

**La data science è uno sport di squadra.
Hai le capacità per entrare a far parte
del team?**

IBM

1

Introduzione

2

Il data scientist

3

**L'ingegnere
informatico**

4

**Lo
sviluppatore**

5

**L'analista
aziendale**

6

**I team di data
science: i nuovi
protagonisti del
cambiamento**

7

Risorse

Introduzione

La sfida più impegnativa della rivoluzione dei Big Data è riuscire a interpretare tutte le informazioni generate dalla vastissima economia digitale di oggi. Va benissimo che un'azienda raccolga ogni più piccola porzione di dati che sia in grado di raggiungere, ma come può riuscire ad estrarre valore da questo enorme volume di informazioni?

Maggiore è la quantità dei dati di cui si dispone e migliore sarà la qualità dei rapporti e delle raccomandazioni strategiche, giusto? Certo, questo se si riesce ad analizzare i dati in maniera intelligente e rapida e a trasformarli in utili informazioni pratiche. In caso contrario, più dati possono voler dire più problemi: dati confusi, difficoltà di archiviazione, rischi



delle sicurezza, team di business frustrati e personale IT oberato di lavoro.

Come si fa quindi a scoprire ciò che si vuole realmente ricavare dai dati e quali sono i dati che possono garantire tali risposte?

Come si fa a trasformare i dati grezzi in informazioni che abbiano (e generino) un effettivo valore di business? Hai le capacità che il tuo business richiede per generare i risultati?

Le aziende di successo formano team di data science che integrano serie di capacità e responsabilità diverse anziché affidarsi a singoli individui di punta.

Nella pratica, più persone lavorano in un team per ottenere nuove informazioni di valore. La qualità delle analisi sarà pari soltanto al team responsabile della raccolta, della costruzione, dell'analisi e del supporto dei dati.

Qual è il contributo offerto dal team di data science? Chi si occupa di programmazione più approfondita, come gli ingegneri informatici e gli sviluppatori front-end, conferisce alla data science la completa fruibilità dei dati. I matematici e i data scientist utilizzano gli algoritmi statistici

per individuare modelli nei dati. Lungo l'intero processo, tutto quanto deve allinearsi con il risultato di business impostato, che viene guidato dall'occhio dell'analista aziendale. Attraverso la collaborazione, il team di data science può superare le sfide e i problemi di oggi per creare nuove opportunità e possibilità per domani.

Quali competenze e capacità definiscono i membri di un team di data science e come si integrano reciprocamente?

[Continua a leggere per scoprirlo.](#)

Il data scientist

Risolvere i problemi e rispondere alle domande tramite l'analisi dei dati sta rapidamente diventando la norma nel mondo di oggi dominato dai dati. Mentre gli esperimenti del mondo reale assumono caratteristiche di ubiquità nel business moderno, il data scientist evolve nel ruolo destinato ad avviare, aggiustare e alimentare questo motore operativo.

Un buon data scientist non si limita a risolvere i problemi di business ma si impegna a mettere a fuoco i problemi che pesano maggiormente sull'organizzazione.

I data scientist vengono spesso definiti “unicorni” perché rappresentano una rara combinazione di talenti: gestiscono una vasta gamma di responsabilità e serie di competenze negli ambiti di matematica, statistica, esperienze di settore, di comunicazione e altro. In sostanza, il lavoro del data scientist è quello di andare alla ricerca dei modelli nascosti. Per farlo, applica tecniche di analytics avanzate tra cui, ad esempio, apprendimento automatico, modellazione, statistica e visualizzazione. Spesso i data scientist costruiscono modelli per prevedere i risultati o scoprire i modelli sottostanti; la loro strategia è quella di produrre informazioni di immediato valore pratico che possono essere utilizzate per migliorare i risultati futuri.



Poiché i data scientist sono coinvolti in tutti i passaggi del processo di creazione dei prodotti informatici, tenderanno ad applicare una visione olistica alla soluzione dei problemi con i dati. Tuttavia, non possono essere esperti in ogni ambito, ed è qui che il team può dare un enorme aiuto.

Conducono esperimenti continui implementando nuovi modelli predittivi, regole aziendali e logica di orchestrazione nelle applicazioni che puntano alle migliori azioni future.

Un data scientist competente esplora ed esamina i dati provenienti da più fonti. Esamina accuratamente tutti i dati in entrata con l'obiettivo di mettere in relazione le nuove informazioni con i dati storici, per individuare una relazione o un andamento capaci di offrire un vantaggio competitivo

cruciale o di risolvere un problema aziendale pressante. Non si limitano a raccogliere e a inviare le segnalazioni sui dati; li osservano da più punti di vista, ne determinano il significato e suggeriscono quindi metodi per applicare i risultati. Devono accertarsi che le loro query siano corrette e devono essere in grado di supportare le proprie conclusioni con solidi modelli e dati affidabili; ai data scientist si chiede infatti spesso di fornire delle raccomandazioni alla direzione e ai team esecutivi.



Capacità principali per i data scientist

I data scientist si distinguono per la loro comprensione del business, oltre alla capacità di comunicare i risultati tanto ai responsabili aziendali quanto ai responsabili IT, tanto da incidere nel modo in cui le aziende affrontano le loro sfide. Il data scientist diventa spesso il collegamento tra il reparto IT e i dirigenti di grado più alto. Quindi il data scientist deve essere in grado di parlare entrambe le “lingue” e di comprendere la gerarchia dei dati: non può essere semplicemente un esperto di dati. Questo significa anche che i data scientist devono avere una chiara comprensione del business e una forte convinzione che li metta in grado di supportare i risultati a fronte di qualsiasi opposizione.

I data scientist sono indagatori e curiosi: esplorano, pongono domande, eseguono analisi what-if e mettono in discussione le ipotesi e i processi esistenti. Le capacità tecniche di un data scientist includono spesso diversi linguaggi di programmazione, familiarità con gli strumenti di gestione e analisi dei Big Data come Apache Hadoop e Spark ed esperienza con strumenti che consentono di visualizzare dati e informazioni.

L'ingegnere informatico

Il data scientist è certamente responsabile dell'individuazione dei modelli che si celano nei dati, ma dove ottiene i dati e cosa succede quando le loro soluzioni devono essere dimensionate per migliaia di utenti o devono gestire informazioni sensibili? Senza il supporto che permette loro di rendere operativo il proprio lavoro, i data scientist possono fare ben poco.

Ed è qui che entra in gioco l'ingegnere informatico. Ad alti livelli, gli ingegneri informatici contribuiscono a raccogliere, organizzare e pulire i dati che i data scientist utilizzeranno poi per costruire le loro analisi. Se da un lato i data scientist estraggono valore dai dati, dall'altro gli ingegneri informatici si assicurano che i dati si riallineino in modo fluido dalla fonte alla destinazione per poter essere elaborati.

Gli ingegneri informatici sono responsabili della configurazione dei sistemi e dei processi che altri operatori di dati, compresi i data scientist, utilizzano e su cui fanno affidamento per lavorare con i dati. Gli ingegneri informatici devono comprendere come perfezionare il flusso di dati per ridurre le latenze degli spostamenti e conferire agilità all'analytics. Collaborano inoltre con gli sviluppatori front-end nello spostamento dei progetti di data science alla produzione.

In molte organizzazioni, un ingegnere informatico è responsabile dell'integrazione dei dati, compresa la progettazione, la costruzione e la misurazione delle pipeline di inserimento e integrazione dati per grandi volumi di dati temporanei provenienti da fonti diverse. Tra gli esempi si contano estratti di database, log dei server alle applicazioni,

Agli ingegneri informatici viene spesso assegnato il compito di gettare le fondamenta affinché un analista dati o un altro data scientist possa recuperare facilmente i dati necessari per le proprie valutazioni ed esperimenti.

immagini digitalizzate, registrazioni vocali, flussi di Twitter, siti web e dati sanitari che provengono da sensori. Una volta stabilite le connessioni continue, da e verso questi enormi "pool" di informazioni filtrate, i data scientist possono estrarre le serie di dati relative alle loro analisi.

Capacità principali richieste agli ingegneri informatici

Gli ingegneri informatici sono ingegneri esperti che comprendono il funzionamento interno del software di database. Compilano e installano i sistemi per database, scrivono query complesse, le ridimensionano per più macchine, gestiscono i backup e implementano i sistemi di disaster recovery. Sviluppano, costruiscono, collaudano e mantengono architetture quali database e sistemi di elaborazione dati di grandi dimensioni.

I migliori ingegneri informatici non smettono mai di imparare e di pensare alle nuove tecnologie che potrebbero consentire di promuovere il business. Questo atteggiamento li spinge a sviluppare un robusto background di programmazione e ad approfondire le tecnologie basate su Hadoop come MapReduce, Hive e Pig. Gli ingegneri informatici hanno spesso una notevole esperienza con le tecnologie basate su SQL e con le tecnologie NoSQL, come pure con le metodologie di data warehousing e con le soluzioni di estrazione, trasformazione e caricamento (ETL).



Lo sviluppatore

Anche se gli sviluppatori non sempre hanno una posizione totalmente dedicata al team di data science, restano comunque una tessera essenziale del puzzle. Un progetto riuscito trasforma in prodotto il lavoro di data science affinché possa essere utile a uno stakeholder interno o ai clienti esterni. Trasformare la data science in processi aziendali è il modo in cui le aziende costruiscono il vantaggio competitivo attraverso i dati.

Gli sviluppatori entrano in gioco generalmente alla fine del flusso di lavoro di data science. Sono responsabili della creazione delle applicazioni in cui risiederanno i modelli: applicazioni che sfruttano le informazioni e i dati raccolti dal resto del team di data science. Ciò richiede una certa quantità di tempo di programmazione e può essere un'attività alquanto complessa. La mancanza di tecnologie integrate può tarpare l'avanzamento degli sviluppatori e rendere difficile l'integrazione delle analisi del team di data science.



Gli sviluppatori rendono la data science fruibile, permettendo agli stakeholder interni ed esterni di sfruttare i dati e l'analytics nelle attività di tutti i giorni.

Un altro requisito prevede che gli sviluppatori abbiano a disposizione una serie completa di capacità di programmazione. Sono incluse l'esperienza nella creazione di servizi web, le capacità di sviluppo front-end e una solida conoscenza delle funzionalità e delle caratteristiche dell'interfaccia utente. Inoltre, gli sviluppatori devono avere familiarità con le interfacce di programmazione delle applicazioni (API) e devono saperle usare per integrare i vari prodotti e le fonti di dati nelle applicazioni. Quando funzionano al meglio, le capacità degli sviluppatori sono in grado di creare un'applicazione così solida da non lasciare nemmeno percepire la presenza di un prodotto informatico sottostante.

Per creare applicazioni mobili, web ed enterprise capaci di cambiare le regole del gioco e di disturbare i mercati, agli sviluppatori servono i dati e gli strumenti per trasformare la loro visione in realtà. Lo sviluppatore deve collaborare con il data scientist, l'ingegnere informatico e l'analista aziendale per garantire l'allineamento tra gli obiettivi di business e il back-end analytics.

L'analista aziendale

L'analista aziendale, a volte definito "citizen analyst", assicura competenze e un orientamento di business ai team di data science. Responsabilità dell'analista aziendale è applicare l'esperienza di settore e prendere decisioni informate grazie ai dati.

L'analista aziendale comprende ciò che serve all'azienda ma non dispone del background tecnico per sviluppare un'analisi dettagliata. I miglioramenti

in ambito tecnologico e una maggiore comprensione del profilo utente hanno trasformato gli analisti aziendali in "consumatori analitici", capaci di prendere decisioni basate su informazioni attuabili. Tecnologie flessibili e intuitive consentono loro di sviluppare alcune analisi di livello business senza codice e senza dover chiedere l'intervento del data scientist. Questo assicura agli analisti aziendali una maggiore indipendenza e incrementa la possibilità che utilizzino i risultati del team di data science nelle loro analisi quotidiane.



I team di data science: i nuovi agenti del cambiamento



Data scientist, ingegneri informatici, sviluppatori e analisti aziendali qualificati sono le figure chiave per la trasformazione del business moderno. Sono il cuore pulsante dell'economia dei Big Data. Non si tratta soltanto del fatto che progettano i nuovi sistemi, ma si impegnano a supportare nuove fonti di dati e nuovi modi di utilizzare quei dati. Naturalmente, il reparto IT deve comunque creare il sistema, ma sono i professionisti della data science a consentire ai reparti di collaborare per risolvere i problemi e accelerare l'innovazione.

I migliori prodotti informatici sono quelli che l'utente finale non nota nemmeno. Ciò richiede un'attenta riflessione sull'interfaccia, sui dati che vengono acquisiti e sull'automazione abilitata dal back-end analytics. La tecnologia necessaria a raccogliere e analizzare enormi volumi di dati aziendali è oggi disponibile e può essere sfruttata a vantaggio della tua azienda. Sei pronto a far parte del team nel gioco della data science dove la posta è più alta?

La data science è uno sport di squadra. Hai le capacità per entrare a far parte del team?

Formazione dei professionisti della data science del futuro

Per mantenerti preparato ed equipaggiato per essere un valido membro della squadra, devi migliorarti attraverso la conoscenza e il potenziamento delle tue capacità. IBM® è una delle tante aziende che consentono ai professionisti della data science di migliorare e di aumentare l'efficienza nel proprio lavoro.

In qualsiasi campo in rapido sviluppo servono professionisti con nuove capacità e competenze. IBM ha garantito un supporto attivo

alla community di data science e ha in programma di continuare a garantire il proprio supporto nei forum di formazione online e in presenza. Attraverso eventi, incontri, corsi, contenuti, contributi alla community open source e molto altro, IBM supporta i professionisti informatici di domani e di oggi e li aiuta a prepararsi per soddisfare la domanda di alti volumi e di alta velocità dei dati del futuro.



Risorse

Pronto a potenziare le tue capacità di data science? Qui sono disponibili alcune risorse per cominciare:

1. Costruisci le tue capacità di data science con [Big Data University](#)
2. Inizia a utilizzare gli strumenti che ti servono con [IBM Data Science Experience](#)



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

IBM, il logo IBM, ibm.com e SPSS sono marchi della International Business Machines Corp., registrata 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" all'indirizzo www.ibm.com/legal/copytrade.shtml.

Il presente documento è aggiornato alla data di pubblicazione iniziale e può essere modificato da IBM in qualsiasi momento. Non tutti i prodotti e servizi sono disponibili in tutti i paesi in cui IBM opera.

Gli esempi di clienti citati sono forniti ai soli fini illustrativi. I risultati effettivi delle performance possono variare in base a configurazioni e situazioni operative specifiche. È responsabilità dell'utente valutare e verificare il funzionamento di qualsiasi altro prodotto o programma con prodotti e programmi IBM. LE INFORMAZIONI CONTENUTE NEL PRESENTE DOCUMENTO SONO FORNITE "AS IS", SENZA GARANZIE, ESPLICITE O IMPLICITE, DI QUALSIASI TIPO, IVI INCLUSE LA GARANZIA IMPLICITA DI COMMERCIALIZZABILITÀ O DI IDONEITÀ A UNO SCOPO PARTICOLARE E LA GARANZIA O CONDIZIONE DI NON VIOLAZIONE. I prodotti IBM sono garantiti conformemente ai termini e alle condizioni dei contratti in virtù dei quali vengono forniti.

© Copyright IBM Corporation 2017



Si prega di riciclare