

Dai dati alla conoscenza: Machine Learning e Data science



Negli ultimi anni vi è grande fermento ed attenzione al mondo dei dati sia per quanto riguarda la loro raccolta e memorizzazione, sia per quanto riguarda il loro utilizzo.

di Massimo Nannini

Sempre più è evidente la presa di coscienza dell'importanza dei dati a supporto delle decisioni umane. Grazie a tecnologie emergenti come i Big Data Analytics, che si focalizzano sull'analisi di enormi quantità di dati tramite le tecniche di apprendimento basate sull'intelligenza artificiale quale il Machine Learning, per i data scientist è diventato più semplice e veloce comprendere, esplorare ed estrarre valore dalla mole di dati che la digitalizzazione dei processi aziendali genera quotidianamente.

Che cosa è il Machine Learning?

Il Machine Learning (ML) è un sottoinsieme dell'intelligenza artificiale (AI) che si occupa di creare sistemi che apprendono o migliorano le loro performance in base ai dati che utilizzano. Intelligenza artificiale è un termine generico e si riferisce a sistemi o macchine che imitano l'intelligenza umana. Spesso vengono utilizzati insieme e in modo interscambiabile, ma non hanno lo stesso significato. Un'importante distinzione: tutto ciò che riguarda il ML rientra nell'intelligenza artificiale ma l'intelligenza artificiale non include solo il ML. Più specificatamente, il ML è una tecnica di intelligenza artificiale che sfrutta il calcolo computazionale per eseguire elaborazioni complesse utili a simulare il modo in cui un essere umano apprenderebbe nelle medesime circostanze.

Perché è importante il Machine Learning?

Attualmente il ML è utilizzato in moltissimi ambiti della nostra vita quotidiana. Quando acquistiamo online, utilizziamo i social media o più semplicemente facciamo delle ricerche sul web, vengono utilizzati gli algoritmi di ML per rendere la nostra esperienza efficiente, facile e sicura. Oltre a questo si stanno sempre più affermando applicazioni aziendali ove il ML è utilizzato per supportare decisioni strategiche di prodotto e di marketing basate sui comportamenti dei propri clienti, ma anche decisioni sull'evoluzione dell'organizzazione aziendale basate sui dati generati in tempo reale dai processi produttivi.

Come funziona il Machine Learning?

Affinché i dati producano informazioni corrette e dunque valore misurabile per le organizzazioni è necessario applicare a loro un processo analitico e strutturare un vero e proprio progetto di ML caratterizzato dalle seguenti fasi:

- Raccolta dei dati
- Preparazione dei dati
- Esplorazione dei dati
- Apprendimento dei dati
- Validazione del modello e test

Come si evince chiaramente dall'elenco sopra riportato alla base del progetto di ML ci sono i dati, che devono essere raccolti e memorizzati in database, ripuliti ed organizzati in modo da comprenderne meglio il significato. Da qui poi il successivo passo dell'apprendimento che è il più importante del processo, ove si opera una rielaborazione detta preprocessing, utile ad ottimizzare le proprietà statistiche, a selezionare gli aspetti ed i fenomeni più significativi, ed infine la trasformazione per adattare il formato dei dati agli algoritmi che dovranno poi utilizzarli.

Infine la fase di validazione, che permette di verificare il corretto funzionamento del modello di ML rispetto ai dati utilizzati ed ai risultati attesi, in altre parole questa fase permette di verificare se il sistema di ML è abbastanza istruito per eseguire analisi predittive su nuovi set di dati.

Massimo Nannini

Ingegnere elettronico e consulente di impresa
info@gemaxconsulting.it