Guida all’analisi predittiva per le aziende

IBM

Contenuti

Perché questa guida? 03

Qual è lo scenario 04

Perché l'analytics predittiva e l'AI sono così importanti 05

Cosa cambia con l'adozione dell'AI 06

Come può l'AI valorizzare le attività dell'azienda? 08

I gradini della scala dell'AI 13

Quali sono le soluzioni disponibili? 14

Il passo successivo 16

I punti chiave 17

Perché combinare l'ottimizzazione decisionale con l'analytics predittiva? 18

Glossario 19

Perché questa guida?

La lungimiranza, nel business, è tutto. Se si riescono a prevedere i prossimi eventi, allora si può:

- prendere decisioni più oculate
- entrare nel mercato più rapidamente
- scavalcare i concorrenti

La moderna analytics predittiva consente all'azienda di valorizzare i dati storici con insight in tempo reale e di sfruttare tutto ciò per prevedere e plasmare il futuro.

L'analytics predittiva è una pietra miliare nel percorso analitico; un punto di confluenza, in cui le tecniche tradizionali di analisi statistica convergono con il nuovo mondo dell'intelligenza artificiale (AI).

Secondo una ricerca Forrester, le imprese hanno raggiunto un punto in cui devono iniziare a combinare l'apprendimento automatico con l'ingegneria della conoscenza. La valorizzazione dei dati mediante le competenze umane accelererà drasticamente lo sviluppo delle applicazioni AI.

Questa guida aiuterà la tua azienda a compiere le seguenti azioni:

- esplorazione del moderno scenario di analytics predittiva
- individuazione di opportunità di crescita e miglior uso dell'AI
- attribuzione di strumenti a gruppi data science e stakeholder per fornire rapidamente valore

L' analytics predittiva utilizza previsioni generate automaticamente che, con l'aiuto dell'intelligenza umana, possono favorire il business aziendale.

Qual è lo scenario

Mentre è in corso la rivoluzione dell'AI, le aziende chiedono ai propri gruppi data science di affrontare grandi questioni.

Per questo, dagli analisti scientifici dei dati ci si attende oggi molto di più del semplice lavoro su progetti di ricerca una tantum. Si chiede loro di trovare modalità ripetibili e automatizzate per ottenere insight in tempo reale a favore dei processi decisionali quotidiani.

Per soddisfare tali aspettative, i responsabili di data science hanno bisogno non solo di poter spiegare agli stakeholder le potenzialità della moderna analytics predittiva, ma anche di fornirne i risultati.

La capacità di definire ed eseguire una strategia di data science di successo sarà, nei prossimi anni, uno dei fattori differenzianti chiave tra leader e non leader.

Ma non è un compito facile. Costruire le proprie competenze di data science prevede le seguenti attività:

- attrarre e trattenere un team diversificato di specialisti capaci
- fornire loro strumenti per collaborare senza soluzione di continuità
- attuare solide strutture di governance per far sì che i modelli predittivi siano sempre attendibili per l'azienda

E soprattutto, i team di data science e quelli aziendali devono trovare nuovi modi per collaborare efficacemente. Tali metodi includono la consapevolezza di cosa l'analytics predittiva sia in grado di fare, identificando gli ambiti nei quali l'AI può apportare vantaggi all'azienda.

1. Come si comportano i nostri clienti?
2. Perché i nostri mercati sono fluttuanti?
3. Cosa fa sì che le nostre strategie aziendali abbiano successo o falliscano?
4. Cosa succederà poi?
5. Come si sostengono finanziariamente le previsioni?
6. Dove sono i centri di acquisto?

Perché l’analytics predittiva e l’AI sono così importanti

L’analytics predittiva non è un concetto nuovo. Gli statistici utilizzano da anni alberi decisionali e regressione lineare e logistica per aiutare le aziende a correlare e classificare i propri dati per ottenerne previsioni.

La novità è la portata dell’analytics predittiva, che si è ampliata. Le rivoluzioni nel campo dell’apprendimento automatico e dell’apprendimento profondo hanno aperto opportunità di utilizzo dei modelli predittivi in ambiti che finora erano impraticabili per la maggior parte dei budget aziendali.

Le imprese assistono oggi a una convergenza inedita di strumenti intuitivi, nuove tecniche predittive e modelli di distribuzione su cloud ibrido che rendono l’analytics predittiva più accessibile che in passato.

Tale situazione ha determinato un punto di svolta. Per la prima volta, aziende di ogni dimensione possono porre in essere le seguenti attività:

- incorporare l’analytics predittiva nei propri processi aziendali
- sfruttare l’AI su larga scala
- trarre valore da "dati oscuri" precedentemente inesplorati, comprendenti qualsiasi cosa, dal testo grezzo fino alle informazioni di geolocalizzazione

Riuscendo ad evolversi da progetti AI confinati in reparti o piccoli gruppi e portandoli a livello di piattaforma di data science aziendale, le organizzazioni hanno davanti a sé notevoli vantaggi competitivi. Chi non approfitta dell’opportunità rischia di rimanere definitivamente indietro.

77,6 miliardi di dollari

saranno spesi nel 2022 in sistemi cognitivi e AI

(Fonte: IDC)

Torna al sommario

Cosa cambia con l'adozione dell'AI

Che tipologie di dati si possono analizzare?

Prima
Principalmente dati relazionali su larga scala; altri tipi di dati richiedevano progetti di ricerca ad hoc.

Ora
Dati relazionali, documenti semistrutturati, testo, dati sensore e altro; è possibile su larga scala sia l'analytics cronologica che quella in tempo reale.

Che strumenti sono disponibili?

Prima
Strumenti disparati e incompatibili fra loro che richiedevano molteplici passaggi di mano tra team dotati di competenze diverse.

Ora
Un mix di interfacce in modalità "drag and drop", e blocchi note open source che rendono molto più praticabile la collaborazione fra team.

Che tecniche analitiche sono utilizzabili?

Prima
Tecniche statistiche di base come la regressione logistica e lineare.

Ora
Tecniche statistiche valorizzate con apprendimento automatico di ultima generazione e algoritmi di apprendimento profondo.

Come implementano le aziende le applicazioni di analytics?

Prima
Applicazioni e analytics erano legate a dati su server on-premise e dispositivi di data warehouse, il che riduceva le opportunità di un'analytics realizzabile in qualsiasi tempo e luogo.

Ora
L'implementazione ibrida e multicloud consente di trasportare l'analytics ovunque risiedano i dati, unendo la sicurezza in loco a flessibilità e scalabilità.

Come possono le aziende integrare l'analytics nei processi aziendali?

Prima
Generando report statici destinati all'analisi manuale da parte di esperti aziendali.

Ora
Integrando totalmente modelli predittivi in nuove app e applicazioni enterprise.

Come implementano le aziende la governance?

Prima
Aderenza ad hoc alle policy a livello di reparto, con minima visibilità e tracciabilità.

Ora
Un framework coerente di governance e sicurezza consente policy a livello di azienda, attuabili su larga scala.

Come le aziende introducono l'intelligenza artificiale nelle moderne applicazioni?

Prima
Disconnessione totale tra sviluppo applicazioni e team di data science, per cui ogni distribuzione era un processo personalizzato.

Ora
Il ciclo di data science è progettato per creare un processo standardizzato e ripetibile per l'integrazione AI.

Come possono le aziende progredire nel loro percorso nell'analytics?

Prima
Ciascuna fase, dalla analytics descrittiva all'analytics predittiva e prescrittiva, richiedeva strumenti, competenze e investimenti separati.

Ora
Una piattaforma integrata supporta la progressione analitica, semplifica l'incorporazione e cresce con l'azienda man mano che questa cresce e sviluppa le competenze.

Torna al sommario

Come può l'AI valorizzare le attività dell'azienda?

In teoria, l'adozione di un approccio moderno all'analytics predittiva dovrebbe essere intuitiva. La tecnologia non è più un ostacolo e migliori strumenti riducono notevolmente le barriere al primo approccio.

Nella pratica, però, può ancora risultare difficile fornire valore aggiunto. Agli stakeholder di un'azienda può facilmente capitare di farsi influenzare dal sentito dire sull'AI e attribuire alla data science aspettative poco realistiche.

Casi d'uso tipici

Il primo obiettivo per data science e dirigenza è quello di collaborare per identificare casi d'uso pratici e concreti in cui la moderna analytics predittiva possa portare valore.

Alcuni, come gli esempi seguenti, possono applicarsi in genere alla maggior parte dei settori:

- raccomandazione prodotti e modelli "next best action" per vendite e marketing
- automazione contact center per assistenza clienti

Altri casi d'uso possono essere specifici di un dato settore, reparto o anche gruppo all'interno dell'azienda. Sono in genere più difficili da eseguire, ma offrono notevoli potenzialità di vantaggi competitivi esclusivi.

Quali funzioni aziendali sono in prima linea negli investimenti in sistemi AI?

46% vendite e marketing

40% assistenza clienti

(Fonte: ricerca Forrester)

Torna al sommario

Casi d’uso generici

Quando un’azienda inizia a investire in una nuova tecnologia, spesso è consigliabile iniziare con il "raccogliere i frutti più vicini", cioè puntare agli obiettivi più a portata di mano.

Lo stesso accade con l’analytics predittiva. Molti casi d’uso sono ampiamente applicabili a vari settori, e i fornitori hanno già predisposto modelli e servizi d’uso generico, preconfezionati.

Questi servizi possono costituire ottimi punti di partenza per le aziende che desiderino trasformare la data science da funzione di ricerca a parte integrante dell’operatività quotidiana. Sono facili da implementare, richiedono una personalizzazione minima e offrono rapidamente valore aggiunto.

Ottimizzazione del contact center

La gestione di volumi imprevedibili di chiamate, e-mail, SMS e messaggi in chat da parte dei clienti è un problema comune nel servizio clienti.

Le chatbot intelligenti sono un modo efficace e conveniente per alleviare la pressione sui dipendenti e ridurre i tempi d’attesa da parte dei clienti. Queste chatbot sfruttano le seguenti funzionalità per comprendere le richieste dei clienti:

- riconoscimento vocale o verbale mediante AI
- elaborazione del linguaggio naturale
- analisi del contenuto per esplorare la knowledge base e trovare risposte utili senza necessità di intervento umano

Alcuni dei casi d’uso transettoriali della moderna analytics predittiva più comuni sono:

- **incremento di cross-sell e up-sell** con raccomandazioni e offerte personalizzate in tempo reale
- **fidelizzazione** grazie alla prevenzione dell’abbandono e agli interventi per prevenirlo
- **ottimizzazione dell’offerta** grazie all’ascolto delle voci dei clienti e all’anticipazione di esigenze future
- **potenziamento del marketing** mediante campagne mirate e personalizzate
- **riduzione dei costi di inventario** e migliore gestione delle risorse con previsioni più accurate
- **aumento della produttività** mediante allocazione dei collaboratori in mansioni e con tempistiche ottimali, e previsioni accurate sulla forza lavoro
- **riduzione dei costi di manutenzione** mediante anticipazione dei guasti prima che si verifichino
- **mitigazione dei rischi** con accurate valutazioni di affidabilità creditizia
- **rilevamento di frodi** grazie all’identificazione di modelli di comportamento sospetti
- **scoperta di nuovi modelli di business** con risposte a segmenti di domanda non ancora sfruttati e integrazione delle previsioni nelle moderne app

Casi d'uso specifici per settore

Le organizzazioni più innovative in numerosi settori stanno già investendo nella costruzione di modelli predittivi proprietari per la risoluzione di problemi aziendali specifici. Le prossime due pagine illustrano quelle che sono solo alcune delle potenziali applicazioni di AI e analytics predittiva presso diversi settori di rilievo.

Servizi bancari a carattere commerciale

Gli istituiti bancari a carattere commerciale utilizzano l'analytics predittiva per le seguenti applicazioni:

- valutazione dei rischi di mercato e delle controparti per operazioni negoziali
- valutazione del rischio di credito per operazioni di prestito
- rilevazione di transazioni fraudolente in tempo reale
- modellazione predittiva per accelerare le procedure di approvazione prestiti

Assicurazioni

Le compagnie assicurative possono utilizzare l'analytics predittiva in particolare per le seguenti applicazioni:

- rilevazione di richieste illegittime di risarcimento
- ottimizzazione di quote e premi mediante valutazione dei rischi specifici per richiedente
- previsione di eventi meteorologici pericolosi ai fini di riduzione delle richieste di risarcimento per assicurazioni auto

Energia e utility

Le utility sfruttano l'analytics predittiva in particolare per le seguenti applicazioni:

- gestione di ampie reti di risorse fisiche
- previsione di pattern di produzione e domanda
- previsione anticipata di interruzioni
- piani relativi a forniture e domanda

Pubblica amministrazione

Le pubbliche amministrazioni si possono affidare a statistiche accurate ai fini di informare la creazione di policy in numerosi ambiti, tra cui i seguenti casi d'uso per l'analytics predittiva:

- rilevazione di illeciti su indennità
- previsione di pattern d'uso di servizi pubblici
- ottimizzazione della raccolta rifiuti e del traffico

Industria manifatturiera

I produttori utilizzano l'analytics predittiva in particolare per le seguenti applicazioni:

- corretto e costante funzionamento delle linee di produzione mediante modellazione della qualità dei prodotti e rilevamento prodotti difettosi
- ottimizzazione della gestione dei magazzini e della logistica
- sviluppo di sensori per veicoli a guida autonoma mediante modelli di apprendimento automatico

Vendita al dettaglio

I dettaglianti utilizzano l'analytics predittiva in particolare per le seguenti applicazioni:

- gestione di programmi di fidelizzazione clienti
- aumento di cross-sell e up-sell mediante raccomandazioni mirate basate sui profili dei clienti e modelli sofisticati di propensione
- previsioni accurate della domanda

Alimentare

L'industria alimentare utilizza l'analytics predittiva per le seguenti applicazioni:

- automazione della raccolta dati e analisi della salute alimentare
- previsione e allerta su potenziali focolai legati alla salute e rapidità d'intervento
- protezione dei dati sensibili dell'azienda per rendere sicura la collaborazione con concorrenti

Sanità

Le organizzazioni del settore sanitario utilizzano tecniche di modellazione statistica per le seguenti applicazioni:

- monitoraggio di flussi di dati da ECG e altri dispositivi medicali
- previsione dei cambiamenti delle condizioni del paziente
- ricerca medica
- analisi in tempo reale di flussi di dati sui pazienti

Retail banking

Le banche retail utilizzano l'analytics predittiva per le seguenti applicazioni:

- accrescere la soddisfazione dei clienti mediante una più rapida valutazione di affidabilità creditizia
- unire flessibilità, solidità e sicurezza con un'infrastruttura cloud ibrida
- ridurre i costi e accelerare lo sviluppo grazie a un'architettura innovativa

Trasporti

Le aziende di trasporti e logistica utilizzano l'analytics predittiva per le seguenti applicazioni:

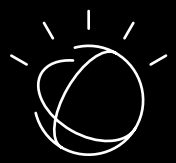
- ottimizzazione della pianificazione degli itinerari
- manutenzione predittiva dei veicoli
- ottimizzazione delle operazioni legate alla catena di fornitura

Istruzione

Le istituzioni del settore dell'istruzione possono utilizzare l'analytics predittiva in particolare per le seguenti applicazioni:

- previsione dei tassi di successo e abbandono degli studenti
- identificazione di studenti che necessitino di supporto speciale
- rafforzamento di rapporti con donatori
- verifica dei movimenti degli studenti per la riduzione dell'assenteismo

I gradini della scala dell'AI



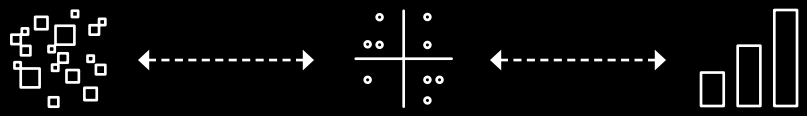
Infondere - Rendere operativa l'AI con fiducia e trasparenza

Analizzare - Scalare gli insight ovunque con l'AI

Organizzare - Creare una base fidata di analytics

Raccogliere - Rendere i dati semplici e accessibili

Dati di ogni tipo, indipendentemente da dove risiedano



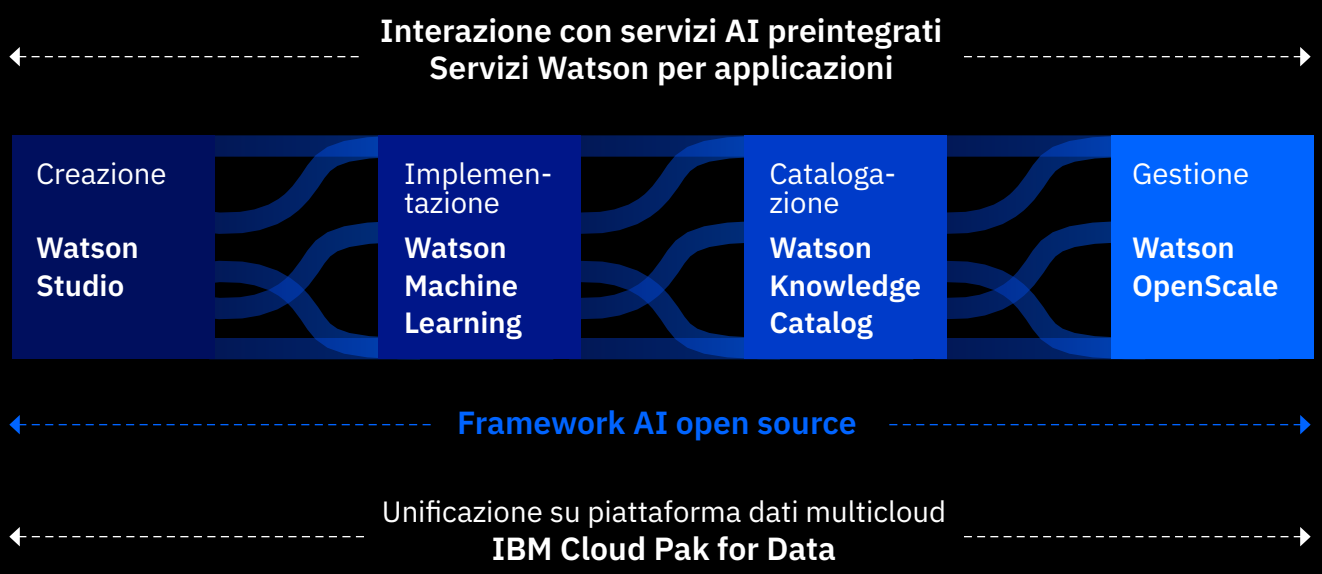
Raggiungere il successo mediante la moderna analytics predittiva è un percorso. È importante modulare la strategia AI al giusto livello per un'azienda, considerandone la maturità sia tecnica che organizzativa. I responsabili data science e aziendali devono collaborare per definire il modo migliore e più rapido per ottenere valore aggiunto.

Dal punto di vista tecnico, la maturità AI è visualizzabile in forma di scala. Il primo gradino della scala è la raccolta dei dati, poiché senza dati non si ha nulla da analizzare o modellare. Il secondo gradino è l'organizzazione dei dati. Aggiungere i metadati per governance e rilevabilità, per far sì che i dati necessari siano sempre disponibili per gli analisti scientifici dei dati.

Ma benché la raccolta e l'organizzazione dei dati siano elementi importanti, vanno al di là dell'ambito di interesse di questa guida. Concentriamoci invece su come scalare i seguenti due gradini più alti:

- analisi dei dati mediante costruzione, addestramento e test di modelli predittivi
- infusione dell'AI nell'operatività implementando e sviluppando tali modelli nella produzione nel contesto delle proprie applicazioni

Quali sono le soluzioni disponibili?



Il portfolio AI di IBM offre tutto ciò che occorre per raggiungere i gradini più alti della scala dell'AI.

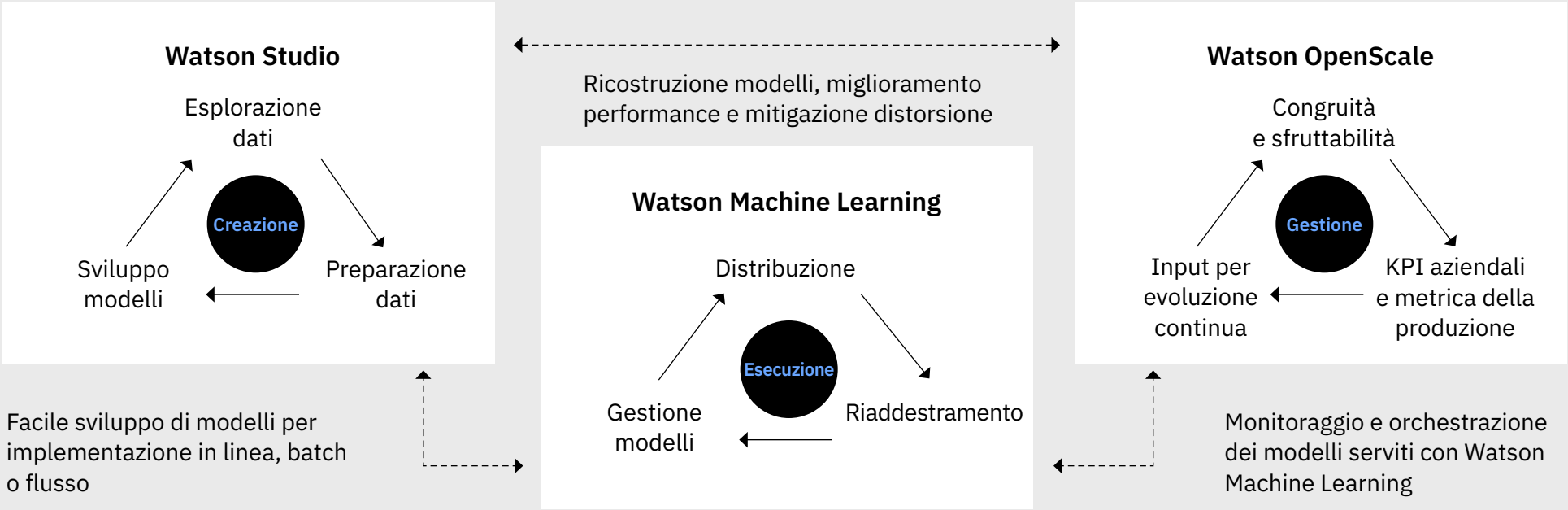
I servizi pre-integrati come Watson Assistant e Watson Visual Recognition aiutano ad affrontare rapidamente ed efficacemente casi d'uso comuni, fornendo valore immediato.

A chi è pronto a sviluppare le proprie soluzioni AI, Watson Studio e Watson Machine Learning offrono workflow ininterrotti per la creazione, l'addestramento e l'implementazione di modelli predittivi. Queste soluzioni arricchiscono le risorse mediante strumenti IBM di ultima generazione e i migliori framework AI open source.

Watson Knowledge Catalog assicura una solida governance dei dati e rilevabilità per modelli e dati, mentre Watson OpenScale aiuta a monitorare e gestire i modelli in tempo reale, ottimizzando l'accuratezza, accrescendo la spiegabilità e mitigando la distorsione.

IBM Cloud Pak for Data unifica l'accesso a tutte e tre queste funzionalità e fornisce una potente piattaforma dati multicloud.

Il componente aggiuntivo IBM Data Science Premium per IBM Cloud Pak for Data fornisce le funzionalità supplementari di produttività della data science quali la SPSS Modeler and Decision Optimization per accelerare il time-to-value e accrescere le possibilità di successo del progetto AI/ML.



I due gradini più alti della scala dell'AI sono Analizzare e Infondere. Per raggiungere questi gradini, le organizzazioni devono aiutare gli analisti scientifici dei dati e gli stakeholder a collaborare efficacemente in ogni fase del ciclo di vita della data science.

Il ciclo di vita completo può essere visualizzato come tre sottocicli che interagiscono fra loro:

Creazione

Gli analisti scientifici dei dati studiano i dati aziendali per identificare caratteristiche interessanti; preparano quindi dati ben strutturati da usare per progettare i modelli predittivi.

Esecuzione

I team operativi addestrano, testano, implementano e gestiscono i modelli, riaddestrando quando necessario.

Gestione

Gli esperti di business monitorano le performance di runtime dei modelli, ricercano eventuali segnali di distorsione o necessità di spiegazione, forniscono feedback e notificano il team di data science della necessità di riaddestramento.

Il passo successivo

A seconda del grado di avanzamento nella scala dell'AI, le aziende possono avere requisiti diversi in base al livello di adozione dell'analytics predittiva nell'intera azienda.

Per cominciare

Quando le aziende iniziano a creare le proprie funzionalità di data science, spesso partono da progetti ad hoc, sviluppando modelli che rispondano a questioni specifiche o che supportino progetti di ricerca. Con soluzioni come Watson Studio Desktop, gli analisti scientifici dei dati possono operare 24/7 sui propri computer e laptop e sincronizzare con un ambito più ampio di collaboratori in caso di necessità.

Per crescere

Quando la data science è applicata in modo ampio, i diversi reparti devono sviluppare i propri modelli, connetterli alle origini dati e infonderli nelle applicazioni di produzione. Watson Studio e Watson Machine Learning facilitano la collaborazione fra data science e IT della divisione, lungo tutto questo ciclo di vita.

Per scalare sull'intera azienda

Una volta incorporata l'AI nei processi business critical, è fondamentale creare una piattaforma centralizzata al fine di gestire e governare modelli e dati. IBM Cloud Pak for Data può fornire l'infrastruttura e gli strumenti necessari per una piattaforma completa e multicloud che operi da singolo punto di controllo.

Per passare alla pratica

In veste di analista scientifico dei dati o di dirigente, il miglior modo per sapere come la gamma di soluzioni di analytics predittiva moderna di IBM possa trasformare la tua azienda è scoprirla di persona. Per iniziare, prova uno dei seguenti tutorial.

Svolgiamo un esercizio di apprendimento automatico

Scopri l'apprendimento automatico eseguendo un esercizio in IBM Watson Studio, utilizzando Apache SystemML. Scopri di più

Creiamo un modello di valutazione per predire un'insufficienza cardiaca

Utilizza IBM Watson Studio per creare un modello predittivo con IBM Watson Machine Learning. Scopri di più

Prevediamo un guasto a un'apparecchiatura mediante i dati del sensore IoT

Scopri come IBM Watson Studio riesce ad analizzare i dati di un sensore multivariato IoT (Internet of Things) prevedendo il guasto all'apparecchio. Scopri di più

Analizziamo set di dati medicali aperti per ottenere informazioni utili

Usa IBM Watson Studio per eseguire classificatori di apprendimento automatico e confrontare i risultati con misure di valutazione. Scopri di più

Plasmiamo e affiniamo dati non elaborati

Lavora con IBM Data Refinery per predisporre vasti set di dati ai fini dell'analisi predittiva. Scopri di più

I punti chiave

Il portfolio IBM di moderna analytics predittiva offre i seguenti vantaggi, che data science e dirigenza possono sfruttare per ottenere vantaggi competitivi nell'era dell'AI.

Scalabilità

- Riduzione del workload e dei costi operativi mediante automatizzazione delle attività di data science e ingegneria dei dati
- Addestramento, test e sviluppo di modelli in modo continuativo presso molteplici applicazioni enterprise
- Espansione delle comuni funzionalità di data science su ambienti multicloud ibridi

Velocità

- Velocizzazione dell'implementazione grazie ad applicazioni preintegrate e modelli preaddestrati
- Creazione di valore più rapida grazie alla collaborazione fra data science e team aziendali
- Razionalizzazione della creazione dei modelli con la combinazione di software di ultima generazione IBM e open source

Semplicità

- I vantaggi di una piattaforma centralizzata per la gestione dell'intero ciclo di vita della data science
- Standardizzazione dei processi di sviluppo e distribuzione
- Creazione di governance dei dati e sicurezza su un unico framework in tutta l'azienda

Watson Studio aiuta le aziende a concentrarsi sulla risoluzione dei problemi e sull'individuazione delle opportunità.

Scopri di più

Watson Machine Learning offre alle aziende gli strumenti idonei per implementare e gestire modelli in grado di fornire rapidamente i risultati di cui hanno bisogno.

Scopri di più

Perché combinare l’ottimizzazione decisionale con l’analytics predittiva?

IBM Decision Optimization è una soluzione di analytics prescrittiva che consente alle aziende che fanno elevato affidamento sui dati di prendere decisioni più informate e raggiungere gli obiettivi risolvendo complesse problematiche di ottimizzazione. I leader utilizzano questi strumenti per migliorare l’efficienza nel proprio uso delle risorse, incluse le seguenti attività, e non solo:

- flusso di magazzino per catena di fornitura
- programmazione della forza lavoro
- instradamento trasporti

La soluzione funziona bene con l’analytics predittiva sfruttando i risultati predittivi delle applicazioni di apprendimento automatico, per fornire risultati ottimizzati. L’apprendimento automatico fornisce insight sul futuro basati sulle osservazioni fornite dagli utenti. Con l’apprendimento automatico, si conosce la risposta e si addestra la macchina a trovarla.

L’ottimizzazione decisionale consente di passare alla fase successiva e dar seguito esecutivo all’informazione ottenuta. Grazie all’ottimizzazione decisionale, anche se non si conosce la risposta, si sa quale sia una risposta buona e una non buona. Si ricava l’output dall’apprendimento automatico e si specifica un’azione da eseguire, conseguente all’ottimizzazione decisionale, che può comprendere anche regole e limitazioni di ottimizzazione, per ottenere i risultati aziendali desiderati.

A seguito di tale azione, l’ottimizzazione decisionale restituisce risposte che forniscono valore aggiunto all’azienda, quali elementi attivabili e consigli di modifiche. Mediante questa attività, l’ottimizzazione decisionale migliora ciò che l’analytics predittiva può offrire.

La soluzione consente ai team di combinare le tecniche di ottimizzazione e di apprendimento automatico con la gestione dei modelli, la distribuzione e altre funzionalità di data science, per sviluppare soluzioni ottimali in grado di migliorare l’efficienza operativa.

Scopri di più

Glossario

Gli **algoritmi** sono insiemi di regole che definiscono una sequenza di operazioni applicabili ai dati per risolvere un particolare problema. In un contesto di data science, il termine comprende una vasta gamma di tecniche, tra cui le seguenti:

- alberi decisionali e modelli di regressione
- media mobile autoregressiva (ARMA), media mobile autoregressiva integrata (ARIMA) e livellamento esponenziale
- funzioni di trasferimento con predittori e rilevamento di valore anomalo
- modello di classificazione binario e gerarchico
- vector machine e modello causale temporale
- serie storiche e AR spaziale per predizione spaziotemporale
- reti generative avversarie (GAN) e reinforcement

La piattaforma di data science dovrebbe fornire facile accesso a tutti questi potenti algoritmi.

Intelligenza artificiale (AI) è la capacità dei sistemi informatici di interpretare i dati e apprendere da essi. L'espressione è usata più comunemente per descrivere sistemi costruiti usando modelli di apprendimento automatico o apprendimento profondo.
Le tecniche di AI possono essere usate per mettere i computer in grado di risolvere un'ampia gamma di problemi prima considerati intrattabili.

Bias è un problema comune quando si progettano, si addestrano e si testano modelli che possono portare a previsioni sbagliate. La mitigazione del bias mediante il monitoraggio e i modelli di verifica durante il runtime è un argomento di importanza crescente per le aziende che intendono adottare l'AI in modo più esteso.

I **modelli di classificazione** mirano a collocare i punti dati all'interno di categorie confrontandoli con un insieme di punti dati già categorizzati. Il risultato è un valore discreto, intendendo con ciò un valore compreso in un elenco limitato di opzioni, piuttosto che un punteggio. Per esempio, un modello di classificazione può rispondere "sì" o "no" alla domanda se è probabile che i clienti facciano un acquisto o siano cattivi pagatori. I modelli di classificazione possono essere costruiti utilizzando varie tecniche, tra cui gli alberi decisionali e la regressione logistica.

Analisi dei contenuti è l'analisi dei dati non strutturati presenti in documenti di vari formati, compresi testi, immagini, file audio e video. Le tecniche di apprendimento automatico possono accelerare molto l'analisi di grandi depositi di contenuti che in precedenza, per essere esaminati e classificati, avrebbero impegnato i lavoratori per centinaia o migliaia di ore.

Data science (scienza dei dati) è una disciplina ad ampio raggio che unifica aspetti di statistica, analisi dei dati e apprendimento automatico per tenere sotto controllo i dati e risolvere problemi aziendali.

Apprendimento profondo è un ramo dell'apprendimento automatico che usa reti neurali con un grande numero di strati nascosti. Queste reti altamente sofisticate sono usate in settori all'avanguardia dell'apprendimento profondo come la visione artificiale, la traduzione automatica e il riconoscimento vocale.

Addestrare una rete neurale profonda è un'operazione a elevato utilizzo di calcolo, che richiede tipicamente cluster di macchine con processori ad alte prestazioni. Una piattaforma per cloud ibrido come IBM Watson Studio o IBM Cloud Pak for Data può rendere questo tipo di infrastruttura più accessibile ed economica per aziende di ogni dimensione.

Deployment è il processo di integrazione di un modello con le applicazioni aziendali e della sua esecuzione su dati del mondo reale. Costruire un modello e fargli attraversare gli ambienti di test, di preparazione e di produzione richiede la collaborazione fra team di data science, sviluppatori di applicazioni e operazioni IT.

Integrare strumenti di data science open source con la pipeline di integrazione continua e di deployment esistente dell'azienda può rivelarsi impegnativo. Per evitare i deployment manuali con molteplici passaggi di consegna fra team, una piattaforma data science coerente con funzionalità di deployment automatizzato può costituire un significativo vantaggio.

Lo **sviluppo** di modelli predittivi prevede l'uso di tecniche statistiche tradizionali o algoritmi di apprendimento automatico per la creazione e l'affinamento di modelli mediante addestramento e test a confronto con i propri set di dati.

Questo processo di sviluppo è altamente iterativo; può essere necessario effettuare l'addestramento di diverse decine o anche centinaia di modelli per raggiungere il livello di accuratezza desiderato. Per questo, l'automatizzazione dei workflow intorno allo sviluppo e all'addestramento dei modelli può costituire un enorme valore aggiunto.

Spiegabilità è un'importante caratteristica di ogni sistema che utilizzi modelli predittivi per formulare raccomandazioni e fornire assistenza al processo decisionale aziendale. Un modello predittivo visto come complesso e misterioso non conquisterà la fiducia degli investitori, degli enti di controllo e dei clienti. Le funzionalità avanzate di monitoraggio del runtime e di registrazione di Watson OpenScale forniscono il contesto intorno a ciascuna decisione, rendendo i modelli AI trasparenti e verificabili.

Esplorazione dei dati è una componente importante del processo di costruzione dei modelli. Questa attività punta a evidenziare caratteristiche interessanti in una certa base di dati, rivelare relazioni nascoste e mettere in luce casi d'uso dove la modellazione predittiva potrebbe fornire valore commerciale.

Durante la fase esplorativa, è fondamentale esercitare le competenze di data science e le conoscenze dell'azienda per definire le domande alle quali si desidera ottenere risposta e i risultati che si intendono prevedere. Ciò può portare a un ciclo iterativo di preparazione ed esplorazione finché non si sia completamente esplorato il dominio e non si disponga dei dati nella forma adeguata a procedere.

Analisi geospaziale è l'analisi dei dati geografici come la latitudine e la longitudine, i codici postali e gli indirizzi. Questa analisi è estremamente utile per risolvere molti tipi di problemi pratici di data science. Una moderna piattaforma di data science dovrebbe rendere facile rilevare, analizzare sintatticamente e calcolare le informazioni geospaziali e dovrebbe presentare una comoda integrazione con strumenti di mappatura per visualizzare i risultati.

Inferenza nell'intelligenza artificiale applica regole logiche alla knowledge base per trarre conclusioni in presenza di incertezza. Con l'inferenza, gli utenti ottengono una previsione che è semplificata, compressa e ottimizzata per le prestazioni di runtime.

Regressione lineare è un processo statistico che usa una variabile indipendente per spiegare o prevedere un valore o un punteggio. Esempi comprendono il numero di SKU (articoli gestiti a magazzino) di un prodotto venduto in una certa settimana o la percentuale di rischio che un cliente chiuda il suo account.

Regressione logistica è una procedura statistica utilizzata per la previsione degli esiti. Differisce dalla regressione lineare in quanto l'unica variabile indipendente presenta un numero limitato di possibili valori anziché infinite possibilità. Gli utenti sfruttano la regressione logistica quando la risposta si situa in categorie quali la numerazione ordinale, cioè primo, secondo, terzo e via dicendo.

Apprendimento automatico utilizza tecniche statistiche per derivarne modelli predittivi e algoritmi sofisticati, da ampi dataset, senza necessità di programmazione specifica.

Tale processo iterativo è tipicamente avviato dividendo un dataset in due sottoinsiemi, per l'addestramento e i test. Si addestra il modello in parallelo con il set di addestramento e se ne testano le performance a confronto, con decine o centinaia di variazioni al fine di valutarne l'accuratezza predittiva. Eseguendo tale procedura e basando la prossima generazione di variazioni sulle voci dalle migliori performance da ciascuna iterazione, il modello impara gradualmente e migliora i risultati.

Possiamo dividere gli approcci all'apprendimento automatico in due categorie: apprendimento supervisionato e apprendimento senza supervisione.

Gestione dei modelli è fondamentale per assicurarne l'accuratezza nel tempo. È importante riaddestrare costantemente i modelli per tenere conto dei nuovi dati: sviluppo, implementazione, distribuzione e gestione del modello devono costituire un ciclo ininterrotto.

Se si usano strumenti disparati e open source, tale gestione può rivelarsi difficile. L'uso di una piattaforma end-to-end di analisi scientifica dei dati può prevenire lacune nel processo. La piattaforma, inoltre, è in grado di assicurare la notifica immediata ai team pertinenti affinché possano agire rapidamente al primo degradarsi delle performance di un modello.

La **NLP (elaborazione del linguaggio naturale)** è un settore dell'AI che si occupa soprattutto di consentire ai computer l'analisi di dati testuali non strutturati. Tra i casi d'utilizzo comuni ricordiamo il riconoscimento vocale, la comprensione del linguaggio naturale e l'analisi delle sensazioni.

Le **reti neurali** forniscono un framework per i modelli di addestramento e consentono un'interazione complessa tra numerosi algoritmi di apprendimento automatico, per contribuire a identificare i modelli ottimali.

La struttura delle reti neurali artificiali si è ispirata alla struttura interconnettiva dei neuroni del cervello umano e di altri animali. I neuroni artificiali sono collegati in strati. I dati attraversano la struttura partendo dal livello di input fino a quello di output, passando da uno o più livelli nascosti. Durante tale percorso, le funzioni matematiche trasformano i dati in previsione, la cui accuratezza è valutabile.

Il **software open source** è divenuto un paradigma sempre più predominante in molti ambiti della modellazione statistica e dell'apprendimento automatico. Linguaggi come R, Python e Scala, architetture big data quali Apache Hadoop e Spark, e framework di apprendimento automatico come TensorFlow e Spark MLlib, sono tutti player di rilievo nel mondo dell'analiytics predittiva e della data science.

I framework open source sono spesso incentrati sullo sviluppo degli strumenti di alta qualità rivolti a parti specifiche del processo di data science, come lo sviluppo o l'addestramento di modelli. Ne consegue che l'utente finale si ritrova responsabile dell'integrazione di tutti gli strumenti per ottenere un workflow coerente. Ciò può presentare problemi quando si cerca di dimensionare l'analytics predittiva presso l'intera azienda e di incorporare l'AI nei processi aziendali.

Analytics predittiva utilizza i dati cronologici per modellare un dominio o un problema specifico e isolare i fattori chiave che nel passato abbiano indotto esiti specifici. I modelli basati su questa procedura prevedono i probabili esiti futuri a partire da nuovi dati.

L'analytics predittiva può includere un'ampia gamma di tecniche, dalla modellazione statistica classica agli algoritmi di apprendimento automatico.

Modelli predittivi sono algoritmi che mappano un input, cioè una porzione di dati, come un record del database, un campione di testo o un'immagine, verso un output o una previsione. Gli output sono tipicamente o variabili continue, quali cifre o percentuali, o categorie discrete, come "sì" o "no". Esistono due tipologie principali di modelli predittivi: modelli di regressione e modelli di classificazione.

Preparazione dei dati è una delle prime fasi del processo di data science. La maggior parte dei progetti inizia dall'affinamento dei dataset, affinché la qualità sia sufficientemente elevata per supportare un'analisi dettagliata.

In molti casi, i dati di origine necessitano di una ripulitura e di essere trasformati in un formato più adatto alla modellazione e all'analisi. Per costruire un modello di apprendimento automatico, può essere anche necessario investire nell'etichettatura manuale dei dati ai fini di utilizzo nell'apprendimento supervisionato.

Modelli di regressione sono utili quando si dispone di un dataset contenente molteplici variabili e si desidera analizzarne le interrelazioni. In particolare, i modelli di regressione sono in grado di rivelare quanto una variante specifica potrà probabilmente modificarsi ove altre variabili siano alterate.

La regressione lineare può essere utilizzata per prevedere un valore o un punteggio. Esempi comprendono il numero di SKU di un prodotto venduto in una certa settimana o la percentuale di rischio che un cliente chiuda il suo account.

Modellazione statistica è un ramo della matematica che si occupa tra l'altro della creazione di modelli basati su ipotesi probabilistiche relative a un dataset. Le aziende dispongono di modelli statistici per l'analisi di importanti funzioni nei loro dataset e per l'identificazione di correlazioni utilizzabili per classificare i dati o generare previsioni.

Apprendimento supervisionato è un metodo per addestrare un modello di apprendimento automatico mediante un dataset nel quale i dati siano già stati correttamente etichettati. Il modello produce una variabile di output, tipicamente categoria o valore, tale che la sua accuratezza è facilmente valutabile mediante il confronto tra l'output e l'input etichettato. La regressione lineare, le random forest e le support vector machine sono noti esempi di algoritmi di apprendimento supervisionato, e la maggior parte dei modelli predittivi è costruita grazie a queste tecniche.

Testing dei modelli predittivi è essenziale per poter determinare l'accuratezza dei dati nei processi AI nel corso dell'addestramento. I modelli predittivi vanno continuamente testati per migliorarne l'accuratezza. Se un modello fallisce, gli analisti devono identificare la causa principale, riaddestrare e testare nuovamente per migliorare i modelli.

Analytics del testo misura il contenuto non strutturato mediante regole linguistiche, elaborazione del linguaggio naturale e apprendimento automatico. Tale processo rivede i dati in modo analogo a quello di un cervello umano, ma più rapidamente. Mediante l'analytics del testo è possibile ottenere maggiori insight e spunti dal contenuto non strutturato, che costituisce di fatto il 90% circa di tutti i dati.

Addestramento dei modelli predittivi è un elemento chiave dell'apprendimento automatico, dell'apprendimento profondo e di altri processi AI al fine di determinare quali dati siano utili. Un modello addestrato per fornire previsioni accurate può essere utilizzato per attribuire punteggi a dati in tempo reale. I modelli vanno riaddestrati periodicamente per adeguarli alla mutevolezza dei pattern comportamentali.

L'**apprendimento senza supervisione** è un metodo di addestramento di modelli di apprendimento automatico con dati non etichettati. Lo scopo è tipicamente di modellare ed evidenziare pattern e strutture interessanti all'interno dei dati. I problemi di clustering e associazione sono comuni nell'apprendimento non strutturato: per trovare, ad esempio, nuovi modi interessanti di segmentare i clienti o per individuare affinità.

La **visualizzazione** è la rappresentazione grafica dei dati, sovente con uso di tabelle e diagrammi. Per comprendere i dati, gli esseri umani hanno bisogno di visualizzarli. Questo processo è importante sia nella presentazione dei risultati agli stakeholder aziendali, sia se si esplora un nuovo dataset durante le fasi iniziali di un progetto.

La piattaforma di analytics predittiva deve prevedere un'interfaccia grafica intuitiva dotata di strumenti di visualizzazione. Così facendo, è possibile dare in pochi minuti senso e comprensibilità anche ai dataset di maggiori dimensioni.

© Copyright IBM Corporation 2019

IBM Italia S.p.A.
Circonvallazione Idroscalo
20054 Segrate (Milano)
Italia

Prodotto negli Stati Uniti d'America
Marzo 2019

IBM, il logo IBM, ibm.com, IBM Cloud, IBM SPSS Modeler e IBM Watson sono marchi di International Business Machines Corp., registrati in molti Paesi del mondo. Altri nomi di prodotti e servizi potrebbero essere marchi di IBM o di altre aziende. Altri nomi di prodotti e servizi potrebbero essere marchi di IBM o di altre aziende. Un elenco aggiornato dei marchi IBM è consultabile sul web alla pagina "Copyright and trademark information" disponibile all'indirizzo: www.ibm.com/legal/copytrade.shtml.

Le informazioni contenute nel documento sono aggiornate alla data della prima pubblicazione e potrebbero essere modificate da IBM senza alcun preavviso. Non tutte le offerte sono disponibili in ogni Paese in cui IBM opera.

I dati relativi alle performance qui menzionati sono da considerarsi derivati da condizioni operative specifiche. I risultati effettivi potrebbero variare. LE INFORMAZIONI FORNITE NEL PRESENTE DOCUMENTO SONO DA CONSIDERARSI "NELLO STATO IN CUI SI TROVANO", SENZA GARANZIE, ESPLICITE O IMPLICITE, IVI INCLUSE GARANZIE DI COMMERCIABILITÀ, DI IDONEITÀ PER UN PARTICOLARE SCOPO E GARANZIE O CONDIZIONI DI NON VIOLAZIONE. I prodotti IBM sono coperti da garanzia in accordo con termini e condizioni dei contratti sulla base dei quali vengono forniti.

È responsabilità del cliente assicurare la conformità a normative e regolamenti applicabili. IBM non fornisce consulenza legale né dichiara o garantisce che i propri servizi o prodotti assicurino al cliente la conformità con qualsivoglia legge o regolamento.

Dichiarazione di applicazione di corrette prassi per la sicurezza: la sicurezza dei sistemi IT implica la protezione di sistemi e dati mediante prevenzione, rilevamento e risposta ad accessi impropri all'interno e all'esterno dell'azienda. Gli accessi impropri possono causare alterazione, distruzione, appropriazione indebita o abuso dei dati e danni o abuso dei sistemi, anche per essere utilizzati per attacchi verso terzi. Nessun sistema o prodotto IT va considerato totalmente sicuro e nessun singolo prodotto, servizio, o misura di sicurezza è da considerarsi completamente efficace nella prevenzione dell'uso o dell'accesso improprio. I sistemi, i prodotti e i servizi IBM sono progettati per far parte di un approccio legittimo e completo alla sicurezza, il quale implica necessariamente procedure operative supplementari, e potrebbe richiedere altri sistemi, prodotti o servizi per fornire la massima efficacia. IBM NON GARANTISCE CHE SISTEMI, PRODOTTI O SERVIZI SIANO ESENTI DA O RENDERANNO L'AZIENDA ESENTE DA, CONDOTTA MALEVOLA O ILLEGALE DI UNA QUALSIASI PARTE.