

4 elementi chiave da considerare per l'implementazione dell'IA

I leader aziendali che vogliono iniziare a utilizzare l'intelligenza artificiale (IA) o incrementarne l'utilizzo hanno tanti dubbi e domande a riguardo. Per esempio, si chiedono come utilizzare l'IA per migliorare l'efficienza operativa e far crescere la propria attività, mantenendo al contempo la stabilità e riducendo i rischi. Di seguito elenchiamo 4 elementi da considerare, a prescindere dal livello di esperienza della tua azienda con l'implementazione dell'IA.

1 Scegli una piattaforma flessibile

Una piattaforma scalabile e flessibile è fondamentale per supportare l'ottimizzazione dei tuoi deployment di IA. Scegliere la piattaforma giusta è importante sia per chi è alle prime armi che per chi ha già esperienza con l'IA.

Se stai muovendo i primi passi con questa tecnologia, opta per un'infrastruttura stabile, flessibile e aperta, che consenta al tuo team di creare ed eseguire piccoli carichi di lavoro. La piattaforma scelta dovrebbe supportare l'addestramento scalabile dei modelli, in modo che si possa adattare all'evoluzione dell'IA e alla crescita dei carichi di lavoro.

Una piattaforma stabile permette ai data scientist e agli sviluppatori di creare modelli che offrono vantaggi effettivi ai clienti. Questi ultimi, a loro volta, devono sentirsi sicuri del fatto che la piattaforma è affidabile e protetta.

Se la tua azienda ha già un po' di esperienza con l'IA, la soluzione giusta è una tecnologia di supporto che consenta di sfruttare al massimo il potenziale dell'intelligenza artificiale. Potrebbe quindi essere necessario fare investimenti importanti nell'acquisto di unità di elaborazione grafica (GPU) e altri acceleratori hardware.

Una piattaforma aperta e scalabile può:

- ▶ offrire alla tua azienda una buona base per la creazione e l'innovazione accelerate di applicazioni che sfruttano l'IA.
- ▶ consentirti di velocizzare il deployment delle applicazioni che sfruttano l'IA.
- ▶ permetterti di ottenere un vantaggio competitivo in un settore in continua evoluzione come quello dell'IA.

Con la giusta infrastruttura, la tua organizzazione sarà pronta per qualunque innovazione nel mondo dell'intelligenza artificiale.

2 Preparati per l'IA ibrida

L'intelligenza artificiale generativa (IA gen) ha incrementato la domanda per le risorse di calcolo e spinto l'elaborazione all'edge alla rete. Gli scenari di utilizzo dell'IA generativa sfruttano sia risorse cloud che on premise, ma alcuni carichi di lavoro (soprattutto quelli che presentano problemi di conformità o data gravity) devono essere eseguiti solo on premise. Con l'IA ibrida, il tuo team può eseguire i carichi di lavoro nella posizione che ritiene migliore, anche vicino all'origine dei dati.

L'IA ibrida offre diversi vantaggi, tra cui:

- ▶ **Flessibilità.** L'esecuzione di carichi di lavoro sperimentali e di produzione nell'ambiente più adatto consente di risparmiare sui costi.
- ▶ **Distribuzione più rapida.** L'inferenza di piccoli carichi di lavoro all'edge accelera la distribuzione di informazioni agli utenti.
- ▶ **Rischio ridotto.** L'esecuzione all'edge delle applicazioni basate sull'IA riduce il bisogno di trasferire grandi quantità di informazioni al core, con un conseguente rafforzamento della protezione dei dati.

Grazie all'IA ibrida, il team IT può addestrare un modello in un data center core o nel cloud pubblico e distribuirlo agli endpoint, dove può inferire e fornire informazioni utili più velocemente. I carichi di lavoro intensivi possono essere sparsi nel cloud centrale e all'edge. In questo modo, si crea un equilibrio tra la necessità di potenza di elaborazione e la richiesta di suggerimenti rapidi e precisi.

3 Vai oltre il primo modello

Probabilmente i tuoi team IT possono produrre facilmente un singolo modello di IA, ma per ottenere buoni risultati spesso la scalabilità è indispensabile. Per ottimizzare il valore di questa tecnologia per l'azienda, i tuoi team devono poter eseguire il deployment di più versioni di quel modello, anche centinaia di volte al giorno.

L'implementazione delle operazioni di apprendimento automatico ([MLOps](#)) può aiutare i tuoi team a ridimensionare facilmente i deployment. MLOps, al pari di DevOps, riunisce sviluppatori, responsabili delle operazioni e data scientist per accelerare lo sviluppo e il deployment efficace delle applicazioni basate sull'IA. Con MLOps, la tua azienda può adattare più facilmente lo sviluppo delle applicazioni, consentendoti di offrire ai tuoi clienti più servizi in meno tempo.

MLOps può essere un metodo di sviluppo conveniente ed efficiente. Se collaborano insieme su un'unica piattaforma, i team possono diventare più efficienti, così da consentire il risparmio di tempo e denaro e quindi l'aumento dell'utile netto dell'azienda.

4 Metti la sicurezza al primo posto

Per fronteggiare l'evenienza di costosi attacchi informatici, i dirigenti aziendali cercano sempre più spesso dei metodi che consentano di proteggere i dati sensibili e ridurre i rischi. L'IA generativa rappresenta un'ulteriore spinta ad accelerare, perché con il suo utilizzo aumenta il rischio che vengano rivelate informazioni confidenziali.

Mantenere un elevato livello di sicurezza è fondamentale per evitare diverse minacce, tra cui:

- ▶ **Data poisoning.** Una forma di attacco che immette dati sbagliati o dannosi nei modelli di IA, che finiscono così per dare informazioni scorrette o suggerimenti non attendibili.
- ▶ **Furto dei modelli.** Si tratta di un'operazione di reverse engineering su un modello di apprendimento automatico e l'estrazione dei relativi dati tramite una copia.
- ▶ **Backdoor attack.** Del codice dannoso viene nascosto in un modello di IA, da cui l'autore dell'attacco ruba le informazioni.

Questi e altri tipi di attacchi rappresentano una minaccia importante per la tua azienda, che corre rischi legali e finanziari. Riduci al minimo le vulnerabilità associate all'IA con il supporto di una piattaforma applicativa affidabile, con un'efficacia comprovata e procedure collaudate per gestire i rilasci e le patch di sicurezza.

Scopri di più

Accelera la strategia per l'IA. [Esplora](#) le soluzioni di IA di Red Hat.



Informazioni su Red Hat

Red Hat consente la standardizzazione in diversi ambienti e lo sviluppo di applicazioni cloud native, oltre a favorire l'automazione, la protezione e la gestione di ambienti complessi grazie a [pluripremiati](#) servizi di consulenza, formazione e supporto.

f facebook.com/RedHatItaly
X twitter.com/RedHatItaly
in linkedin.com/company/red-hat

ITALIA
it.redhat.com
italy@redhat.com

**EUROPA, MEDIO ORIENTE,
E AFRICA (EMEA)**
00800 7334 2835
it.redhat.com
europa@redhat.com