

Sistemi di raccomandazione:

Intelligenza Artificiale, Deep Learning e personalizzazione dei contenuti

Paolo Casagrande, Sabino Metta
Rai - Centro Ricerche, Innovazione Tecnologica e Sperimentazione

Anno 2020... al giorno d'oggi i *sistemi di raccomandazione*, i cosiddetti *Recommender Systems* (di seguito *RS*), non hanno più bisogno di presentazioni. Nel nostro precedente articolo [1], abbiamo già introdotto e descritto le tecnologie per raccomandare all'utente finale il contenuto per lei/lui più rilevante. Tali tecnologie, che afferiscono al dominio dell'*Intelligenza Artificiale (AI)*, sono in continua evoluzione. Se da un lato permettono di ridurre il sovraccarico informativo a cui ogni utente è quotidianamente esposto, dall'altro lato tali tecnologie rappresentano, per il fornitore di contenuti (il cosiddetto *content provider*), lo strumento necessario per valorizzare la propria offerta e quindi difendere il proprio guadagno.

Dal "lontano" 2009, anno in cui **Netflix** ha assegnato un premio da *1 Milione di dollari* al miglior algoritmo in grado di predire le valutazioni degli utenti per i film, la comunità scientifica internazionale ha esplorato nuovi algoritmi e soluzioni per migliorare la previsione delle preferenze degli utenti.

I sistemi di raccomandazione svolgono l'importante funzione di filtrare e personalizzare in modo automatico le informazioni permettendo in questa maniera di far fronte al sovraccarico informativo a cui ogni utente è quotidianamente esposto.

Recentemente, la ricerca internazionale sta sperimentando l'utilizzo di tecnologie di apprendimento profondo, anche conosciuto come deep learning, per accrescere le potenzialità dei sistemi tradizionali di raccomandazione.

L'articolo presenta brevemente il funzionamento, la classificazione ed i limiti di quest'area pervasiva dell'Intelligenza Artificiale.

Secondo la IDC (*International Data Corporation*), la prima società mondiale specializzata in ricerche di mercato, la spesa che ruota attorno alle tecnologie di AI raggiungerà entro il 2023 i 100 miliardi di dollari.

Nel 2019 circa 2 miliardi di dollari sono stati spesi nel dominio dei sistemi di raccomandazione. Tra i vari approcci, negli ultimi anni un fiorente filone di ricerca sta sperimentando l'utilizzo delle tecnologie dell'apprendimento profondo (*deep learning*) e delle reti neurali profonde (*deep neural networks*) per superare alcuni limiti e alcune criticità dei sistemi tradizionali di raccomandazione.

SISTEMI DI RACCOMANDAZIONE TRADIZIONALI

In generale, i sistemi di raccomandazione rappresentano un'area dell'Intelligenza Artificiale e comprendono tecniche e strumenti software in grado di suggerire all'utente gli oggetti, e quindi anche i contenuti, più rilevanti tra i tanti a disposizione.

Tali sistemi non intendono necessariamente sostituirsi alla raccomandazione classica dove, per fare un esempio a noi più vicino, i suggerimenti vengono scelti e creati manualmente dall'editore del canale televisivo valendosi della competenza di professionisti esperti. In un RS, il contenuto più rilevante per un determinato utente viene inferito sulla base dei dati a disposizione ed attraverso l'implementazione di specifiche logiche. Per avere maggiori dettagli, raccomandiamo la lettura di un nostro precedente articolo [1] e, potendo approfondire, del classico di Ricci e altri [2].

Nella sua forma più semplice un sistema di raccomandazione produce, per ogni utente, una lista di oggetti e, per ciascuno di questi oggetti, una valutazione dell'utilità (*ranking*). Un oggetto può essere rappresentato da un qualunque contenuto multimediale e/o servizio digitale che possa essere in qualche modo suggerito all'utente finale: ad esempio, il prodotto di un supermercato, un video, un programma radiofonico, una canzone, una notizia.

I sistemi di raccomandazione possono essere classificati sulla base della specifica logica e dello specifico algoritmo attraverso i quali producono la lista finale di oggetti:

- *Collaborative Filtering (CF)*. I suggerimenti per uno specifico utente vengono generati sulla base di giudizi o valutazioni (*ranking*) che altri utenti, in passato, hanno dato a determinati oggetti a disposizione.
- *Content-based Filtering (CB)*. Il RS suggerisce ad uno specifico utente una lista di oggetti simili ad altri che l'utente ha scelto in passato. In questa categoria vengono confrontati gli oggetti stessi o le loro descrizioni. Rispetto alle metodologie *collaborative filtering*, i metodi *content-based* risultano essere computazionalmente più rapidi ed interpretabili, è cioè più facile capire il motivo di una specifica raccomandazione. Inoltre, gli algoritmi CB possono essere facilmente adattati a nuovi oggetti e/o utenti.
- *Demographic*. Il suggerimento del RS si basa sulla caratterizzazione demografica dell'utente, ad esempio l'età, il sesso, la residenza.
- *Knowledge-based*. I suggerimenti si basano su regole derivate da una conoscenza e descrizione esplicita dell'utente o degli oggetti (ad esempio, avendo conoscenza della mia preferenza verso la tecnologia, mi vengono consigliati articoli di quel tipo).
- *Community-based* o *Social*. I suggerimenti tengono conto delle scelte e delle valutazioni degli amici dell'utente sui *social networks* [2].
- *Hybrid*. Le tecniche descritte sopra vengono combinate per ottenere suggerimenti più precisi, cercando di far fronte alle debolezze di ciascun metodo. Molti RS commerciali sono di questo tipo.
- *Context-Aware (CARS)*. I suggerimenti tengono in conto il contesto [3]. Il contesto è definito da svariati fattori quali stato emotivo, attività, posizione geografica, condizioni atmosferiche, e qualsiasi altra cosa concorra alla definizione dello stato dell'ambiente, dell'utente, dell'oggetto o dell'interazione tra essi.

Lo ricordiamo, tali logiche di raccomandazione sono automatiche ed implementate attraverso precisi algoritmi. Recentemente la ricerca internazionale ha intrapreso una promettente sperimentazione delle tecniche di *deep learning* per risolvere alcune criticità presenti negli approcci tradizionali.

SISTEMI DI RACCOMANDAZIONE BASATI SUL DEEP LEARNING

Negli ultimi tempi le *reti neurali* e il *deep learning* (DL) sono diventati tecnologie emergenti in grado di risolvere compiti molto complessi. Importanti risorse economiche sono state investite per studiare l'applicabilità di tali tecnologie a diversi ambiti e casi d'uso (medico, manifatturiero ecc.). Sono nati, così, i sistemi di raccomandazione basati sul deep learning, i cosiddetti *Deep Learning based Recommender Systems* (DLRS). Rispetto ai tradizionali RS, i sistemi di raccomandazione basati sul deep learning sono riusciti a raggiungere una notevole precisione e presentano sicuramente enormi potenzialità [3][4][5].

I sistemi di raccomandazione basati sul deep learning presentano almeno quattro punti di forza rispetto ai sistemi tradizionali:

- *capacità di modellazione delle nonlinearità* eventualmente presenti nei dati attraverso funzioni di attivazione non lineari (rettificatore, sigmoide, tangente iperbolica, ecc.). In altre parole, una *modellazione nonlineare* è in grado di superare l'ipersemplificazione introdotta dai modelli lineari (tra questi la *fattorizzazione di matrice*). Questo permette di catturare in maniera più efficace le relazioni complesse ed intricate tra utenti (*users*) ed oggetti (*items*) nonché di individuare caratteristiche latenti di utenti e oggetti, vedi [4];
- *capacità di apprendimento* di caratteristiche rappresentative ed esplicative nei dati di input. Oggigiorno, la disponibilità di app e servizi in internet fornisce un'ampia base dati dove è possibile recuperare preziose informazioni relative ad utenti e/o ad oggetti (eventualmente raccomandabili). In tal senso, le tecnologie di

DL possono essere impiegate per estrarre tali informazioni ed in questa maniera potenziare gli algoritmi di raccomandazione. Oltre ad alleviare una eventuale attività manuale di estrazione di informazioni, le tecnologie DL permettono di costruire algoritmi di raccomandazione in grado di combinare fonti eterogenee (testo, immagini, audio, video);

- *capacità di modellazione sequenziale*. Le tecnologie DL giocano un ruolo fondamentale nell'individuazione di strutture sequenziali presenti all'interno dei dati. Tra le varie applicazioni menzioniamo la *traduzione automatica* e la *comprensione del linguaggio naturale*. Per quanto riguarda i RS, la modellazione sequenziale è importante per intercettare la dinamica di comportamento dell'utente e l'evoluzione di cambiamento di determinati oggetti;
- *alta flessibilità*. Oggigiorno, è possibile accedere a numerose piattaforme di sviluppo di tecnologie DL (citiamo ad es. **TensorFlow**, **Keras**, **PyTorch**, ecc.). Tali piattaforme sono costruite in maniera modulare rendendo così maggiormente efficienti le attività di sviluppo ed integrazione di differenti modelli di raccomandazione.

Di seguito accenniamo brevemente ai *principali metodi di deep learning* che possono essere utilizzati nei sistemi di raccomandazione:

- *Autoencoder*. Il loro scopo è di replicare l'input sull'output passando attraverso una rappresentazione più semplice e possono essere impiegati per ridurre la dimensionalità e la complessità del problema o colmare i vuoti di una matrice sparsa.
- *Restricted Boltzmann Machines (RBM)*. Le *Macchine di Boltzmann Ristrette* sono rappresentate da grafi non direzionali caratterizzati da un solo layer visibile ed uno nascosto. In ambito accademico tali tecnologie sono state utilizzate con successo per ottenere raccomandazioni di tipo *collaborative filtering*.
- *Recurrent Neural Network (RNN)*. Le *Reti Neurali Ricorsive* trovano applicazione nella traduzione automatica o nel riconoscimento vocale, e in generale per modellare dati sequenziali dal

momento che elaborano sia un input sia uno stato dipendente dai precedenti dati. Le *RNN* sono state proposte per *RS location-based*, dipendenti cioè dalla sequenza di spostamenti dell'utente;

- *Convolutional Neural Network (CNN)*. Le *Reti Neurali Convoluzionali* trovano, invece, vasta applicazione nel riconoscimento di immagini e, per questo, nei *RS* possono trovare applicazione nella ricerca di *features degli oggetti* (ad esempio descrizione con parole chiave di immagini da usare nel content-based filtering);
- *Deep Belief Network (DBN)*. Le *Reti di Credenze Profonde* utilizzano *RBM* come componenti di base e sono state usate per estrarre *features* della musica, quindi possono trovare utilizzo nel content-based filtering.

In generale, è possibile costruire *RS* che integrino tecnologie classiche e deep learning, oppure *RS* basati interamente su deep learning. La parte di deep learning può, a sua volta, essere basata su un solo tipo di tecnologia in modo monolitico (*deep single model*), oppure integrare diverse tecnologie di deep learning che si completino a vicenda (*deep composite model*).

Nel seguito chiariamo intuitivamente l'utilizzo del *deep learning* nelle diverse classi di sistemi di raccomandazione:

- *Content-Based Filtering*: il *DL* permette di descrivere gli item da raccomandare estraendo caratteristiche salienti e, anche, di catturare relazioni non lineari tra utenti ed items.
- *Collaborative Filtering*: alcuni ricercatori hanno proposto metodi basati sul *DL*, in particolare per far fronte al problema della sparsità della matrice o del *cold start* nel collaborative-filtering. Ad esempio, gli *RNN* sono stati usati per dimostrare la capacità di ottenere raccomandazioni basate non solo su una preferenza dell'utente, ma anche sulla sequenza temporale delle sue preferenze.
- *Hybrid*: viene utilizzata la capacità del *DL* di estrarre caratteristiche salienti dagli item o dagli utenti e potenziare il collaborative-filtering o

il content-based filtering alla base dell'hybrid recommender.

- *Context-aware*: i *context-aware recommenders (CARS)* hanno avuto recentemente un altissimo interesse per la capacità di incorporare il contesto nelle raccomandazioni. Il *DL* può essere utilizzato per estrarre gli elementi latenti del contesto e usarli per migliorare la raccomandazione.
- *Community based/Social Recommender Systems*: utilizzano la rete sociale degli utenti per migliorare le raccomandazioni, ad esempio le relazioni di fiducia, le reazioni degli utenti e la loro collocazione spazio-temporale. Il *DL* può essere inserito per modellare queste caratteristiche.

Un esempio di *DLRS* molto promettenti riguarda i raccomandatori fra domini diversi (*cross-domain*), in particolare quelli che utilizzano la conoscenza di un dominio per riversarla su un altro, che possono beneficiare della capacità dei metodi di deep learning.

Attualmente la ricerca sui *DLRS* sta cercando di risolvere alcuni problemi. La *scalabilità* del *DLRS* è importante per le applicazioni pratiche: si cerca di ridurre il numero di parametri oppure di comprimere la complessità dei dati. Un altro aspetto critico nel deep learning in generale, ma importante in special modo per i media pubblici, è la *trasparenza dell'algoritmo*, cioè la possibilità di spiegare le ragioni per un certo suggerimento: la ricerca sta lavorando anche su questo.

Oltre alle potenzialità di questi metodi, è necessario tenere in conto che il vantaggio pratico nell'impiego dei *DLRS* è tutt'altro che chiaro e che i risultati della ricerca sono a volte difficilmente riutilizzabili. Nei casi analizzati da un articolo molto discusso presentato alla *Conferenza RecSys 2019* [6], i risultati ottenuti con *DL* spesso non sono ripetibili, cioè non è immediato ricostruire l'esperimento con risultati simili: meno della metà è risultata ripetibile ed inoltre, nei casi rimanenti, il problema poteva essere risolto meglio con algoritmi classici e computazionalmente meno complessi. Le potenzialità dei *DLRS* sono grandi ma la ricerca dovrà lavorare ancora per coglierle appieno.

CONCLUSIONI

L'articolo introduce il concetto di *sistema di raccomandazione (RS)*, mostrando le ragioni della sua pervasività. Ad una introduzione al funzionamento e alla classificazione dei *RS* segue una panoramica dei più recenti sviluppi nell'applicazione delle *reti neurali* e del *deep learning (DL)* ai sistemi di racco-

mandazione. Mentre i risultati sono promettenti e in futuro le tecnologie di *DL* potrebbero risultare decisive per accrescere l'efficacia o i campi applicativi delle raccomandazioni, alcuni problemi sono tutt'ora aperti e oggetto di ricerca e di un vivace dibattito.

BIBLIOGRAFIA

- [1] P. Casagrande e S. Metta, *Leggi questo articolo, una tua amica lo ha trovato interessante*, in "Elettronica e Telecomunicazioni", Anno LXV, n. 2/2016, 2016, pp. 27-37, <http://www.crit.rai.it/eletel/2016-2/162-4.pdf>
- [2] F. Ricci, L. Rokach e B. Shapira, *Recommender systems: introduction and challenges*, in F. Ricci, L. Rokach e B. Shapira (ed) "Recommender Systems Handbook", Springer US, 2015, pp. 1-34, DOI: [10.1007/978-1-4899-7637-6_1](https://doi.org/10.1007/978-1-4899-7637-6_1)
- [3] G. Adomavicius ed altri, *Context-Aware Recommender Systems*, in "AI Magazine", vol. 32, n. 3, 2011, pp. 67-80, DOI: [10.1609/aimag.v32i3.2364](https://doi.org/10.1609/aimag.v32i3.2364)
- [4] R. Mu, *A survey of recommender systems based on deep learning*, in "IEEE Access", vol. 6, 2018, pp. 69009-69022, DOI: [10.1109/ACCESS.2018.2880197](https://doi.org/10.1109/ACCESS.2018.2880197)
- [5] Z. Batmaz ed altri, *A review on deep learning for recommender systems: challenges and remedies*, in "Artificial Intelligence Review", vol. 52, n. 1, 2018, pp. 1-37, DOI: [10.1007/s10462-018-9654-y](https://doi.org/10.1007/s10462-018-9654-y)
- [6] M. Ferrari Dacrema, P. Cremonesi e D. Jannach, *Are we really making much progress? A worrying analysis of recent neural recommendation approaches*, in "RecSys '19: Proceedings of the 13th ACM Conference on Recommender Systems", 2019, pp. 101-109, DOI: [10.1145/3298689.3347058](https://doi.org/10.1145/3298689.3347058)

