



CENTRO DI DOCUMENTAZIONE INTERDISCIPLINARE
DI SCIENZA E FEDE



SCUOLA INTERNAZIONALE SUPERIORE
PER LA RICERCA INTERDISCIPLINARE

Angelo MONTANARI

Artificial Intelligence and Machine Learning: fini e confini

A.A. 2022/23

TRIENNIO

***ORIGINS: Le grandi domande su cosmo, vita e intelligenza
nella scienza, nella filosofia e nelle culture***

ANNO III: L'origine dell'intelligenza e la domanda sulla coscienza

3 dicembre 2022

Documento n. 37

Testo ad uso dei partecipanti al seminario

Sin dai suoi primi passi, l'Intelligenza Artificiale (*Artificial Intelligence*, AI in breve) si è andata sviluppando lungo due direzioni significativamente diverse: l'AI simbolica, anche nota come "*model-based AI*", e l'AI sub-simbolica, anche nota come "*data-driven AI*". Il cuore dell'AI simbolica sono i metodi e le tecniche per la rappresentazione della conoscenza e per il ragionamento automatizzato, mentre l'AI sub-simbolica si basa sulle reti neurali, sugli algoritmi genetici e sul *machine/deep learning*. La prima adotta la logica matematica come il formalismo più consono per rappresentare la conoscenza umana, e i metodi logici come gli strumenti più appropriati per ragionare su quest'ultima; l'AI sub-simbolica, invece, mira ad imitare il comportamento del cervello umano e della sua complessa rete di neuroni interconnessi. Negli ultimi anni, la *data-driven AI* sta assumendo una posizione dominante grazie ai suoi innegabili successi in molti domini applicativi. Tuttavia, sussiste un equilibrio delicato tra l'efficacia e l'efficienza dell'AI sub-simbolica e la sua "interpretabilità" – vale a dire, la possibilità per l'essere umano di accedere facilmente ai processi computazionali che portano agli esiti del sistema di AI sub-simbolica (decisioni, classificazioni, previsioni). Tale mancanza di interpretabilità si rivela un importante problema, specialmente in contesti critici o che coinvolgono sicurezza e salute. Ciò ha promosso la nascita dell'area di ricerca nota come "explainable/trustworthy AI" (XAI, AI "esplicitabile" o affidabile).

La nascita dell'AI simbolica (*model-based*) si individua solitamente nell'articolo di Alan Turing sulla relazione tra sistemi computativi e intelligenza. Di fatto, in questo articolo Turing si interroga circa la possibilità di costruire un sistema di intelligenza artificiale, ma non indica esplicitamente la strada da percorrere per realizzarlo – pur se la stretta relazione che egli individua tra intelligenza e linguaggio assegna in effetti un ruolo fondamentale alla dimensione simbolica umana. Alcuni anni più tardi, Newell e Simon formularono la cosiddetta *ipotesi del sistema fisico simbolico* ("physical symbol system hypothesis"), che stabilisce che l'abilità di processare strutture di simboli (espressioni) è sufficiente per rendere intelligente un computer; inoltre, assunsero che lo stesso tipo di manipolazioni di simboli si trova alla base dell'intelligenza umana. Da allora, moltissimo lavoro è stato fatto per conferire ai sistemi artificiali la capacità di eseguire compiti avanzati che coinvolgono rappresentazioni simboliche e metodi di ragionamento, quali la pianificazione automatica, la calendarizzazione di attività, capacità diagnostiche e la configurazione di sistemi. Alcuni dei risultati più importanti di questa branca della ricerca in AI sono i "sistemi esperti" e i "sistemi multi-agente", ivi compresi i sistemi automatici avanzati e, specialmente, i sistemi robotici. Durante il seminario, vedremo una breve panoramica dei metodi e delle tecniche di AI simbolica, e svolgeremo un'analisi critica delle sue potenzialità e debolezze.

Negli ultimi anni, la *data-driven AI* sta assumendo un ruolo preponderante nell'universo dell'AI. Di fatto, il *machine learning* è stato presente sin dall'esordio dell'AI, ma la sua crescita e diffusione attuale è motivata da almeno due fattori cruciali: il forte aumento del potere di calcolo dei computer e la disponibilità nel web di ingenti quantità di informazioni (testi, immagini, registrazioni audio e video). La capacità di analizzare enormi archivi di dati (*big data*) per scoprire somiglianze significative tra loro e usarle per identificare nuovi dati dello stesso tipo – così come la possibilità di fare previsioni affidabili – ha indotto un radicale balzo in avanti nella ricerca e, soprattutto, nelle applicazioni. Riassumeremo brevemente i metodi e le tecniche di *machine learning* e *deep learning*, e discuteremo una serie di questioni sia pratiche che teoriche.

Riferimenti essenziali

Alan Turing, "Computing Machinery and Intelligence", *Mind* 59 (1950), 433-460
(disponibile a: <https://inters.org/files/computing-machinery-and-intelligence.pdf>)

Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning* (Adaptive Computation and Machine Learning Series), The MIT Press, Cambridge, MA and London, UK 2016

Angelo Montanari, *Per un vocabolario filosofico dell'informatica*, in G. Cicchese, A. Pettorossi, S. Crespi Reghizzi, e V. Senni (eds.), *Istanze Epistemologiche e Ontologiche delle Scienze Informatiche e Biologiche*, Città Nuova, Roma 2011, pp. 82-106.

Angelo Montanari, "Artificial Intelligence and Creativity", *Acta Philosophica*, 2023 (in pubblicaione per gennaio 2013).

Shai Shalev-Shwartz and Shai Ben-David, *Understanding Machine Learning. From Theory to Algorithms*, Cambridge University Press, 2014.