali nella determinata posizione fino ad esaurire le possibilità per il Nero, determinando così la soluzione di matto 1. Th6 gxh6 2.g7#

I programmi di scacchi moderni di tipo classico (dove il termine si riferisce a qualsiasi programma precedente il *Mastering Chess and Shogiby Self-Play with a General Reinforcement Learning Algorithm*⁸) utilizzano il concetto di funzione di valutazione per assegnare un valore numerico ad ogni posizione sulla scacchiera, tale valutazione essendo unica nel caso di giochi a somma zero, come in questione⁹. Una posizione sulla scacchiera consiste in un insieme statico di caratteristiche, come ad esempio quantità di materiale, sicurezza del re, possesso della coppia degli alfiere o di torri su colonne aperte (o qualsiasi altro concetto equivalente) e in un insieme dinamico, ovvero la presenza, in suddetta posizione, di una combinazione forzata di mosse che porti alla fine della partita o ad un vantaggio permanente. A ciascuno degli elementi di cui sopra è assegnato un valore o peso numerico e la valutazione della posizione è semplicemente la somma di ciascuno dei contributi menzionati; a questo punto la ricerca della “mossa migliore” si riduce alla ricerca della mossa che generi una posizione la cui valutazione numerica è massimizzata.

Da questo punto di vista il programma scacchistico non ha alcuna conoscenza o esperienza del gioco in sé, piuttosto lo riduce a mero calcolo di varianti scegliendo alla fine quella che massimizzi il valore numerico di

⁸ D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, and D. Hassabis. *Mastering Chess and Shogiby Self-Play with a General Reinforcement Learning Algorithm*, in *ArXiv e-prints*, December 2017.

⁹ Un gioco si dice essere a somma zero quando il guadagno del primo giocatore è bilanciato esattamente dalla perdita del secondo.

una serie di parametri immessi ad hoc. Il calcolo deterministico di tutte le possibilità, assegnata una posizione iniziale, procede dal punto di vista concettuale nell'esplorazione forzata di ciascuna di esse ed in questo riguardo non esiste alcuna differenza con i programmi primitivi che analizzavano ad una ad una tutto l'insieme di mosse legali. Il vantaggio di avere una valutazione della posizione giace però nella possibilità di ottimizzare il calcolo e di tralasciare continuazioni la cui valutazione è peggiore di altre. Il programma fissa una profondità di ricerca a priori (equivalente più o meno al numero di mosse che vengono analizzate in anticipo) e restituisce la mossa che massimizza la valutazione della posizione finale dopo aver effettuato tutte le possibili continuazioni che possono essere generate. In questa fase viene utilizzata l'ottimizzazione dell'*alpha-beta pruning*: a ciascun giocatore viene assegnato il peggior valore possibile della posizione, α infinitamente negativo per il giocatore di turno e β infinitamente positivo per l'avversario; si noti che, in questa fase, $\alpha < \beta$. A questo punto la ricerca della mossa migliore viene inizializzata e vengono calcolate le possibili continuazioni a partire dalla posizione di partenza; a ciascuna variante corrisponde una coppia di valori di α e β , le quali corrispondono alle valutazioni della posizione di ciascuno dei due giocatori, data la variante in considerazione; se viene incontrata una continuazione per la quale esiste almeno una mossa tale che la valutazione massima del secondo giocatore risulta essere minore della valutazione minima del primo giocatore, ovvero $\beta \leq \alpha$, tale variante viene scartata dalle analisi in quanto fornisce almeno una soluzione peggiore della posizione iniziale per il giocatore di turno. Sebbene la profondità iniziale di ricerca sia fissa, qualsiasi programma di scacchi commerciale riesce, grazie al suddetto metodo, a dimostrarsi di gran lunga superiore a qualsiasi Gran Maestro professionista. Al momento della stesura un recente articolo¹⁰ è stato pubblicato dalla compagnia londinese *Google Deep Mind*, nel quale si presentano nuove tecniche di computazione che hanno permesso la costruzione di un programma di scacchi (*Alpha Zero*) basato sull'utilizzo del *machine learning* piuttosto che di una funzione di valutazione *ad hoc*. Sebbene i dettagli dell'algoritmo non siano ancora stati resi pubblici, la metodologia generale è basata sul *Monte Carlo tree search*, qui spiegata.

Il programma viene inizializzato solamente con le regole generali del gioco, senza alcuna valutazione posizionale prefissata, e, partendo da mosse completamente casuali (ma legali), viene generato un immenso numero

¹⁰ D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, and D. Hassabis, *Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm*. cit.

di partite, le quali vengono calcolate e giocate a forza bruta fino alla fine. Ad ogni partita successiva, assegnata una posizione iniziale, il programma confronta l'insieme di varianti possibili con alternative simili presenti in partite precedentemente giocate da sé stesso (e conservate in memoria); successivamente la ricerca viene ottimizzata sulle mosse che hanno portato al maggior numero di vittorie, i parametri vengono ri-aggiornati automaticamente dopo l'ultimo ciclo di calcolo e la valutazione della posizione è basata sull'esperienza acquisita dell'insieme di tutte le partite giocate piuttosto che in una collezione di parametri fissati a priori. Il vantaggio di questo tipo di approccio – noto come *reinforcement learning* – è di fornire una valutazione molto più vicina ai piani pratici di gioco piuttosto che derivata da concetti astratti generali; al contrario, però, questo approccio funziona soltanto una volta generata una collezione abbastanza ricca di partite che permette un confronto oggettivo per più o meno qualsiasi posizione.

Da questo punto di vista un algoritmo di questo tipo risulta fornire un gioco molto più “umano”, ovvero basato sull'esperienza accumulata durante lo studio e sul riconoscimento di schemi comuni, ovvero la realizzazione che la geometria della scacchiera permette risultati simili se le condizioni della posizione iniziali sono anch'esse simili tra loro; tutto sommato, è così che i giocatori professionisti scelgono la mossa migliore.

Le conseguenze dell'utilizzo di tali programmi negli scacchi professionistici ha portato ad un miglioramento della comprensione delle posizioni. Contrariamente all'immaginario comune (che dipinge i giocatori trovare le mosse migliori sulla scacchiera) gli scacchi sportivi sono prevalentemente preparazione scientifica: dal punto di vista dell'apertura molte varianti classiche considerate buone per il Bianco sono state confutate, in quanto si sono trovate continuazioni di difesa che evitano vantaggi forzati per il primo giocatore; allo stesso modo, lo studio dei finali è arrivato alla computazione e alla tabulazione esatta di posizioni contenenti fino a sei figure¹¹. Lo scopo finale di un programma scacchistico (così come in qualsiasi altra materia) è chiaramente la soluzione del gioco, ovvero il determinare l'esistenza o meno di una strategia che permetta ad un giocatore di non perdere, indipendentemente dalla strategia utilizzata dall'avversario. Alcuni giochi sono stati completamente risolti (alcune versioni della dama sportiva e del Go su scacchiera piccola), per i quali si è mostrato l'esistenza di una strategia che permette la vittoria o la patta assegnata la posizione iniziale; la complessità degli scacchi rende l'obiettivo ancora lontano, correntemente: una stima approssimativa può essere effettuata tenendo conto

¹¹ <http://tb7.chessok.com/>

che, assegnata una posizione qualsiasi, il numero di mosse legali a disposizione è 35^{12} , il quale deve essere ripetuto per qualsiasi coppia di mosse; un rapido calcolo, assumendo una partita di 40 coppie di mosse (quindi 80 posizioni diverse) mostra un ordine di grandezza di $35^{80} = 10^{123}$ possibili combinazioni (tale ordine di grandezza essendo noto come numero di Shannon). Ciò nonostante è bene sottolineare che il gioco degli scacchi ammette di per sé una soluzione (sebbene essa sia al momento difficile da calcolare), ovvero esiste una strategia che permette ad uno dei due giocatori di non perdere (ovvero forzare una vittoria od una patta): tale risultato, noto in letteratura come teorema di Zermelo, si estende a tutti i giochi di due giocatori a somma zero, nel caso in cui non siano presenti alcuni contributi casuali.

Un altro risultato eclatante raggiunto con la tecnica sopra menzionata, il *reinforcement learning*, è il recente sorpasso – nell’antico gioco asiatico del Go – da parte dell’algoritmo *AlphaGo Zero* ai danni di giocatori umani, battendo ripetutamente i migliori algoritmi esistenti e anche il campione mondiale Lee Sedol. L’algoritmo – anche questo ideato dal già citato *Google Deep Mind* – è passato dallo stato di amatore a quello di Gran Maestro in soli 3 giorni e ha imparato da solo, in 40 giorni, quello che i media hanno giustamente descritto come tremila anni di storia e tradizione accumulati in questa attività ludica. Ad *AlphaGo Zero* sono state semplicemente insegnate le regole del gioco ed è stato esso stesso a ricavare strategie e tattiche semplicemente giocando contro se stesso partendo da tabula rasa e quindi diventando esso stesso il proprio istruttore di Go¹³. In taluni casi esso ha addirittura proposto delle mosse o configurazioni originali. Software di intelligenza artificiale precedenti ad *AlphaGo Zero* venivano addestrati nutrendoli di un database enorme di partite (umane) già giocate; l’elemento invece innovativo di *AlphaGo Zero* sta nell’aver imparato senza aver avuto accesso ad alcuna cronologia umana, ma semplicemente da solo. Questa attività di ricerca non è semplice *divertissement*. Uno degli intenti manifestati dall’amministratore delegato di *Google DeepMind*, Demis Hassabis, è quello di far avanzare ulteriormente il livello di questo tipo di algoritmi per affrontare problemi reali riscontrati ad esempio nell’ambito della medicina e farmacologia, così come dell’ingegneria e della scienza. Andy Okun e Andrew Jackson, membri della *American Go Association*, hanno commentato

¹² A meno di trovarsi in posizioni particolari di stallo, scacco o simili

¹³ D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel and D. Hassabis, *Mastering the game of go without human knowledge*, in *Nature*, 550:354, October 2017.

l'eccitazione derivata da questi progressi con la seguente affermazione su *Nature News and Views* [traduzione nostra]: «I giocatori di Go, provenendo da così tante nazioni diverse, comunicano attraverso le mosse del gioco, anche quando non parlano alcuna lingua in comune. Sulla scacchiera essi condividono così idee, intuizioni e, in definitiva, i propri valori — non solo particolari aperture o tattiche, ma se preferiscono il caos all'ordine, il rischio alla certezza, la complessità alla semplicità. È ancora lontano il giorno in cui gli uomini avranno una conversazione sensata con un'intelligenza artificiale, è sempre sembrata roba da fantascienza. Per i giocatori di Go, quel giorno è già arrivato»¹⁴.

5. *Training* non supervisionato

Gli esempi discussi fino ad ora, nella maggior parte dei casi, trattano di classificazione supervisionata da parte degli algoritmi, ovvero il processo di *training* avviene assegnando un certo numero di caratteristiche etichettate “a priori” come determinanti una certa categoria (ad esempio la categoria “mela”), in modo tale che la macchina possa calcolare, successivamente, similitudini tra nuovi esempi e dati assegnati, per poter così classificare in base alle maggiori somiglianze (si veda il paragrafo 2).

È doveroso citare che esistono anche algoritmi di *training* non supervisionato, ovvero dove non viene assegnato alcuna categorizzazione iniziale, bensì è l'algoritmo stesso a dover separare il *set* di dati in diverse categorie. Come si può immaginare, tali meccanismi sono concettualmente più complicati e la classe di problemi ai quali essi possono essere applicati con precisione è minore: ciò nonostante un caso molto comune è quello della determinazione degli *outliers* in un *set* di dati, ovvero del sottoinsieme dell'insieme di partenza le cui caratteristiche sono “diverse” dal resto generale.

Un esempio concettualmente efficace è fornito dal *random forest*: l'insieme di dati di partenza viene separato dall'intelligenza artificiale in sottogruppi che condividono le stesse caratteristiche (tale separazione è nota in letteratura con il termine di alberi decisionali). Come già esposto in precedenza, i dati possono essere rappresentati come punti in uno spazio cartesiano, dove le coordinate di tali punti rappresentano la misura di una certa variabile. Questi punti vengono quindi appaiati a due a due, a tre a

¹⁴ D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel and D. Hassabis, *Mastering the game of go without human knowledge.*, cit.

tre e così via a seconda dell'insieme di caratteristiche comuni che essi condividono, dove per caratteristiche si intende qualsiasi insieme di parametri utilizzato per descrivere il set di dati (siano essi componenti di vettori o altro); l'aggregazione dei punti avviene a seconda di metodi più o meno complicati, che vanno dal semplice considerare quante componenti coincidono a più complessi meccanismi geometrici come l'appartenere alla stessa superficie. A partire da un punto generico iniziale, quindi, si disegnano diverse ramificazioni, ciascuna delle quali corrisponde al sottoinsieme dei restanti punti che condividono delle caratteristiche con il punto iniziale; tale rappresentazione, infatti, è molto simile al disegnare alberi decisionali, da cui l'algoritmo prende il nome. Una volta esaurite tutte le possibilità combinatorie, alcuni meccanismi di validazione vengono introdotti, com e ad esempio il mischiare casualmente alcune componenti o ripetere la scelta iniziale (per essere sicuri di non aver introdotto alcun errore sistematico), o il semplice fare la media delle distanze dei punti finali dalla ramificazione iniziale. Alla fine del processo, sotto l'ipotesi iniziale che l'insieme degli *outlier* sia piccolo, vengono scelti quei dati che, secondo tutti i processi di validazione descritti sopra, hanno mostrato di condividere il "minor numero" di caratteristiche con il resto del gruppo e, in quanto tali, essere naturalmente "diversi", per così dire. Un esempio di applicazione del random forest si ha nel tracciamento dei movimenti e trasferimenti bancari di un certo numero di individui sospetti per poter evidenziare potenziali frodi: i trasferimenti bancari, in termini di coordinate, quantità e caratteristiche, vengono mappati e vettorizzati, in modo tale che ogni operazione possa essere descritta da un *set* di componenti. A questo punto l'insieme dei trasferimenti viene sottoposto a *random foresting*, con l'obiettivo di determinare un sottoinsieme piccolo di *outlier*, ovvero di trasferimenti le cui caratteristiche si discostino dai restanti e che quindi, sotto l'assunzione che la maggior parte delle persone effettui azioni bancarie quotidiane, possano rappresentare possibili frodi di danaro o riciclaggio dello stesso.

A questo punto, dall'insieme così determinato vengono tracciati i proprietari così che essi possano essere soggetti ad una attenzione più ravvicinata da parte degli enti legali e di dovere. Chiaramente si può ben immaginare che tali applicazioni possono estendersi a tutti i problemi di natura simile, dove bisogna scartare un certo numero piccolo di dati le cui caratteristiche risultino differenti dalla massa. Simili esercizi vengono effettuati, ad esempio, anche nei test medici su farmaci nuovi, quando non esistono risultati sperimentali su pazienti con condizioni note ma dove ciononostante bisogna determinare la possibilità di continuare la ricerca su uno o su un altro prodotto. Al di là del *random forest* esistono altri algoritmi

non supervisionati: vale forse la pena citare il *clustering generico*, che tratta a grandi linee del problema di raggruppare dati in sottoinsiemi di densità quanto più alta possibile, e le reti neurali, un complesso insieme di algoritmi portati a risolvere problemi di matematica non lineare. In entrambi i casi, il lettore può scegliere di approfondire lo studio su ciascuno degli esempi di sopra, in quanto essi meriterebbero da soli una lunga dissertazione.

7. Attuali limiti

Attuali algoritmi di *Natural Language Processing* sono in grado di estrarre la data di scadenza di un contratto legale nonostante tale data non venga esplicitamente menzionata nel testo del contratto stesso. Ad esempio: se nella stipulazione compare la dicitura “allo scadere del dodicesimo mese dalla sottoscrizione di questo contratto”, la macchina può interpretare correttamente questa frase e “calcolare” la data in questione. Il lettore a questo punto potrebbe chiedersi se tutto ciò implichi quindi che la macchina sia in grado di “capire” il significato del testo. Dopotutto la macchina potrebbe essere in grado di rispondere in maniera corretta a semplici domande di verifica della comprensione, come quelle che si pongono agli scolari dopo la lettura di un testo. Al momento la risposta a questa domanda è negativa. La macchina lavora ancora su un piano meramente sintattico e protosemantico e non (ancora?) su un piano semantico o metalinguistico. Già negli anni '70 ci si poneva domande di questo tipo, anche nell'ambito del *Natural Language Processing*. Si veda ad esempio l'articolo esemplare *Artificial Intelligence meets Natural Stupidity*¹⁵ – L'intelligenza artificiale alle prese con la stupidità naturale – dove fallacie di vario tipo vengono attribuite ai programmatori di intelligenze artificiali (applicate ad esempio alla comprensione dei testi). Di fronte a questo tipo di quesiti, puntualizziamo ulteriormente che gli autori di questo breve saggio non stanno qui discutendo il problema dell'intelligenza. Al momento della stesura di questo scritto, nessun programma software di conversazione ha ancora superato il celebre test di Turing quando sottoposto a giuria esperta. Il lettore interessato può consultare numerosi testi specializzati o la recente opera divulgativa di Luciano Floridi¹⁶, ovvero proprio uno dei membri della giuria del test di

¹⁵ D. McDermott, *Artificial intelligence meets natural stupidity*, in *SIGART Bull.*, (57):4–9, April 1976.

¹⁶ L. Floridi, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Raffaello Cortina Editore, Milano 2017.

Turing per la competizione Loebner del 2008.

Negli esempi mostrati fin'ora è evidente come il termine intelligenza artificiale sia stato applicato ad algoritmi che permettono la classificazione (nel caso di *Natural Language Processing* o categorizzazione stellare) o la stima di parametri (come nel caso del gioco degli scacchi) basati sul training e su un certo numero di patterns forniti alla macchina durante l'addestramento. È evidente che molti (se non tutti) gli algoritmi di machine learning hanno come scopo l'indovinare (o meglio, il calcolare con probabilità massima) quale porzione di informazione possa essere predetta una volta assegnato il contesto circostante: ci si riferisce a questa area del *machine learning* con il termine di intelligenza artificiale di "tipo debole", per sottolineare appunto come i risultati siano sempre dipendenti dall'addestramento iniziale e dalle condizioni in cui la macchina opera, in generale. Al giorno d'oggi tutte le applicazioni esistenti sono di tipo debole.

A questo si contrappone, dal punto di vista concettuale, una intelligenza artificiale di "tipo forte" che sia in grado di interpretare il contesto alla quale viene sottoposta e di scegliere un addestramento mirato (e di conseguenza produrre risultati) basandosi sull'unica informazione di partenza, senza che le venga assegnato una serie di patterns da riconoscere e sui quali fondare il training.

8. Considerazioni sull'impatto sulla società

Esistono diverse posizioni riguardo all'impatto delle intelligenze artificiali sulla società. Ultimamente si sono lette opinioni allarmiste come quelle del visionario Elon Musk o del fisico Stephen Hawking negli ultimi anni della sua vita, dove – come in una realizzazione dello scenario da incubo de *L'apprendista stregone* o di *Frankenstein* – l'uomo perde il controllo dell'automa che ha creato per scopi benefici o per pura curiosità scientifica. Altre posizioni sono invece caute e scettiche, dove si prevede una saturazione naturale dello sviluppo di queste tecniche. Esistono poi posizioni entusiaste di chi reputa che le intelligenze artificiali alla lunga creeranno più posti di lavoro di quelli che distruggeranno, così come quelle di chi intravede la chiave del futuro di queste tecnologie nella pacifica e fruttifera convivenza tra l'uomo e la macchina (due agenti che si completerebbero a vicenda senza valicare le competenze altrui). Quest'ultimo scenario, che rientra nel filone del cosiddetto *artificial intelligence augmentation* o *amplification* (amplificazione delle capacità umane tramite intelligenze artificiali) è qualcosa che stiamo già sperimentando quando il nostro cellulare ci suggerisce

quale sia la parola successiva a quelle già scritte in un certo messaggio o quando l'astronomo fa uso delle tecniche descritte sopra et cetera. Esiste molta confusione su questo campo oggigiorno e spesso gli articoli allarmisti ignorano o sorvolano alcuni ostacoli pratici e teorici che le intelligenze artificiali sono ben lungi dal risolvere.

Gli autori di questo scritto sono degli entusiasti della tecnologia, purché ovviamente utilizzata con criterio e maturità. Ed è per questo che essi trovano pericoloso o per lo meno insoddisfacente il fatto che il pubblico e l'ambiente accademico spesso non siano molto aggiornati sui recenti risultati di chi lavora in questo campo. È dovere di chi lavora in ambito tecnico-scientifico quello di informare, allarmare e anche entusiasmare il lettore. Come già accennato sopra, si parla qui di un campo che si basa su un nuovo paradigma, quello in cui la macchina impara dall'esperienza come fa il bambino, senza spiegarle alcuna conoscenza "a priori" (come si faceva un tempo). Oltretutto si tratta di un campo che – recentemente e in alcuni ambiti circoscritti – è diventato veloce, perché i nuovi risultati creano da subito le basi per i successivi, portando a una crescita esponenziale.

Proiettando coraggiosamente la nostra immaginazione, nel futuro il legislatore rischierebbe di non riuscire a stare al passo con queste tecnologie. Si pensi alla regolamentazione dello spazio aereo cittadino che bisognerebbe implementare per i droni intelligenti, possibili fattorini del futuro. Si pensi ai problemi etici che si pone chi lavora nell'ambito dello sviluppo di automobili che si guidano completamente da sole, le quali promettono un futuro con pochissimi incidenti stradali: chi avrebbe però la responsabilità di un eventuale incidente stradale tra due automobili autonome? Peggio ancora: di fronte all'alternativa tra il salvataggio del passeggero o quello della persona investita, quale scelta dovrà effettuare il veicolo intelligente un istante prima dell'impatto? Problemi ancora più stringenti si porranno quando le intelligenze artificiali potrebbero dare eccessivo potere a certi agenti, non importa se criminali o governativi. Su *YouTube* è possibile ad esempio trovare il video divulgativo *Slaughterbots* (fusione dell'inglese *to slaughter* – massacrare – e *robots*), ideato dal Professore Stuart J. Russell dell'Università di Berkeley, da 35 anni uno dei pionieri dell'intelligenza artificiale. Il video immagina uno scenario agghiacciante in cui tecnologie già esistenti – quella del riconoscimento facciale tramite intelligenza artificiale, quella dei droni e quella degli esplosivi a piccola scala – vengono combinate in un unico prodotto: robot volanti intelligenti in grado di scovare e uccidere i "cattivi" (un criminale ma anche un oppositore politico o un militante). Il video termina con un monito dello stesso Professore, che invita ad agire per rendere le intelligenze artificiali

un'opportunità per l'umanità, piuttosto che un pericolo. Un tema molto simile è stato anche brillantemente affrontato in alcuni episodi della serie tv *Black Mirror*, che propone allo spettatore scenari distopici in cui la tecnologia moderna interferisce troppo aggressivamente con la società. Per concludere, torniamo allora allo scenario di apertura immaginato da Asimov. Da esso scaturisce il quesito spontaneo che tutti gli addetti ai lavori (e non solo) si sono posti: l'esistenza concreta di una enorme intelligenza artificiale di "tipo forte" che tutto connette e controlla è effettivamente una minaccia per l'umanità e per l'indipendenza di quest'ultima? Saremo tutti ridotti in schiavitù da questa fantomatica intelligenza sovraumana? A questo proposito, riportiamo una citazione molto più pragmatica letta sul *Financial Times*: «La vera minaccia è costituita dagli esseri umani [...]. Gli umani hanno dimostrato che essi possono tramutare ogni innovazione e avanzamento tecnico in un'arma. Useranno sicuramente macchine intelligenti contro altri umani così come hanno usato pugnali, polvere da sparo, energia nucleare e virus informatici» (traduzione nostra, tratto da *Humans can take any innovation and turn it into a weapon. Letters to the Editor*¹⁷). Forse è vero che – parafrasando un articolo citato in questo saggio – ciò di cui l'umanità deve occuparsi maggiormente e con più urgenza non è tanto l'intelligenza artificiale quanto la stupidità naturale.

ABSTRACT. Artificial intelligence has been living in the last (few) years a golden age, full of developments. New algorithms and ingenious methodologies are overwhelmingly gaining great visibility, and the latest results can evoke disturbing scenarios but also great opportunities. This essay aims to analyse the impact and the effects of artificial intelligence on the contemporary society.

KEYWORDS: artificial intelligence, algorithms, contemporary society.

¹⁷ J. Eskenazi, *Humans can take any innovation and turn it into a weapon. Letters to the Editor*, Financial Times, 2016.