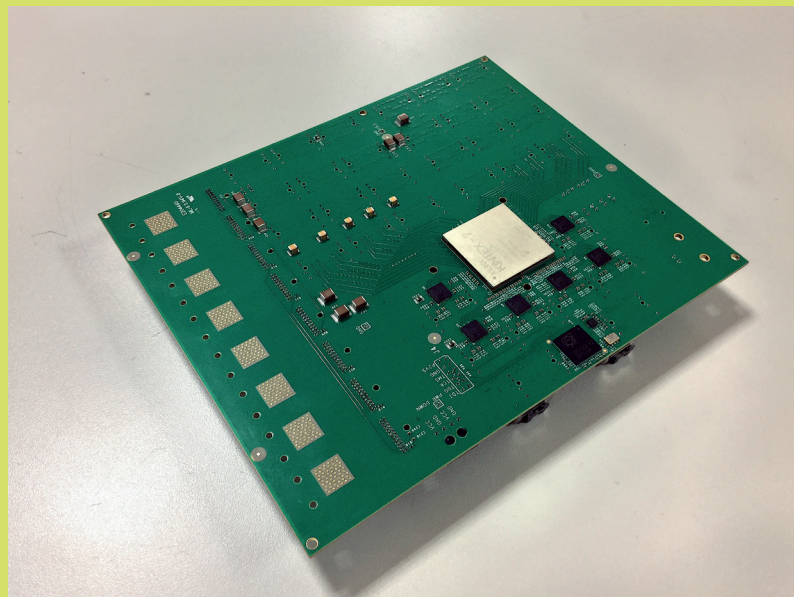


Caduti nella rete (neurale)

Gli algoritmi dell'intelligenza artificiale

di Stefano Giagu



a.
Prototipo di una scheda elettronica del sistema di trigger per muoni che verrà installato nell'esperimento Atlas. Nel centro della scheda è visibile il processore di tipo Fpga.

Il concetto di intelligenza artificiale esiste dalla metà degli anni '50 e lo studio e sviluppo di algoritmi di *machine learning*, che ne costituiscono il motore computazionale, è in corso da più di venti anni, ma è solo recentemente che le applicazioni hanno cominciato a diffondersi esponenzialmente in tutti i settori della società, incluso l'ambito della ricerca di base in fisica delle particelle. L'idea di "rete neurale artificiale" nasce nel 1943 come un modello di tipo connessionista, proposto dal neurofisiologo Warren McCulloch e dal logico matematico Walter Pitt, per descrivere il funzionamento dei neuroni nel cervello umano. Tale idea si è evoluta negli anni, grazie soprattutto all'algoritmo "*perceptron*" di Frank Rosenblatt, psicologo americano pioniere nella ricerca sui sistemi cognitivi alla Cornell University, in grado di imparare ad associare informazioni in input con risposte, in modo apparentemente simile all'apprendimento negli umani, e successivamente nei moderni algoritmi di apprendimento basati su "*backpropagation*", che costituiscono la spina dorsale delle moderne architetture di reti neurali artificiali in uso oggi. Le più recenti evoluzioni degli algoritmi di *machine learning* basate sul *deep learning* (in pratica, un qualunque algoritmo basato su una o più reti neurali) stanno trovando un terreno molto fertile nelle grandi collaborazioni al Large Hadron Collider (Lhc) del Cern, caratterizzate da apparati costituiti da milioni di elementi, i cui segnali producono enormi quantità

di dati complessi che devono essere analizzati e semplificati per poterne distillare l'informazione utile. Tradizionalmente il problema viene affrontato tramite una serie di passaggi, che prendono il nome di "analisi dati", che operano sugli eventi per ridurne la complessità. Questo approccio, che ha permesso scoperte fondamentali, segna però il passo di fronte all'aumento della complessità e della quantità dei dati prodotti dagli esperimenti, e in tali condizioni il *deep learning*, grazie alla capacità di apprendere e condensare le informazioni rilevanti, può fornire l'ingrediente necessario a colmare il divario tra l'approccio tradizionale e quello ottimale. Applicazioni di intelligenza artificiale in fisica delle particelle non si limitano all'analisi dati. La flessibilità e la velocità delle reti neurali, che possono essere eseguite in tempo reale su dispositivi di calcolo veloci come le Field Programmable Gate Array (Fpga) (vd. fig. a), possono essere utilizzate per rimpiazzare i sistemi hardware di selezione degli eventi prodotti nelle collisioni di Lhc, più costosi e complessi da progettare e mantenere (vd. in *Asimmetrie* n. 17 p. 33, ndr). Il *deep learning* viene utilizzato anche nel campo della ricostruzione "offline" degli eventi, per individuare le traiettorie delle particelle elementari prodotte nelle collisioni, per identificarne la tipologia e, più recentemente, per simularne il comportamento all'interno dei rivelatori di Lhc, un passo cruciale per confrontarsi con le previsioni dei modelli teorici.

Tra i diversi algoritmi di *machine learning*, i Boosted Decision Tree (Bdt) (vd. p. 19, ndr) hanno tradizionalmente avuto un ruolo di primo piano nell'analisi dati degli esperimenti per le intrinseche caratteristiche di semplicità e trasparenza. I Bdt combinano insieme un numero molto grande di algoritmi elementari chiamati “alberi di decisioni binarie”. Questi emulano la sequenza di domande/decisioni utilizzata per esempio nel gioco “Indovina Chi?” per cercare di individuare un personaggio segreto sulla base di una serie di domande sull'aspetto del personaggio, alle quali si possono ottenere solo risposte binarie del tipo “vero-falso”. Ad ogni domanda successiva il campione di possibilità (i personaggi nel gioco) si restringe sempre più fino idealmente ad arrivare alla risposta corretta. La stessa procedura è utilizzata dai fisici sperimentali quando cercano di identificare eventi appartenenti a un processo fisico interessante, per esempio eventi contenenti un bosone di Higgs che decade in due fotoni, tra i tantissimi eventi prodotti a Lhc. Gli eventi interessanti vengono selezionati tramite una sequenza di richieste su diverse grandezze fisiche (l'equivalente delle domande nel gioco), scelte in modo da massimizzare la purezza del segnale che si vuole selezionare. Quali grandezze fisiche utilizzare e quale tipo di richieste fare costituisce l'addestramento dell'algoritmo, che viene effettuato sulla base di campioni di eventi di esempio, spesso simulati in modo da riprodurre il segnale anche prima di averlo trovato. Un Bdt combina diversi alberi di decisione binaria per produrre un *ensemble* di classificatori più efficace e potente. Ancora più promettenti sono gli algoritmi basati sulle “reti neurali profonde”, in particolare le cosiddette “reti convoluzionali” o Cnn (Convolutional Neural Network), sviluppate nel campo della

computer vision per il riconoscimento di immagini fotografiche. Le Cnn sono specializzate nell'identificazione di caratteristiche geometriche ricorrenti presenti nelle immagini con cui vengono addestrate. Caratteristiche che possono essere combinate in strutture complesse, per esempio volti di persone (vd. in Asimmetrie n. 20 approfondimento p. 33, ndr) o specifici oggetti, animali, ecc., con lo scopo di riconoscerle successivamente in altre immagini. Molti dei segnali ricostruiti dai rivelatori degli esperimenti di fisica delle particelle si prestano a essere rappresentati come immagini fotografiche, in cui ad ogni pixel dell'immagine corrisponde per esempio l'energia rilasciata da una particella in una delle zone sensibili del rivelatore. Una rete Cnn può quindi essere addestrata a riconoscere caratteristiche specifiche nelle “immagini” dei rivelatori, e quindi per esempio a identificare in tali immagini la presenza di un decadimento del bosone di Higgs, al pari di quanto fa una Cnn in uno smartphone per riconoscere il volto del proprietario. Una variante interessante è costituita dalle “reti neurali ricorrenti” o Rnn (Recurrent Neural Network), algoritmi sviluppati per l'elaborazione del linguaggio naturale. Esempio tipico è il funzionamento dei traduttori online, dove la traduzione di una frase viene raffinata man mano che si aggiungono parole (vd. p. 19, ndr). Analogamente, le Rnn analizzano ogni elemento di una sequenza di dati di input mantenendo “memoria” di ciò che è stato calcolato precedentemente, e sono molto efficaci nell'elaborare le lunghe sequenze di dati correlati che compaiono in molti problemi in fisica delle particelle. Tipico esempio è quello dell'identificazione di quark beauty prodotti nelle interazioni di alta energia di Lhc. I quark beauty si presentano sotto forma di



b.
Il gioco “Indovina Chi?” ci fornisce un esempio di “albero di decisioni binarie”: per cercare di indovinare il personaggio segreto dell'avversario, il giocatore fa una serie di domande sull'aspetto del personaggio, scartando man a mano i personaggi non pertinenti (p.es. i personaggi con gli occhiali o con i capelli biondi, ecc.).



un flusso collimato di particelle, chiamato “getto adronico”, che vive per un certo tempo prima di decadere in altre particelle. Identificare questi getti adronici è molto impegnativo dal punto di vista computazionale per cui negli algoritmi tradizionali si introducono semplificazioni che ne limitano le prestazioni. L’uso di Rnn ha permesso di migliorare sostanzialmente la precisione con cui i quark beauty vengono identificati negli esperimenti di Lhc (in particolare in Atlas e Cms), estendendone la sensibilità di scoperta per effetti di nuova fisica.

Infine, un’altra classe di reti neurali chiamate “reti generative avversarie” o Generative Adversarial Network (Gan) (vd. fig. c), possono essere sfruttate per simulare in modo preciso e veloce il comportamento atteso dei rivelatori in presenza di diversi segnali fisici interessanti. Una Gan è fatta da due reti neurali profonde, il generatore e il discriminatore. La prima genera immagini, mentre la seconda riconosce un’immagine reale da una generata. Le due reti vengono addestrate “affrontandosi” l’una contro l’altra. In pratica il discriminatore cerca di identificare le differenze tra le immagini di un campione vero (per esempio immagini ricostruite dai segnali di un rivelatore colpito da un dato tipo di particelle) e quelle generate dal generatore, mentre quest’ultimo cerca di ingannare il discriminatore nell’accettare le sue immagini, e così facendo impara a generare immagini realistiche. I risultati ottenuti con questa tecnica sono molto incoraggianti e, se le promesse iniziali verranno confermate, l’utilizzo delle Gan permetterà in futuro di ridurre sostanzialmente il costo computazionale necessario per simulare l’enorme mole di eventi richiesta per la fase di funzionamento ad altissima intensità dell’acceleratore Lhc.

c. Immagini di persone non esistenti, generate da una rete neurale di tipo Gan.

Biografia

Stefano Giagu insegna alla Sapienza Università di Roma e svolge attività di ricerca in fisica delle particelle elementari e nello sviluppo di metodi di intelligenza artificiale e *machine learning* in fisica di base e applicata. Ha partecipato agli esperimenti L3 al Cern e Cdf al Fermilab ed è membro della collaborazione Atlas, al Large Hadron Collider del Cern di Ginevra, della quale dal 2019 è responsabile nazionale.

Link sul web

https://it.wikipedia.org/wiki/Apprendimento_automatico
https://en.wikipedia.org/wiki/Artificial_neural_network
https://en.wikipedia.org/wiki/Deep_learning
<https://cerncourier.com/the-rise-of-deep-learning/>
<https://www.datascience.com/blog/convolutional-neural-network>

DOI: 10.23801/asimmetrie.2019.27.3