

Intelligenza artificiale? Una manna dal cielo

Ne parliamo con Crescenzo Tortora, astrofisico e ricercatore all'Inaf – Osservatorio di Capodimonte. Come istruiamo e usiamo gli algoritmi per interpretare i dati di Euclid sulla materia oscura

Alberto Agliotti

Autore di oltre 150 pubblicazioni scientifiche, Crescenzo Tortora si occupa principalmente di temi legati all'evoluzione delle galassie, con particolare attenzione alla materia oscura, al lensing gravitazionale e alle proprietà stellari e strutturali delle galassie, anche attraverso l'uso dell'intelligenza artificiale.

Ha lavorato con i dati provenienti dal telescopio VST (VLT Survey Telescope), ideato e costruito a Napoli e dal 2011 operativo in Cile, in particolar modo alla survey extra-galattica KiDS (Kilo Degree Survey), ed è coinvolto in diversi progetti legati all'Osservatorio Vera Rubin e al telescopio spaziale Euclid. È tra gli autori di uno studio che di recente ha dimostrato come l'IA sia in grado di imparare dalle simulazioni cosmologiche a misurare la massa delle galassie.

L'intelligenza artificiale sta cambiando il modo di fare ricerca in tutti i campi. L'astronomia non fa eccezione.

L'intelligenza artificiale a un certo punto ha invaso l'astronomia. Tempo fa mi sono divertito a cercare tra gli articoli scientifici. Fino alla metà degli anni Duemila quelli che nell'abstract e nelle parole chiave contenevano l'espressione "machine learning" erano tra 200 e 300 ogni anno. Nei 10 anni successivi sono raddoppiati, mentre negli ultimi anni il numero è salito a 5000. Se per certi versi può essere una moda, è inequivocabile che l'intelligenza artificiale in astronomia sia diventata fondamentale, perché offre una serie di strumenti che permettono di fare ricerca meglio di quanto possiamo fare con tecniche standard visuali. Abbiamo telescopi che andranno a osservare sempre più grandi aree di cielo, i dati che forniranno saranno sempre più numerosi, dunque servono necessariamente tecniche automatiche per riuscire ad analizzarli.

Come usate l'intelligenza artificiale nel vostro gruppo di ricerca?

Tra i vari progetti di cui ci occupiamo c'è la ricerca di nuove lenti gravitazionali nelle immagini astronomiche. La lente gravitazionale è il fenomeno che si genera quando la luce di una sorgente lontana passa nelle vicinanze di una distribuzione di massa più vicina, come una stella, una galassia o un ammasso di galassie, detta "lente". La gravità della lente deforma i percorsi della luce creando una specie di miraggio, deformando, amplificando e moltiplicando l'immagine della sorgente lontana. La "firma" di una lente gravitazionale, il tipo di immagine che genera, è nota. Conosciamo la teoria che la descrive, la Relatività generale. Possiamo quindi costruire un modello per generare lenti simulate che diamo in pasto a quelle che vengono chiamate "reti neurali convoluzionali", un tipo particolare di reti neurali che si ispirano alla corteccia visiva animale e che sono perfette per trovare forme nelle immagini. La rete neurale è in grado di apprendere in modo autonomo, stabilendo "regole" a partire dagli esempi.

Ha bisogno di "vedere" tantissimi esempi, per poter esplorare tutti gli esempi possibili di lenti gravitazionali. Alla fine dell'apprendimento, nelle immagini vere andrà a cercare tutto quello che ritiene corrispondere a una lente gravitazionale.

Il telescopio spaziale Euclid sarà una miniera d'oro nella scoperta di queste rare sorgenti astrofisiche. Ci aspettiamo di trovare circa 100.000 nuove lenti gravitazionali, cento volte di più rispetto a quelle che abbiamo scoperto fino a ora.

L'intelligenza artificiale può imparare da sola a identificare un tipo di oggetto?

Molto di quello che facciamo si basa sui cosiddetti "metodi supervisionati", in cui si



Alberto Agliotti, ingegnere di formazione, lavora nel campo della divulgazione scientifica come giornalista, redattore e traduttore

forniscono alla rete modelli di oggetti noti. Quello che restituisce la rete non sarà quindi molto diverso dal modello che ha ricevuto in pasto. Esistono però metodi non supervisionati, nei quali non si fornisce un esempio. La rete analizza un set di dati senza avere un indirizzamento e può individuare insiemi di oggetti raggruppati secondo parametri diversi. Alcuni colleghi stanno sperimentando queste tecniche per trovare nei dati quelli che in inglese sono chiamati "outlier" (sorgenti anomale), vale a dire oggetti con proprietà totalmente diverse rispetto a quelle della classe a cui pensiamo appartengano.

Che sia una ricerca supervisionata o non supervisionata, applicata a grandi moli di dati l'IA può permettere di scoprire oggetti intrinsecamente molto peculiari. Sarà probabilmente il caso di Euclid, che osserverà circa un terzo dell'intera volta celeste con una risoluzione spaziale mai sperimentata prima.

Se esistono oggetti intrinsecamente rari, su un'area così ampia possiamo trovarne in gran numero e studiarli. Per esempio, le galassie ellittiche massive tipicamente sono molto grandi perché hanno subito processi di *merging*, cioè di fusione con altre galassie. Ne esiste una piccolissima percentuale che pur avendo una massa grande sono di piccole dimensioni, 3-4 volte più piccole di tutte le altre galassie ellittiche di quella specifica massa. Perché? Perché per puro caso non hanno subito fusione e sono rimaste uguali a come erano quando si sono



CRESCENZO TORTORA

formate. Ne conosciamo poche perché finora abbiamo potuto studiare aree di cielo limitate, con bassa risoluzione spaziale. Con Euclid ne potremmo trovare tantissime.

▲ Crescenzo Tortora

L'IA sarà sempre più presente nella ricerca, ma non si potrà mai prescindere dall'intervento umano.

L'IA è uno strumento potentissimo, veloce ed efficiente, in grado di mostrare oggetti che si perderebbe troppo tempo a trovare, o che ►

Mela, una bilancia per le galassie

Uno studio di Inaf, in collaborazione con la Sun Yat-sen University, ha dimostrato che l'intelligenza artificiale può imparare dalle simulazioni cosmologiche a misurare la massa di galassie reali

Gli algoritmi e le applicazioni di intelligenza artificiale fanno ormai parte della nostra vita quotidiana. La comunità scientifica, tuttavia, ne fa largo utilizzo già da diversi anni e l'Italia è all'avanguardia nel settore. L'Istituto nazionale di astrofisica (Inaf) ha partecipato a un progetto guidato da Nicola R. Napolitano, da cinque anni presso la Sun Yat-sen University, in Cina, che per la prima volta è riuscito a dimostrare che l'intelligenza artificiale può imparare dalle simulazioni cosmologiche di formazione ed

evoluzione dell'universo a misurare correttamente la massa delle galassie. Lo studio, pubblicato sulla rivista «Astronomy & Astrophysics», descrive un nuovo metodo per stimare la massa delle galassie, incluso il contenuto di materia oscura, usando il *machine learning* e dimostra per la prima volta che il metodo funziona su cataloghi di galassie reali. Gli esperti hanno confrontato le stime del nuovo codice, denominato Mela (Mass Estimator machine Learning Algorithm), con stime di

procedure dinamiche classiche verificando che Mela può riprodurre con incredibile accuratezza le masse dei metodi classici, in alcuni casi molto più laboriosi e basati su dati molto più complessi. L'Inaf, e in particolare la sede di Napoli, vanta una lunga esperienza in materia di dinamica delle galassie, con la partecipazione a progetti nati sul solco della tradizione della fisica delle galassie. I ricercatori italiani, in particolare Crescenzo Tortora e Nicola Napolitano, sono diventati negli anni

specialisti a livello mondiale e collaborano con i gruppi di dinamica delle galassie più importanti nel contesto internazionale e con progetti unici al mondo. «Da questo lavoro abbiamo capito che l'intelligenza artificiale è pronta a imparare la fisica a partire dai dati», conclude Napolitano. «Nella fattispecie abbiamo verificato che Mela può utilizzare le leggi fisiche che conoscevamo, ma presto l'intelligenza artificiale potrà imparare anche la fisica che non conosciamo». (INAF)

non si riuscirebbe nemmeno a trovare perché non sarebbe possibile analizzare tutti i dati. Può anche offrire un cambio di prospettiva, far vedere le cose in un altro modo dando lo spunto per scoprire qualcosa di nuovo. Ma è sempre l'essere umano che deve analizzare i risultati. Nel nostro caso, per esempio, la rete neurale non riesce ancora a trovare un campione puro di lente gravitazionale, e dopo la fase di analisi automatizzata dobbiamo condurre un'ispezione visuale. Però ci aiuta a ridurre drasticamente il campione da analizzare, facilitando estremamente il nostro lavoro.

Oltre a scoprire nuovi oggetti, grazie all'intelligenza artificiale, si possono anche migliorare le conoscenze sui processi fisici che regolano l'universo.

Alcuni astronomi la usano proprio a questo scopo. Cercano di capire quale tra gli universi simulati si avvicina di più all'universo reale.

Nelle simulazioni, infatti, inseriamo le leggi fisiche che conosciamo molto bene, come la legge di gravità, ma anche altri parametri su cui invece abbiamo una serie di incertezze: come si addensa il gas, come i venti di supernova e i nuclei galattici attivi regolano la formazione stellare, o ancora la percentuale di materia ed energia oscura. Una variazione in questi parametri può portare a universi con galassie che hanno proprietà diverse, per esempio universi in cui ci sono galassie mediamente più massive, o meno galassie ellittiche e più galassie a spirale, o ancora un numero totale di galassie diverso.

Noi prendiamo i risultati delle diverse simulazioni e li confrontiamo con l'universo nel quale viviamo. Possiamo così fare una sorta di *reverse engineering* per individuare i valori più corretti dei parametri forniti come input alla simulazione e capire meglio come funzionano certi processi fisici.

Sebbene l'IA permetta di gestire una quantità sempre maggiore di dati, c'è la possibilità che non tutti i dati raccolti vengano analizzati in breve tempo, o sotto tutti i punti di vista.

Dobbiamo essere sicuri di poterli tramandare.

È così. Ogni gruppo di ricerca sviluppa codici per compiti specifici e analizzare i dati ottenuti da una certa area di cielo, ma non è detto che quel compito esaurisca tutto il potenziale offerto da quell'area. Bisogna conservare i dati perché qualcuno li possa analizzare in futuro.

Questo significa avere prima di tutto disponibilità di spazio fisico su dischi e server. In genere università ed enti di ricerca investono



ESA/HUBBLE NASA

▲ Lente gravitazionale: l'anello azzurro è un effetto dovuto alla galassia al centro che deforma la traiettoria della luce di una galassia molto più lontana alle sue spalle

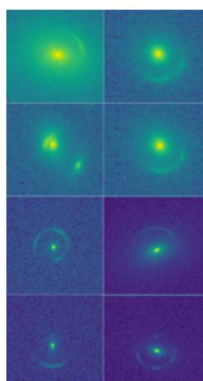
molto in questa direzione. C'è poi un problema di leggibilità dei dati, che devono essere agevoli da gestire e accessibili anche a versioni successive dei software. A questo scopo, in astronomia usiamo il formato Fits, che sta per "Flexible Image Transport System" (sistema flessibile di trasporto delle immagini). È un formato non proprietario, perfetto per la conservazione, lo scambio e l'analisi dei dati, che permette di includere nei file la grande quantità di parametri che caratterizzano un oggetto celeste come una stella o una galassia (coordinate, luminosità, colore, massa, distanza e così via), parametri che diversi programmi possono agilmente maneggiare. È così versatile da essere stato adottato di recente per digitalizzare 80.000 manoscritti della Biblioteca Vaticana.

Un altro aspetto da considerare è come processare i dati. Per l'intelligenza artificiale servono computer con prestazioni elevate, processori potenti ma anche schede grafiche all'altezza. Possiamo cavarcela con i nostri computer, ma per compiti molto dispendiosi servono macchine dedicate.

Quali competenze deve avere chi fa il tuo mestiere? Ti definiresti un astroinformatico?

L'astroinformatico è una figura che, a ragione, si sta facendo strada all'interno della comunità astrofisica, con nuove competenze per lo sviluppo dell'astronomia dei prossimi decenni.

Io non mi definisco un astroinformatico. Come molti di coloro che usano la *machine learning* in astronomia mi limito a scrivere il codice che sfrutta le strutture già pronte in linguaggi come Python. Pur lavorando molto con l'IA rimango un astrofisico: la mia attività principale è cercare galassie, studiarle, caratterizzarle, capire come si sono formate e così via. Un astrofisico che ha trovato nel *machine learning* una manna caduta dal cielo. ☺



▲ Lenti gravitazionali simulate per addestrare la rete neurale che le cercherà nelle immagini astronomiche

F. GENTILE ET AL., MNRAS, 2023