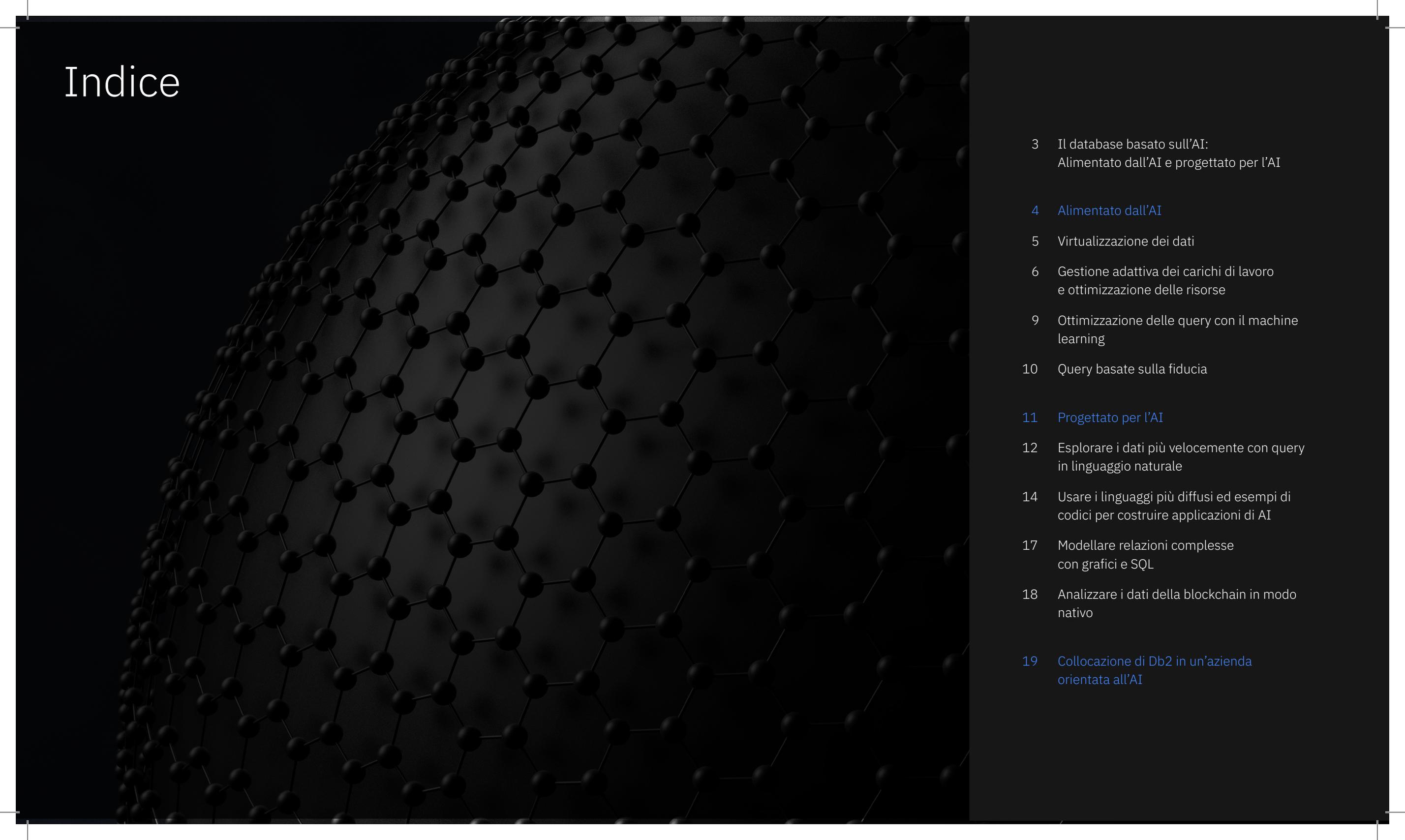


Db2

Il database
basato
sull'AI

Indice



- 3 Il database basato sull'AI:
Alimentato dall'AI e progettato per l'AI
- 4 [Alimentato dall'AI](#)
- 5 Virtualizzazione dei dati
- 6 Gestione adattiva dei carichi di lavoro
e ottimizzazione delle risorse
- 9 Ottimizzazione delle query con il machine
learning
- 10 Query basate sulla fiducia
- 11 [Progettato per l'AI](#)
- 12 Esplorare i dati più velocemente con query
in linguaggio naturale
- 14 Usare i linguaggi più diffusi ed esempi di
codici per costruire applicazioni di AI
- 17 Modellare relazioni complesse
con grafici e SQL
- 18 Analizzare i dati della blockchain in modo
nativo
- 19 [Collocazione di Db2 in un'azienda
orientata all'AI](#)

Il database basato sull'AI:

Alimentato dall'AI e progettato per l'AI

L'aumento dei volumi e della varietà dei dati richiedono competenze, processi e infrastrutture innovativi per supportare la trasformazione digitale in tutti i settori. Il machine learning (ML) e l'intelligenza artificiale (AI) sono componenti fondamentali di questo processo, poiché queste tecnologie permettono ai computer di elaborare i dati e di apprendere dall'esperienza senza la necessità di effettuare continui interventi umani.

AI e ML supportano operazioni informatiche tradizionali e non, per esempio analisi e logistica, interazione con il linguaggio naturale, composizione musicale, progettazione di strutture chimiche e individuazione di frodi. Forniscono informazioni basate sui dati, che consentono alle aziende di capire le dinamiche di eventi passati, presenti e futuri.

Il viaggio verso l'AI è come una scala in cui ogni componente è un piolo. Chi intraprende questo viaggio rafforza la scala e ne accresce l'utilità.

Integra

Rendi operativa l'AI in tutto il tuo business

Analizza

Crea e scala l'AI con fiducia e trasparenza

Raccogli

Rendi tutti i tuoi dati semplici e accessibili

Modernizza

Prepara i tuoi dati per il cloud ibrido e per l'AI

Organizza

Crea una base di analytics pronta al business

Per intraprendere il viaggio verso l'AI è necessaria una solida infrastruttura di dati, che consenta alle aziende di accedere, raccogliere, organizzare e rendere tutti i dati disponibili per l'analisi, indipendentemente dal tipo, dalla fonte o dal deployment. Supportando il data management con il ML, l'AI diventa sia l'obiettivo che il mezzo per raggiungerlo.

Le soluzioni integrate potenziate con AI e ML rendono la gestione dei dati più semplice, più veloce e più intelligente, permettendo a sviluppatori e data scientist di utilizzare i dati in nuovi modi, costruendo funzionalità di AI che portano benefici all'intero sistema aziendale, dalle operazioni alla costruzione di modelli allo sviluppo di applicazioni.

Alimentato dall'AI

Le funzionalità di machine learning e di AI integrate in IBM® Db2® supportano diverse funzioni di data management innovative, tra cui:

- Virtualizzazione dei dati full-stack
- Gestione automatica dei carichi di lavoro e ottimizzazione delle risorse
- Performance delle query 10 volte migliori
- Risultati delle query basati sulla fiducia

Virtualizzazione dei dati



Di cosa si tratta

Una combinazione di federazione dei dati con l'aggiunta di un livello di astrazione. Fornisce a tutti gli utenti e alle applicazioni la possibilità di interagire con più fonti di dati contemporaneamente da un unico punto di accesso, indipendentemente da tipo, formato, dimensione o collocazione dei dati sottostanti. Questa maggiore capacità di accesso aiuta a scomporre i silos di dati e consente di utilizzare i dati in modo più rapido ed efficiente, per accelerare la propria scalata verso l'AI.

Perché è importante

Un unico punto di accesso ai dati offre i seguenti vantaggi:

– Semplice accesso ai dati per i vostri professionisti

Un'unica panoramica di tutti i dati consente la ricerca e l'accesso ai dati attraverso diverse fonti: strutturate o non strutturate, relazionali o NoSQL, senza dover spendere tempo e risorse nel trasferimento dei dati. Gli sviluppatori e gli ingegneri di dati ottengono l'accesso ad ogni possibile tipo di dati, sia storici che nuovi, e possono trarre insight da combinazioni inaspettate di tali dati.

– Un unico punto di controllo per la governance e la sicurezza

Le disposizioni di sicurezza e di governance sono molto più semplici e solide e meno inclini agli errori, dal momento che gli amministratori possono indirizzare tutti gli utenti e le applicazioni verso un unico punto di accesso per tutti i repository di dati.

– Riduzione del trasferimento di dati

L'elaborazione dei dati può essere effettuata nei repository stessi, invece di trasferire prima i dati soggetti a query in un repository locale. Questo può portare a una significativa riduzione della latenza e dei costi della larghezza di banda.

Come funziona

Gli ambienti di dati delle aziende moderne dipendono da più datastore. Questi includono dati transazionali in un datastore relazionale, sentiment data in un cluster Hadoop, dati storici dei clienti in un data warehouse e dati clickstream di e-commerce in un cluster di fast data.

Senza virtualizzazione, chiunque voglia interagire con questi dati, come amministratori di database (DBA), sviluppatori di applicazioni e data scientist, dovrebbe implementare un metodo di accesso personalizzato per ogni data store. Nella migliore delle ipotesi, questo comporterebbe non poche difficoltà per tutte le persone coinvolte. Limiterebbe gli insight che gli utenti potrebbero trarre confrontando fra loro diversi tipi di dati (ad esempio, la storia dei clienti con il comportamento di navigazione in tempo reale). E nella peggiore delle ipotesi si rischierebbe di corrompere i dati e di provocare nuove violazioni della sicurezza.

Le funzionalità di federazione dei dati sono quindi diventate piuttosto popolari per fornire un unico punto di accesso a dati diversi in tutta l'azienda. Tuttavia, la virtualizzazione dei dati fornita da IBM va oltre la federazione, fornendo un livello di astrazione e permettendo agli utenti di definire la propria terminologia e metodologia per l'accesso ai dati (vedi Figura 1).

In azione

Virtualizzazione dei dati

La banca europea ING sta collaborando con IBM per implementare un unico punto di accesso ai dati per i propri utenti in tutto il mondo. Questo permette alla banca di gestire un'infrastruttura di dati globale, pur offrendo agli utenti le prestazioni, la scalabilità, la sicurezza e le funzionalità di governance di cui hanno bisogno. Questo contribuisce inoltre a creare una maggiore connessione tra le diverse aree aziendali.

Con la soluzione IBM, possono essere aggiunti nuovi dati alla piattaforma ING da qualsiasi parte del mondo e sono accessibili a livello globale a qualsiasi utente della banca, senza la necessità di modificare i singoli schemi di accesso ai dati o i permessi di sicurezza.



Figura 1. Rappresentazione di come la virtualizzazione consente un accesso e una governance unificati per diverse fonti di dati

Gestione adattiva dei flussi di lavoro e ottimizzazione delle risorse

Di cosa si tratta

Tecnologia intelligente che alloca automaticamente le risorse di dati per gestire diversi carichi di lavoro. La tecnologia è sia automatizzata che adattiva, in quanto risponde basandosi su condizioni passate e prevede anche le richieste future.

Perché è importante

Con la gestione adattiva del carico di lavoro, si ottiene un sistema pronto all'uso stabile e ad alte prestazioni, senza la necessità di spendere tempo ed energie per la messa a punto e la configurazione. Questo per le aziende moderne è fondamentale, a causa della complessità delle loro piattaforme dati e dell'alto tasso di errore delle strategie di gestione manuale del carico di lavoro.

Quando arrivano troppi carichi di lavoro in una sola volta, un database deve determinare il modo in cui gestire le richieste extra. I cosiddetti schemi "aperti" o manuali non hanno alcun

meccanismo di feedback; essi richiedono semplicemente all'utente di impostare limiti predeterminati per il numero o la dimensione dei carichi di lavoro. Ma questi limiti possono essere infranti quando ci si trova di fronte ad un insieme complesso di carichi di lavoro, e anche nelle migliori circostanze richiedono un costante monitoraggio e regolazione manuale. Ciò comporta problemi di performance del database, sottoutilizzo o anche errori.

Al contrario, Db2 rileva e prevede le tendenze di utilizzo, in modo da poter notificare agli utenti il problema o addirittura correggerlo automaticamente prima che la situazione si aggravi. Questo semplifica le cose per gli amministratori di database (DBA) e abbassa il costo di proprietà di ogni database, riducendo il tempo necessario per la configurazione e la messa a punto. I test IBM hanno dimostrato che questo miglioramento nell'utilizzo ha portato a miglioramenti delle prestazioni complessive del database fino al 30%.





Come funziona

La gestione adattiva del carico di lavoro integra il machine learning sotto forma di meccanismo di feedback; in questo modo il database monitora costantemente il tempo di esecuzione previsto ed effettivo dei vari carichi di lavoro, quindi organizza le risorse disponibili per far fronte ad eventuali carenze. I database Db2 sono preconfigurati e possono essere eseguiti con una configurazione necessaria minima per la maggior parte dei carichi di lavoro.

Se per situazioni particolarmente complesse è necessario configurare manualmente, questa tecnologia adattiva rende più facile l'operazione. Gli amministratori di database (DBA) possono creare più classi di carico di lavoro e indicare degli obiettivi di performance per ciascuna di esse. L'addetto alla gestione adattiva controlla quindi i carichi di lavoro in arrivo e assegna le risorse per raggiungere tali obiettivi.

In azione

Gestione automatizzata dei carichi di lavoro

La gestione automatizzata e adattiva dei carichi di lavoro affronta nello specifico molti dei punti deboli tipici della gestione di un database aziendale.

Ad esempio, nelle impostazioni del data warehouse il database può ricevere un'ampia gamma di incarichi diversi, come l'acquisizione di dati in streaming, l'elaborazione di report a bassa latenza con tempi di risposta inferiori al secondo (come il caricamento di pagine web) o l'elaborazione di report batch più pesanti che necessitano di più minuti per essere completati. L'amministratore deve assicurarsi che il sistema non si blocchi su lavori più grandi impedendo il completamento di lavori urgenti.

La gestione adattiva dei carichi di lavoro rende tutto questo molto semplice, facendo in modo che gli amministratori possano creare raggruppamenti di risorse (chiamati anche classi di carico di lavoro o di servizio) e assegnare una parte del sistema ad ogni raggruppamento. Il database assegna quindi le risorse in modo intelligente per raggiungere gli obiettivi di performance. Tutto questo avviene senza la necessità di monitorare e regolare individualmente e manualmente i carichi di lavoro e le risorse.

L'AI è sia
l'obiettivo che
il mezzo per
raggiungerlo.



Ottimizzazione delle query con il machine learning



Di cosa si tratta

L'ottimizzatore del machine learning Db2 fornisce un ulteriore salto di qualità nell'ottimizzazione intelligente che si serve di machine learning non supervisionato per fornire strategie di esecuzione delle query che migliorano la tradizionale ottimizzazione delle query basata sui costi.

Perché è importante

I tradizionali ottimizzatori dei costi del carico di lavoro possono suggerire strategie di esecuzione per una determinata query, ma non sono sensibili alle modifiche recenti del database e non possono imparare dall'esperienza. Possono suggerire quale dovrebbe essere la strategia di esecuzione più veloce, tuttavia continueranno a raccomandare la stessa strategia anche se non funziona come previsto.

L'ottimizzatore del machine learning Db2 al contrario recepisce il feedback delle prestazioni effettive delle query in modo da consigliare strategie di esecuzione che forniscano all'infrastruttura di dati i risultati effettivamente migliori, perfezionando il percorso delle query ad ogni esecuzione.

Con un ottimizzatore di machine learning, i gli amministratori di database (DBA) non devono più perdere tempo a monitorare le prestazioni del sistema e a cercare di ottimizzare le query. Possono invece concentrarsi su attività di maggior valore per l'azienda, come implementare applicazioni di AI, sviluppare strategie di utilizzo dei dati e assistere gli utenti di tutta l'azienda aiutandoli a utilizzare al meglio i dati disponibili.

Come funziona

L'ottimizzatore di machine learning Db2 imita i modelli delle reti neurali per ottimizzare i percorsi delle query, con il risultato che alcune query vengono completate 8-10 volte più velocemente (test interni IBM).

In azione

Ottimizzazione delle query con il machine learning

Uno strumento di ottimizzazione dei costi tradizionale si serve della modellazione statistica e delle risorse per valutare le strategie di esecuzione di una determinata query. È in grado di fornire la prima opzione mostrata nella figura 2 qui sotto: l'accostamento di due coppie di tabelle e il successivo accostamento dei risultati per ottenere una risposta. Tuttavia, un algoritmo di machine learning può imparare dall'esperienza e rilevare la disponibilità di eventuali strategie di esecuzione migliori: accostare una coppia di tabelle, accostare una terza tabella al risultato e infine accostarne una quarta al nuovo risultato ottenuto per arrivare alla stessa risposta.

Senza machine learning



Con machine learning



Figura 2. Esempio di strategie di esecuzione delle query con ottimizzazione dei costi (in alto) e con ottimizzazione del machine learning (in basso)

Queste modifiche nella strategia di esecuzione possono essere applicate in tempo reale, in risposta alle effettive esecuzioni delle query. Per applicazioni sensibili al fattore tempo, come il rilevamento delle frodi, un affinamento dell'elaborazione in tempo reale può essere fondamentale.



Query basate sulla fiducia

Di cosa si tratta

Una funzionalità avanzata che fornisce risultati di query SQL in termini di probabilità, o “migliori corrispondenze”, piuttosto che una semplice risposta sì/no.

Perché è importante

Questo tipo di matching basato sulla fiducia era già disponibile nelle impostazioni di machine learning, ma Db2 estende questa caratteristica anche alle formulazioni in linguaggio SQL. Per un ingegnere SQL, questa caratteristica amplia in modo consistente la gamma di possibili operazioni sui dati che possono essere eseguite senza coinvolgere un data scientist; questo rende gli ingegneri SQL più preziosi per l'azienda e sgrava i data scientist, già sovraccarichi di lavoro.

Come funziona

La query basata sulla fiducia aggiunge estensioni di machine learning a SQL implementando reti neurali feed-forward profonde. Questa varietà di deep learning non supervisionato trova aree in cui sia la “somiglianza” che la probabilità di trovare una corrispondenza sono elevate. L'esecuzione di query basate sulla fiducia può essere applicata a:

- Query di somiglianza/diversità
- Query di ragionamento induttivo (per esempio clustering semantico, analogie, intrusi)
- Operazioni semantiche con group-by
- Anomalie dello schema (ad esempio, rilevamento di frodi)
- Immagini, audio, video

In azione

Query basate sulla fiducia

Un database della polizia è un ambiente ricco di variabili: altezza, peso, età, colore della pelle, colore degli occhi, segni distintivi e molte altre ancora; tutte queste variabili sono fondamentalmente incerte a causa dell'inaffidabilità delle testimonianze. Una ricerca tradizionale in questo tipo di database utilizza una query SQL molto complessa per tenere conto di questa incertezza, creando manualmente un “intervallo” di valori che corrisponda a una data dichiarazione del testimone. Ad esempio, un peso riportato di 195 libbre potrebbe far corrispondere i sospetti con un peso effettivo di 185-205 libbre. Questi intervalli non sono agevoli da usare e possono eliminare erroneamente determinati individui sospettati se un valore fornito da un testimone è troppo al di fuori dell'intervallo. Al contrario, un agente che cerca di far corrispondere le dichiarazioni dei testimoni con il profilo di un individuo sospettato può usare una query basata sulla fiducia per generare un semplice statement SQL probabilistico che trovi il soggetto che meglio corrisponde al profilo complessivo del testimone, anche quando i singoli data point non rientrano nell'intervallo originale di corrispondenze accettabili.

Progettato per l'AI

Db2 è anche progettato per l'AI, il che significa che i professionisti dei dati possono abilitare più facilmente potenti aggiornamenti del carico di lavoro basati sull'AI e miglioramenti del flusso di lavoro. Tra cui ad esempio:

- Esplorare i dati più velocemente con query in linguaggio naturale
- Usare i linguaggi più diffusi ed esempi di codici per costruire applicazioni di AI
- Modellare relazioni complesse con grafici e SQL
- Analizzare i dati della blockchain in modo nativo

Esplorare i dati più velocemente con query in linguaggio naturale

Di cosa si tratta

Augmented Data Explorer è un'interfaccia web simile a Google per l'esplorazione di dati aziendali complessi. Offre la possibilità di ricerca in linguaggio naturale per gli utenti e API REST per gli sviluppatori. Una domanda posta in linguaggio naturale genererà visualizzazioni di dati e riepiloghi che attingono da tutte le fonti di dati disponibili.

Perché è importante

Man mano che i dati nel loro insieme diventano sempre più complessi, molti utenti si trovano di fronte a set di dati sempre più vasti. Le competenze per sviluppare query SQL o applicazioni di analisi non sono molto diffuse, ma sono sempre più numerose le persone che vogliono risposte dai dati in tempi rapidi.

Augmented Data Explorer è un nuovo strumento di esplorazione dei dati. Dà accesso a sofisticate query di dati anche a utenti meno esperti, che magari hanno semplicemente bisogno di inserire una query in linguaggio naturale come “vendite per il 2019” o “quanto spesso piove in ottobre?” Questo

genererà uno statement SQL al fine di interrogare tutti i dati rilevanti e restituire il risultato desiderato in linguaggio naturale, tramite una visualizzazione dinamica dei dati o entrambe le cose.

In passato, gli utenti aziendali in genere avevano bisogno di una combinazione di vari processi manuali per tradurre le richieste formulate in linguaggio naturale nella query SQL desiderata e della consulenza di un esperto per aiutare a risolvere gli inevitabili imprevisti che si presentavano. Augmented Data Explorer è stato progettato per evitare queste due necessità e offrire nuove strade per ottenere ulteriore valore dai dati già in possesso.

I data scientist e gli utenti più esperti possono utilizzare Augmented Data Explorer per eseguire query iniziali su set di dati sconosciuti o di grandi dimensioni, per farsi un'idea del contenuto di tali set di dati e capire dove iniziare le indagini. Augmented Data Explorer offre anche API REST, in modo che gli sviluppatori possano facilmente integrare nelle proprie applicazioni queste funzionalità di ricerca e di previsione in linguaggio naturale, invece di perdere tempo per codificare una propria soluzione.



Come funziona

Augmented Data Explorer utilizza la tecnologia del ML per analizzare e indicizzare i dati per portare alla luce insight nascosti. La sua funzionalità di AI è focalizzata sul supporto del linguaggio naturale (comprensione del contesto, dei sinonimi e della sintassi per un'ampia varietà di discipline e campi), oltre che sulla previsione di argomenti di interesse correlati sulla base dei termini di ricerca. Lo strumento è containerizzato per semplificare deployment e gestione, e dispone di una semplice interfaccia web.

Include inoltre funzionalità grafiche, rivelando a colpo d'occhio a quali dati il motore di esplorazione ha effettivamente avuto accesso e quali ha analizzato. L'utente può facilmente vedere quali dati (come schemi e tabelle) vengono utilizzati per produrre le risposte fornite dal sistema.



Valore di Augmented Data Explorer

Augmented Data Explorer è molto utile in situazioni in cui gli utenti dei dati non sono sicuri delle giuste domande da porre. Facilita l'elaborazione di domande per gli utenti e li aiuta a trovare relazioni all'interno dei dati e a ottenere utili insight. Inoltre, amplia il pool di persone in grado di affrontare queste sfide di business. In contesti aziendali in costante cambiamento, in cui gli utenti hanno molte responsabilità, Augmented Data Explorer rappresenta un primo passo per affrontare l'enorme quantità di dati ora disponibili e trarre valore aggiunto da questi ultimi.

Prendiamo per esempio un responsabile delle vendite presso un'azienda di retail. Questo individuo può porsi molte domande diverse giorno per giorno o mese per mese, per esempio sulla suddivisione delle vendite per regione, sulla comprensione dei tassi di fidelizzazione dei clienti, sull'individuazione delle cause dell'abbandono del carrello, sulla previsione dei motivi per cui i clienti abituali a volte scompaiono, sulla spiegazione di come il tempo atmosferico influisce sui tempi di consegna dell'inventario e molto altro ancora.

Augmented Data Explorer è progettato per identificare e rispondere rapidamente a nuove domande, estendendo così le funzionalità dell'AI a un segmento molto più ampio di operazioni commerciali.

Usare i linguaggi più diffusi ed esempi di codici per costruire applicazioni di AI

Di cosa si tratta

Db2 11.5 include supporto nativo per i linguaggi e le librerie più diffusi, inclusi Python, JSON, GO, Ruby, PHP, Java, Node.js, Sequelize e Jupyter Notebooks. I professionisti dei dati possono costruire applicazioni di AI in un linguaggio che già conoscono, per poi collegarle senza soluzione di continuità ai dati in Db2.

Perché è importante

Con il supporto nativo per i più diffusi linguaggi di machine learning, IBM crea un collegamento fra il data management e il rapido sviluppo di applicazioni con un ecosistema di dati focalizzato su ML e AI. Gli sviluppatori e i data scientist possono conservare i dati in un database aziendale e continuare a utilizzare i linguaggi di sviluppo e le librerie esistenti e diffusi.

Il supporto di linguaggi e librerie nativi offre molteplici vantaggi che portano a uno sviluppo più rapido e semplice delle applicazioni AI:

- Gli sviluppatori possono sfruttare i linguaggi e le librerie più diffusi, oltre ad esempi di codice specifici del settore su GitHub, per costruire rapidamente applicazioni a proprio uso.
- Gli sviluppatori possono facilmente scrivere applicazioni che utilizzano i dati di Db2 in modo nativo, eliminando la necessità di un codice back-end personalizzato per accedere a questi dati ed escludendo un fattore che comunemente è causa di errori o che comunque comporta necessità di gestione per le applicazioni.
- I nuovi database se necessario possono essere rapidamente implementati e scalati sul cloud e poi collegati senza soluzione di continuità alle app.
- Gli sviluppatori hanno accesso agli ambienti di sviluppo per il machine learning (IBM Watson® Studio) all'interno dello stesso ecosistema di database.
- I data scientist possono dedicare il loro tempo a lavorare su modelli di machine learning piuttosto che a risolvere problemi di accesso ai database.
- Le aziende possono assumere persone dotate delle competenze necessarie per analizzare i dati, senza bisogno di formarle su nuove competenze o linguaggi necessari solo per accedere ai dati.

Come funziona

Il processo di sviluppo delle applicazioni di AI comporta diverse domande importanti:

- Dove vengono memorizzati i dati rilevanti?
- Come posso accedervi ed esplorarli?
- Come si applicano le funzionalità di machine learning nella costruzione di applicazioni?
- Quali linguaggi posso usare?

All'interno dell'ecosistema IBM, ci sono risposte di classe enterprise a tutte queste domande. I dati possono essere memorizzati in Db2 o Db2 Warehouse, Event Store o Hadoop, su cloud o on-premise, praticamente in qualsiasi combinazione di database. Ciò che hanno in comune è il fatto di essere di classe enterprise, con il supporto completo, l'affidabilità e le performance che ne derivano. La code base comune Db2 e il motore comune SQL sottostante consentono alle applicazioni di elaborare query rivolte ai dati di qualsiasi componente dell'ecosistema che non necessitano di essere riscritte per essere eseguite su altre parti dell'ecosistema. Inoltre, grazie a una serie di nuovi driver disponibili per i principali linguaggi e framework di programmazione open-source, ora per gli sviluppatori è più facile analizzare e costruire modelli di machine learning nelle applicazioni che utilizzano Db2.

I professionisti dei dati possono utilizzare alcune delle altre funzionalità descritte in questo ebook, come Augmented Data Explorer, per ottenere una rapida comprensione del contenuto dei dati. Una volta memorizzati in Db2, i dati possono essere collegati direttamente all'ambiente di sviluppo di machine learning di Watson Studio. Librerie specifiche del settore, esempi di codice e modelli forniscono agli sviluppatori un vantaggio nella costruzione di un'applicazione funzionale. E infine, un set di driver nativi converte gli statement SQL in un modello specifico per il linguaggio di ML (come Python), per un ulteriore sviluppo nel linguaggio scelto.

Il valore del supporto nativo per il linguaggio

Il mercato potenziale per le applicazioni abilitate all'AI è enorme. Ecco alcuni esempi:

Retail: Motori di raccomandazione, contenuti mirati, promozioni

Assicurazioni: Analisi di frodi, elaborazione di applicazioni

Logistica: Fleet management, manutenzione predittiva

Trasporti: Assistenza in tempo reale su itinerario e percorso in base al traffico, alle condizioni meteorologiche e altro ancora

Una moltitudine di sviluppatori e data scientist lavora nel settore per sviluppare applicazioni e algoritmi intelligenti, di solito utilizzando Python, Java o un altro linguaggio di largo utilizzo. I dati che alimentano queste applicazioni risiedono in uno o più database. Esistono database che si collegano in modo nativo a Python, ma in generale questi non sono attrezzati per fornire performance e affidabilità di classe enterprise. Tuttavia, senza l'integrazione nativa tra i linguaggi di sviluppo di ML scelti e il database aziendale sottostante, è difficile e macchinoso per gli sviluppatori collegare l'applicazione stessa ai dati necessari.

Grazie all'integrazione nativa tra Db2 e i più diffusi linguaggi di sviluppo, lo sviluppatore può accedere ai dati senza dover coinvolgere un amministratore di database (DBA) o uno specialista di SQL. Il risultato è una migliore produttività per gli sviluppatori e i data scientist e applicazioni più solide e di rapido sviluppo.

L'AI fornisce
agli utenti di
tutti i livelli gli
strumenti per
sfruttare i dati
come mai
prima.



Modellare relazioni complesse con grafici e SQL

Di cosa si tratta

Db2 integra profondamente le funzionalità dei grafici con i dati relazionali e SQL; in questo modo le applicazioni grafiche possono essere eseguite direttamente a partire dai dati relazionali, e il motore SQL di Db2 può interrogare direttamente i dati dei grafici.

Perché è importante

Il grafico è uno strumento incredibile che fornisce importanti insight, ma che prima era incompatibile con i sistemi transazionali e OLTP. Per aggirare questa incompatibilità, le aziende conservano solitamente la loro normale infrastruttura relazionale, da cui estraggono alcuni dati inserendoli in un database di grafici, eseguendo poi le applicazioni grafiche in questo contesto. Questo tipo di duplicazione comporta un ritardo e un sovraccarico di lavoro non più accettabili. Le aziende hanno bisogno di ottenere in pochi secondi insight basati sui grafici nei loro dati relazionali.

Gli utenti di diversi settori vogliono aggiungere gli insight grafici al loro toolkit esistente di insight analitici, ma senza compromettere la loro infrastruttura di dati o aggiungere un nuovo database solo per supportare il grafico. Db2 rende tutto questo possibile.

Come funziona

I dati dei grafici sono memorizzati in tabelle all'interno del framework relazionale di Db2 o Db2 Event Store. Ciò significa che il motore SQL può interrogare direttamente i dati dei grafici e le applicazioni grafiche possono interrogare i dati direttamente nelle tabelle relazionali. Questa architettura supporta anche linguaggi di query grafiche open-source, come Gremlin e Tinkerpop.

La profonda integrazione tra i dati relazionali e le funzionalità dei grafici in Db2 offre diversi importanti vantaggi:

- Il grafico può essere applicato direttamente ai dati relazionali esistenti, aggiungendo un nuovo livello di insight oltre all'analisi basata su SQL.
- L'analisi SQL può essere applicata direttamente ai dati dei grafici memorizzati nelle tabelle relazionali.
- Le applicazioni grafiche possono servirsi di un'API personalizzata o di Spark per connettersi a Db2.
- Le transazioni con proprietà ACID aggiornano il grafico in tempo reale senza disturbare le applicazioni relazionali esistenti: i grafici possono essere così utilizzati per l'elaborazione delle transazioni in tempo reale in aggiunta all'analisi.
- Sono disponibili soluzioni specifiche per i settori che necessitano maggiormente di grafici, come quello sanitario e quello finanziario.



Il valore del supporto nativo dei grafici

Ci sono molti casi in cui gli insight basati su grafici potrebbero consentire nuovi vantaggi e flussi di lavoro con i dati relazionali. Ad esempio, un rivenditore potrebbe utilizzare relazioni grafiche in tempo reale per notare uno schema dei profitti per un particolare articolo in diversi luoghi. Questo insight può quindi aggiornare istantaneamente il database transazionale del rivenditore e far sì che l'articolo venga rimosso dagli scaffali il giorno stesso.

In un campo diverso, un investigatore di frodi assicurative potrebbe utilizzare il grafico per raggruppare diversi set di dati: numero e dimensione dei sinistri, relazioni tra i richiedenti, identità dei fornitori di servizi e altro ancora. Le relazioni che emergono nel grafico possono rivelare la probabilità di frode.

Analizzare i dati della blockchain in modo nativo

Di cosa si tratta

L'API nativa di connessione della blockchain in Db2 fornisce chiarezza e insight senza precedenti sui dati altamente compressi contenuti nei registri della blockchain.

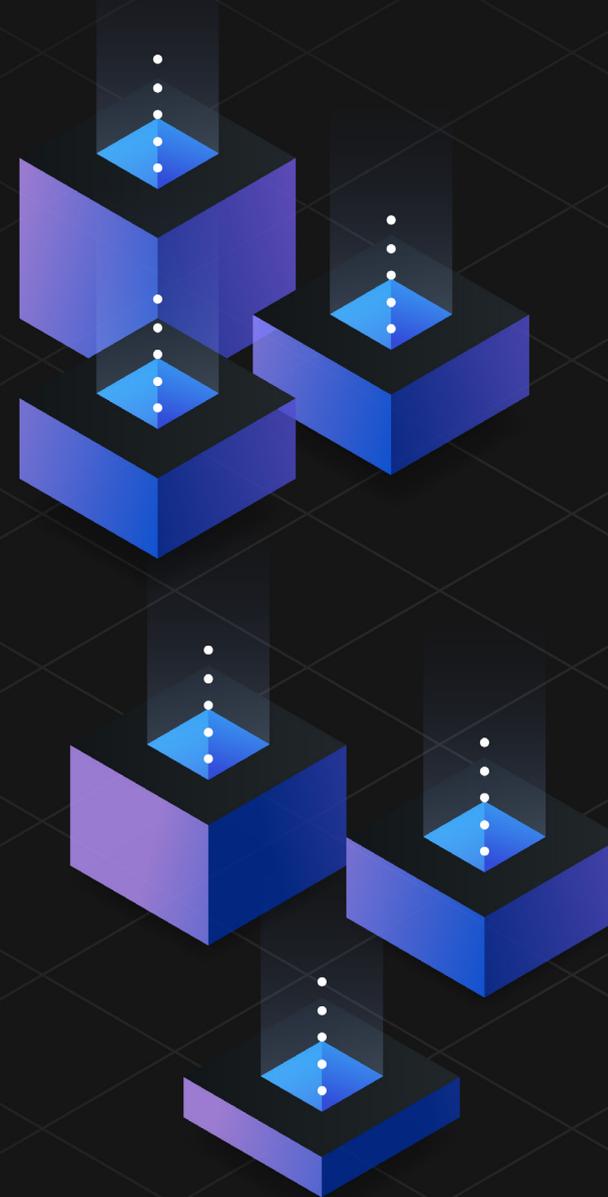
Perché è importante

La tecnologia blockchain sta venendo rapidamente adottata in diversi settori, ma questo sta causando rallentamenti e intoppi nei processi di analisi. Enormi quantità di informazioni potenzialmente preziose sono contenute in forma compressa in ciascun registro della blockchain, ma fino ad oggi non esisteva un modo semplice per visualizzare e analizzare tali informazioni. Il connettore per la blockchain di Db2 cambia la situazione. I registri della blockchain, compatti, chiari, controllabili e inalterabili, sono ora disponibili per l'analisi come qualsiasi altra fonte di dati all'interno dell'azienda. Le aziende ottengono così tutti i vantaggi del motore Db2, in fatto di performance, flessibilità, scalabilità e sicurezza, senza la necessità di sviluppare soluzioni di reporting ad hoc per la sola blockchain.

Il connettore per la blockchain è reso ancora più prezioso dalla connettività nativa e dall'interoperabilità con altri tipi di dati disponibili in Db2. Le aziende possono facilmente integrare dati aggiuntivi provenienti da altri archivi di dati per ampliare il contesto e aprire nuove possibilità analitiche per i dati della blockchain. E infine, questo connettore apre la possibilità di utilizzare i dati della blockchain per applicazioni di AI, un caso d'uso che in precedenza sarebbe stato molto difficile da gestire perché i dati della blockchain erano archiviati in un proprio silo. Gli sviluppatori di AI possono ora integrare facilmente i set di dati della blockchain, sia come fonte di dati primaria per le loro applicazioni, sia per fornire dettagli aggiuntivi secondari.

Come funziona

Il connettore per la blockchain estrae i dati transazionali, che vengono compressi e memorizzati all'interno del registro della blockchain, e li presenta nella forma di una tabella relazionale in Db2. Una solida strategia di caching utilizza le funzionalità esistenti di Db2 per creare una tabella di cache per i dati della blockchain al fine di migliorare le performance delle query, consentendo al contempo agli utenti di recuperare i dati più recenti quando necessario.



Il valore del supporto nativo della blockchain

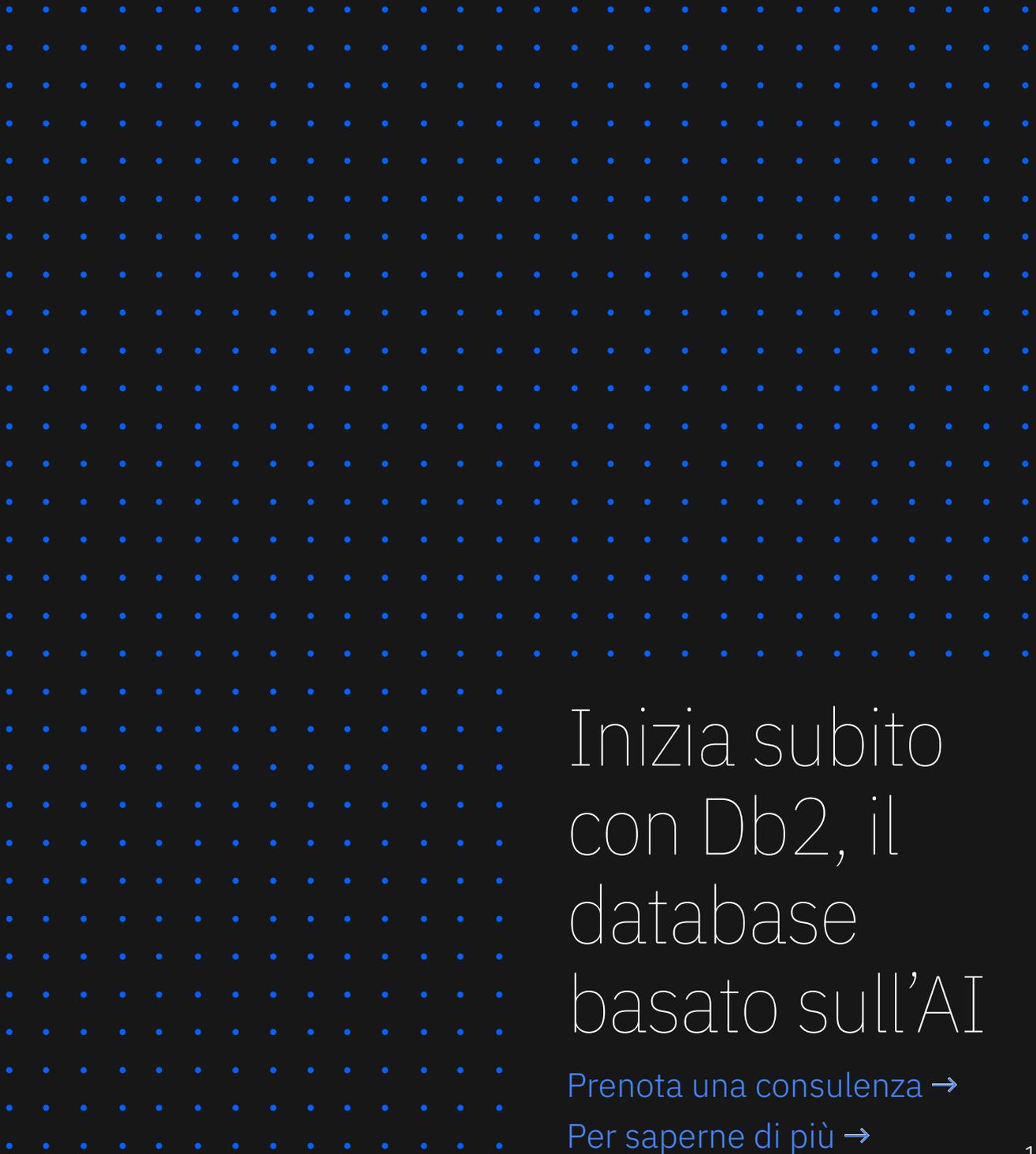
La blockchain è già utilizzata in settori come quello assicurativo, finanziario, logistico e sanitario per memorizzare un registro immutabile di transazioni, ma questi settori hanno anche la necessità di eseguire analisi per comprendere meglio i dati. Ad esempio, le compagnie assicurative vogliono sapere quanti sinistri sono stati aperti per cliente in una particolare area geografica; le compagnie di spedizione vogliono tenere traccia di container e veicoli; le aziende sanitarie vogliono comprendere le cartelle cliniche dei pazienti.

Invece di costruire un back-end personalizzato per estrarre questi dati, possono semplicemente collegare Db2 al data store della blockchain ed eseguire le loro analisi come farebbero con qualsiasi altro insieme di dati. Grazie alla connettività e alla flessibilità proprie di Db2, possono correlare altri dati contestuali con i dati transazionali della blockchain (ad esempio le condizioni atmosferiche durante una particolare spedizione o gli individui che potrebbero aver presentato richieste di risarcimento simili a quelle di un particolare richiedente) e ottenere una comprensione molto più ricca e completa di ciò che ogni transazione comporta per l'attività. Le applicazioni di AI che lavorano direttamente sui dati della blockchain sono ora una possibilità reale.

Collocazione di Db2 in un'azienda orientata all'AI

È possibile implementare le funzionalità di Db2 come deployment tradizionale on-premise, come database relazionale cloud o istanza del warehouse, sulla piattaforma Hybrid Data Management ottenibile su abbonamento, su IBM Cloud Pak™ for Data o utilizzando altri modelli di deployment; la tecnologia sottostante è eccezionale e i vantaggi per il business sono notevoli.

Db2 è progettato in modo da contribuire a rendere l'AI accessibile per gli utenti delle aziende di tutte le dimensioni. La proposta di valore è chiara. Costruire, eseguire e gestire l'AI aziendale su qualsiasi infrastruttura: on-premise o su cloud. È importante sfruttare tutte le fonti di dati, indipendentemente da dove si trovano fisicamente. È importante sfruttare le funzionalità dell'AI integrate nel database per ottenere insight predittivi e proattivi per le decisioni basate sui dati. Con Db2 come base di dati le aziende possono costruire e connettere in modo più efficiente le applicazioni di AI, ottenendo rapidamente insight fruibili e rendendo l'AI accessibile a un numero sempre maggiore di persone.



Inizia subito
con Db2, il
database
basato sull'AI

[Prenota una consulenza →](#)
[Per saperne di più →](#)



© Copyright IBM Corporation 2020.

IBM Italia S.p.A.
Circonvallazione Idroscalo
20090 Segrate (MI)
Italia

Prodotto negli Stati Uniti d'America, gennaio 2020

IBM, il logo IBM, **ibm.com**, Db2, IBM Cloud Pak e IBM Watson sono marchi di International Business Machines Corp., registrati in varie giurisdizioni in tutto il mondo. I nomi di altri prodotti e servizi possono essere marchi di IBM o di altre aziende. L'elenco aggiornato dei marchi IBM è disponibile nella sezione "Copyright and trademark information" dell'indirizzo web [ibm.com/legal/copytrade.shtml](https://www.ibm.com/legal/copytrade.shtml).