Thomas H. Davenport
Jeanne G. Harris

COMPETERE con gli ANALYTICS

La nuova scienza per vincere nel business

Business 4.0
FrancoAngeli

Business 4.0

Sfide e Opportunità per il Business nell'era dell'Industry 4.0

Coordinamento editoriale: **Alessandro Giaume** (BIP)

Come impattano la trasformazione digitale e l'Industry 4.0 sul business? Quali prospettive si aprono per gli operatori dei diversi settori, in particolare in Italia? Quali ostacoli vanno superati? Quali opportunità possono essere colte? E quali temi si possono definire "chiave" per comprendere meglio tutto questo?

Una Collana per definire le tematiche emergenti, capire i cambiamenti in atto e rispondere con proposte, casi e suggerimenti alle evoluzioni che ne derivano. Testi di autori italiani, pensati per il nostro mercato, ma anche traduzioni internazionali, per osservare e comprendere le novità che si dispiegano per il business a livello globale. Destinatari principali i manager e i consulenti che, dentro e fuori le imprese, devono confrontarsi con professionalità forti dal punto di vista tecnico: per condividere il loro linguaggio e muoversi senza imbarazzo su territori nuovi. Una Collana trasversale sui settori (HR, Operations...) e ampia per temi – IoT e business, Robotica e servizi, Artificial Intelligence e Deep/Machine Learning solo per citarne alcuni – che dia ai propri lettori strumenti efficaci per affrontare una nuova era del business.

I lettori che desiderano informarsi sui libri e le riviste da noi pubblicati
possono consultare il nostro sito Internet: *www.francoangeli.it* e iscriversi nella home page
al servizio “Informatemi” per ricevere via e.mail le segnalazioni delle novità
o scrivere, inviando il loro indirizzo, a “FrancoAngeli, viale Monza 106, 20127 Milano”.

Thomas H. Davenport
Jeanne G. Harris

COMPETERE
con gli
ANALYTICS

La nuova scienza per vincere nel business

Edizione italiana
a cura di Alessandro Giaume

Business 4.0

FrancoAngeli

Titolo originale: *Competing on Analytics. The New Science of Winning*
Updated with a New Introduction

Original work copyright © 2017 Harvard Business School Publishing Corporation
Published by arrangement with Harvard Business Review Press

Traduzione dall'inglese di Pierluigi Micalizzi

Progetto grafico della copertina: Elena Pellegrini

1a edizione. Copyright © 2019 by FrancoAngeli s.r.l., Milano, Italy

L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore. L'Utente nel momento in cui effettua il download dell'opera accetta tutte le condizioni della licenza d'uso dell'opera previste e comunicate sul sito www.francoangeli.it.

Indice

Prefazione, di *David Abney* pag. 7

Introduzione. Quattro ere in dieci anni
La rivoluzione degli analytics » 11

Parte 1 – La natura della competizione analitica

1 **La natura della competizione analitica**
Usare gli analytics per costruire una capacità distintiva » 29

2 **Che cosa fanno i competitor analitici?**
La definizione delle qualità comuni fondamentali di queste aziende » 54

3 **Gli analytics e la performance aziendale**
Trasformare la capacità di competere con gli analytics in un vantaggio competitivo duraturo » 76

4 **Competere con gli analytics nei processi interni**
Applicazioni per amministrazione e controllo, fusioni e acquisizioni, operations, ricerca e sviluppo, e risorse umane » 98

5 **Competere con gli analytics nei processi esterni**
Applicazioni per i clienti e i fornitori » 137

Parte 2 – Costruire una capacità analitica

6 Un piano d'azione per conseguire capacità analitiche avanzate
Un percorso attraverso le cinque fasi della maturità analitica pag. 165

7 Analitici e integrati
Come gestire una risorsa limitata indispensabile per il funzionamento degli analytics » 194

8 L'architettura degli analytics e dei big data
Creare un solido ambiente tecnologico conforme alle strategie aziendali » 223

9 Il futuro della competizione analitica
Approcci determinati da tecnologia, fattori umani e dalla strategia aziendale » 251

Ringraziamenti » 267

Gli autori » 270

Prefazione

Sono convinto che il compiacimento rappresenti una delle minacce più serie per qualunque impresa. Le aziende così assorte nel mantenere la rotta da non accorgersi di ciò che accade alle loro spalle pagano un costo elevato per la loro malriposta soddisfazione. Le aziende che producevano pellicole fotografiche non hanno compreso appieno la minaccia rappresentata dalla fotografia digitale e forse ancor meno quella proveniente dagli smartphone. I quotidiani che dipendono dalle inserzioni pubblicitarie non hanno reagito in maniera proattiva ai siti web come Craigslist. Le aziende che noleggiavano film non hanno reagito abbastanza prontamente alle aziende di streaming come Netflix.

La storia è quasi sempre la stessa. I vincitori diventano gli sconfitti e infine casi di studio perché era senz'altro più facile continuare a fare ciò che li rendeva vincenti. La strada della minore resistenza era più comoda rispetto al mettere in discussione ciò che garantiva il successo.

Il fondatore di UPS, Jim Casey, era consapevole dei pericoli del compiacimento e si appellava a quella che ha definito "insoddisfazione costruttiva". Casey ha reinventato e ricostruito UPS tutte le volte che è servito per contrastare una serie di minacce competitive. Con il tempo, uno dei nostri maggiori vantaggi competitivi lo abbiamo ottenuto adottando fin dall'inizio gli analytics come strumento di valutazione continua e miglioramento costante di ogni aspetto del business.

Gli analytics ci hanno consentito di progettare uno dei primi palmari, che abbiamo messo in mano ai nostri autisti nel 1990. Ma per sfruttare appieno la potenza di questi dispositivi abbiamo dovuto collegare tra loro quelle che erano in gran parte piccole reti locali per formare un network proprietario in grado di coprire tutto il territorio degli Stati Uniti. Questo ci ha messi in grado di raccogliere informazioni sulla movimentazione di ogni collo all'interno del nostro sistema, che a sua volta ci ha consentito di

creare una delle più grandi data warehouse al mondo. Le indicazioni che abbiamo ottenuto tramite gli analytics ci hanno permesso di raggiungere nuovi livelli di efficienza e di condividere queste conoscenze con i clienti.

Dopo un assaggio di ciò che era possibile, la nostra fame di analytics è aumentata. Abbiamo creato una suite di applicazioni basata su modelli predittivi denominata Package Flow Technologies. Lanciate nel 2003, queste applicazioni hanno ridotto i percorsi di 85 milioni di miglia all'anno. E aprendo i nostri motori predittivi all'uso dei clienti e fornendogli le stesse informazioni in nostro possesso, essi potevano interagire con le nostre operation e personalizzare la loro esperienza di consegna.

Ma per la nostra azienda il jackpot analitico è un modello di analytics prescrittivo che abbiamo battezzato ORION (abbreviazione di On-Road Integrated Optimization and Navigation), che indica ai nostri autisti il percorso che dovrebbero seguire per effettuare le consegne nel minor tempo possibile, e con il minor consumo di carburante. Non è una esagerazione sostenere che ORION ha rivoluzionato il nostro business. In questo libro, Tom Davenport e Jeanne Harris hanno definito ORION "l'antenato" delle applicazioni che operano in tempo reale e affermano che si tratta "forse del più imponente progetto analitico commerciale mai sviluppato".

Prima che la rivoluzione digitale prendesse piede, i nostri manager indicavano i percorsi infilando puntine da disegno sulle mappe nei nostri depositi. Poi abbiamo trasferito queste informazioni sui palmari forniti ai nostri autisti. Gli autisti dovevano seguire tutti i giorni più o meno lo stesso percorso, effettuando prima le consegne commerciali e poi quelle private perché normalmente i clienti commerciali erano più esigenti.

Ma con l'avvento della rivoluzione dell'e-commerce i clienti privati hanno preteso un maggiore controllo sulle spedizioni che li riguardavano. Nel tentativo di soddisfare le esigenze degli uni e degli altri, gli autisti si trovano a ripercorrere più volte la stessa strada, sprecando tempo e carburante.

I nostri autisti svolgevano bene il loro lavoro e i clienti ne erano soddisfatti sia in termini di affidabilità sia di credibilità. Ma abbiamo immaginato il giorno in cui non avrebbero potuto spremere una goccia in più di efficienza dai loro percorsi. Quando il *loro* meglio sarebbe stato il *nostro* meglio. Che fare, dunque? (Dovete rammentare che a quel tempo Watson di IBM non era ancora diventato un campione di *Jeopardy* e non eravamo ancora riusciti a trasferire la potenza di calcolo dei primi mainframe nelle nostre tasche). Inoltre, in azienda alcune persone ritenevano che nessun computer avrebbe potuto migliorare ciò che i nostri autisti esperti erano già in grado di fare così bene.

Per fortuna, i nostri ingegneri hanno creato un algoritmo di trenta pagine che ci consente di stabilire il percorso ottimale per ognuno dei nostri cinquantaduemila autisti che effettuano consegne quotidianamente. Oggi,

ORION accorcia mediamente il percorso giornaliero di ogni autista di sette-otto miglia, che complessivamente ci fa risparmiare 100 milioni di miglia e 10 milioni di galloni di carburante (circa 38 milioni di litri). A conti fatti, ORION genera per UPS oltre 400 milioni di dollari di risparmi e contenimento dei costi.

Nel frattempo, abbiamo compreso che i dati da noi raccolti mentre un collo attraversa il nostro network sono spesso tanto preziosi per il cliente quanto lo è il collo stesso. Servendosi dei dati estratti dalla nostra rete, i clienti stanno migliorando i loro processi aziendali e possono prendere decisioni più informate, il che rende UPS un partner di maggior valore.

Per anni, le nostre campagne pubblicitarie hanno proclamato UPS un'azienda che amava la logistica. L'amiamo ancora, ma abbiamo un debito incalcolabile nei confronti degli analytics. E ora non ci limitiamo a impiegarli per ottimizzare la rete fisica, ma li utilizziamo anche per molti altri aspetti del nostro business.

Questa versione aggiornata e ampliata del lavoro originario di Tom e Jeanne racconta un nuovo capitolo della storia degli analytics, con nuove applicazioni, indicazioni e previsioni. Gli autori illustrano come gli analytics continuino a trasformare la tecnologia da strumento di supporto a vantaggio strategico.

In questo libro, gli autori non solo introducono i leader aziendali nel mondo degli analytics, ma li rendono competitor analitici in perfetta armonia con le loro aziende e con il valore strategico. Il libro *Moneyball* di Michael Levis è forse servito per rendere "cool" gli analytics (la versione cinematografica si avvantaggia del contributo di Brad Pitt), ma ci volevano Tom Davenport e Jeanne Harris per farli diventare un imperativo del business.

Aprile 2017

David Abney
CEO UPS

Introduzione.
Quattro ere in dieci anni

La rivoluzione degli analytics

Il mondo che ha ricavato conoscenze dai dati è rimasto relativamente stabile nei suoi primi trent'anni di esistenza. Ci sono stati alcuni progressi tecnologici, ma la tecnica per creare un istogramma o per effettuare un'analisi della regressione non è cambiata molto. Nel 1977, un analista inseriva il programma di analisi e i dati in un computer attraverso schede perforate; nel 2005, l'analista utilizzava invece una tastiera, ma gli altri dettagli erano rimasti pressoché identici.

Tuttavia, a partire dal nuovo millennio, il ritmo dei cambiamenti ha subito una decisa accelerazione. Se identifichiamo come "Analytics 1.0" il modo in cui si effettuava l'analisi e l'attività di business intelligence prima del 2007, nei dieci anni trascorsi da allora abbiamo visto l'avvento delle versioni 2.0, 3.0 e 4.0: tre mutazioni ragguardevoli nel modo di analizzare i dati aziendali in un solo decennio. Quando lavoravamo alla stesura di *Competing on Analytics* nel 2005 e nel 2006, illustravamo l'era primitiva e prendevamo in considerazione le aziende migliori in questo ambito (ritorneremo sul concetto di "Analytics 1.0" poco oltre). Le aziende che si contendevano il vantaggio competitivo utilizzando l'analisi dei dati adottavano quelle metodologie antiquate per gestire i dati e trasformarli in qualcosa che avesse valore.

Da tutto questo possiamo trarre un insegnamento. Estrarre valore dalle informazioni non è primariamente una questione di quanti siano i dati disponibili o di quali tecnologie si usino per analizzarli, sebbene si tratti di fattori che hanno una loro utilità. Si tratta piuttosto dell'incisività con la quale si sfruttano queste risorse e da quanto le si usa per creare approcci nuovi o migliori per fare business. Le aziende leader in questo ambito, di cui parlavamo nella precedente edizione, non usavano sempre gli strumenti più aggiornati, ma erano molto abili nel creare strategie e modelli di business a partire dalle loro capacità analitiche. Quelle aziende erano gestite da

dirigenti che ritenevano che i fatti fossero la via maestra per giungere alle decisioni e alle azioni. I dati e la loro analisi erano una componente costitutiva della loro cultura.

Ciò detto, quando il mondo degli analytics cambia, le aziende migliori cambiano di conseguenza. Non abbiamo verificato se tutte le imprese che abbiamo menzionato nella prima edizione sono passate a versioni successive alla 1.0, ma sappiamo che molte lo hanno fatto. In questa introduzione forniremo una panoramica delle nuove opportunità di utilizzo dei dati e di rivoluzionare il business che sono emerse nel corso degli ultimi dieci anni. Passeremo brevemente in rassegna anche le ere precedenti, non per tenere una lezione di storia ma per capire quali insegnamenti possiamo trarne.

Analytics 1.0 e le sue implicazioni per il presente

A metà degli anni 2000, quando stavamo scrivendo *Competing on Analytics*, le aziende più progredite avevano ormai acquisito una perfetta padronanza degli Analytics 1.0 e cominciavano a pensare alla fase successiva. Ma ancora oggi, molte aziende sono solidamente ancorate all'ambiente 1.0. E per quanto siano disponibili tecnologie analitiche e processi più avanzati, tutte le aziende devono ancora ricorrere per alcune attività alla versione 1.0. Pertanto, anche se vi siete lasciati alle spalle l'era 1.0, vale la pena comprenderne il significato.

Gli Analytics 1.0 facevano (o fanno, se li usate ancora oggi) ampio ricorso all'*analisi descrittiva* – report e illustrazioni che spiegavano ciò che accadeva nel passato – e in misura minore usavano strumenti di analisi per predire il futuro (*analisi predittiva*) o che fornivano indicazioni su come svolgere in modo più efficace un determinato compito (*analisi prescrittiva*). Pur avendo impiegato gli ultimi dieci anni sollecitando le aziende a procedere oltre l'analisi descrittiva, questa è ancora necessaria: serve per sapere che cosa è accaduto nella vostra azienda nel recente passato e come confrontarlo con ciò che è accaduto in precedenza.

Le aziende avanzate dei nostri giorni generano ancora analisi descrittive, ma cercano di limitarne il volume e cercano di servirsi degli utenti (invece che di analisti professionisti) per la loro creazione. È disponibile una serie di nuovi strumenti per semplificare "l'analisi fai da te", in particolare per la creazione di materiale visivo. Naturalmente, molte analisi condotte dagli utenti impiegavano come strumento d'elezione i fogli di calcolo, ed è ciò che ancora accade, nonostante la questione dei possibili errori e la probabilità che con questi strumenti si creino "molteplici versioni della verità".

I dati costituiscono un problema che è rimasto immutato in tutte le ere: ottenerli, depurarli, archiviarli nei database per l'uso futuro e via dicendo. Vista la proliferazione dei dati negli scorsi decenni, si doveva trovare una soluzione nell'ambito degli Analytics 1.0. La soluzione principale per l'immagazzinamento dei dati di quel periodo è stata l'invenzione dei database relazionali: un passo avanti significativo rispetto agli approcci precedenti, ma la nuova metodologia poneva anche sfide ragguardevoli. Immagazzinare i dati attraverso un processo chiamato ETL (Extract, Transform and Load) richiedeva molto tempo e risorse. Tutti i dati dovevano essere strutturati nello stesso modo (righe e colonne) prima di essere archiviati. Alla fine, i database erano talmente grandi e diffusi che era difficile sapere quali risorse contenessero. E benché lo scopo del database fosse quello di mantenere i dati da analizzare separati dai sistemi per le transazioni, gli analytics sono diventati così importanti che una parte dei dati immagazzinati è stata impiegata per applicazioni destinate alla produzione.

I problemi degli Analytics 1.0 non erano solo di natura tecnologica. La cultura dell'epoca era reattiva e lenta. Un'analista cresciuta in quel periodo ha dichiarato che il suo ruolo era quello di "prendere ordini". I manager richiedevano l'analisi di un problema di cui si stavano occupando e l'analista si ripresentava – spesso dopo un mese di arrotondamenti e aver fatto una qualche analisi dei dati – con una risposta. Probabilmente il manager non comprendeva il metodo analitico usato e magari non usava i risultati per prendere una decisione, ma almeno faceva la figura del dirigente orientato ai dati.

In questa prima era, gli analytics erano spesso considerati un *supporto alle decisioni*. E il termine *supporto* è convenientemente debole. Questi dati venivano usati unicamente per supportare le decisioni interne e spesso venivano ignorati. Generalmente, i manager non avevano un rapporto stretto con gli analisti che perlopiù operavano dietro le quinte. Di conseguenza, molte delle decisioni si prendevano ancora basandosi sull'intuito e sulle sensazioni.

Nonostante queste difficoltà, le aziende che abbiamo preso in esame e che si affrontavano sul mercato *usando effettivamente* gli analytics a metà degli anni 2000 cercavano di ottenere il massimo da una situazione difficile. Cercavano di immaginarsi in quali ambiti gli analytics potessero aiutarli a prendere decisioni migliori e producevano questo tipo di dati in gran quantità. Si trattava probabilmente di un processo più lungo e faticoso del dovuto, ma il loro impegno su questo fronte era tale da garantire risultati di qualche utilità. I loro sforzi sono stati fonte di ispirazione per noi e per molti lettori, e per quanti credevano all'idea della "concorrenza tramite gli analytics". Ma nella Silicon Valley, il mondo stava già cambiando.

Analytics 2.0: l'alba dei big data nella Silicon Valley

Più o meno dieci anni fa, nella Silicon Valley, i principali protagonisti del mondo online (Google, eBay, PayPal, LinkedIn, Yahoo! e via elencando) si stavano lasciando alle spalle gli Analytics 1.0. Queste aziende avevano adottato un nuovo paradigma per i dati e la loro analisi, fondato sul bisogno di rendere comprensibile il flusso di navigazione degli utenti a cui essi avevano dato origine. Questi erano dati voluminosi, in rapido movimento e mutamento, e non sempre erano ordinati in righe e colonne. In breve, si trattava di *big data*. La nuova era degli Analytics 2.0 riguardava queste aziende pionieristiche. Non consigliamo ad altri tipi di imprese di adottare direttamente questo approccio, ma anche loro possono trarre molti insegnamenti dagli Analytics 2.0.

Per immagazzinare, analizzare e agire su tutti quei dati, le aziende avevano bisogno di alcune nuove tecnologie. Pertanto, nel 2006, Doug Cutting e Mike Cafarella hanno creato Hadoop, un programma open-source per immagazzinare grandi quantità di dati da diversi server. Hadoop non analizza i dati, ma può svolgere una elaborazione di base degli stessi e rappresenta un modo economico e flessibile per immagazzinare i big data.

Hadoop è diventato il capostipite di una serie di tecnologie di elaborazione dei big data dai nomi più bizzarri. Pig, Hive, Python, Spark, R e una varietà di altri strumenti sono diventati il modo preferito (almeno nella Silicon Valley) di immagazzinare e analizzare i big data. Gli analytics prodotti non erano generalmente molto raffinati (un amico data scientist la chiamava la sindrome "dei big data equivalenti alla matematica elementare"), ma la flessibilità e il basso costo delle tecnologie, e l'applicazione degli analytics a forme di dati meno strutturate, erano dei grandi passi avanti. Lo sviluppo e l'acquisizione di queste tecnologie open-source hanno prodotto un cambiamento che è in corso ancora oggi. In molte applicazioni si adotta una combinazione di analytics proprietari e di strumenti di gestione dei dati. Per programmare questi strumenti e svolgere anche l'analisi dei dati, sembra che sia necessario un nuovo profilo professionale. I professionisti dell'analisi dei big data si sono dati il nome di *data scientist*. Come hanno osservato Tom (Davenport) e il suo coautore (chief data scientist della Casa Bianca fino al 2017) nel loro articolo "Data Scientist: The Sexiest Job of the 21st Century" queste persone erano diverse dal comune analista quantitativo[1].

Innanzitutto, non erano contente di rimanere dietro le quinte. Patil continuava a ripetere a Tom che volevano stare "sul ponte di comando", accan-

1. Thomas H. Davenport, D.J. Patil, "Data Scientist: The Sexiest Job of the 21st Century", *Harvard Business Review*, ottobre 2012, https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century.

to al CEO o a qualche altro top manager e dare una mano al timone della nave. Per esempio, lo stesso Patil era stato un data scientist in LinkedIn che è poi passato a lavorare nel settore del venture capital, quindi a una start-up come responsabile del prodotto e da qui alla Casa Bianca (dove ha ammesso di aver avuto un ufficio nel seminterrato, ma almeno nell'edificio giusto).

In secondo luogo, i data scientist che abbiamo intervistato non erano interessati all'attività di supporto alle decisioni. Uno di loro ha definito questa attività di consulenza direzionale "la zona morta". In molti casi, preferivano lavorare sui prodotti e le loro caratteristiche, demo e via dicendo, elementi che riguardano i consumatori. LinkedIn ha sviluppato *data products* come People You May Know, Jobs You May Be Interested In, e Groups You Might Like, offerte che hanno determinato una rapida crescita dell'azienda e hanno portato alla sua acquisizione da parte di Microsoft per 26 miliardi di dollari. Praticamente tutto ciò che Google produce – tranne forse i suoi telefoni e i termostati – sono prodotti realizzati a partire dai dati e dagli analytics. Facebook ha la sua versione di People You May Know e anche prodotti come Trending Topics, News Feed, Timeline, Search e molti approcci diversi inerenti al targeting pubblicitario.

È evidente che gli analytics erano al centro delle strategie di molte di queste aziende. Google, per esempio, si è evoluta a partire dall'algoritmo PageRank. Forse queste aziende hanno cercato il vantaggio competitivo attraverso gli analytics più di quanto non abbiano fatto le aziende di cui abbiamo dato conto nella prima edizione di questo libro. Una tale visione alternativa degli obiettivi e dell'importanza degli analytics costituisce un insegnamento fondamentale offerto dai professionisti degli Analytics 2.0.

Ma gli Analytics 2.0 rappresentavano anche una cultura dell'esperimento molto impaziente di emergere. Abbiamo potuto rilevare che la maggioranza dei data scientist aveva un PhD in fisica sperimentale. Facebook, un importante referente per questa nuova professione, si riferiva ai data scientist e agli sviluppatori come "hacker" e il suo motto era "fate in fretta e fate piazza pulita". Si tratta di una componente interessante della cultura della Silicon Valley, che probabilmente non sarebbe però adatta a molte delle grandi aziende.

Analytics 3.0: big (e small) data

Più o meno dopo il 2010, molte aziende hanno compreso chiaramente che la questione dei big data non era una moda e che c'erano diverse tecnologie importanti e molte cose da imparare da questo movimento. Tuttavia, visto il disallineamento tra grandi aziende consolidate e Analytics 2.0, a questo punto si è reso necessario un nuovo modo di intendere gli analytics.

Per molti aspetti, gli Analytics 3.0 sono una combinazione delle due versioni precedenti: si tratta di "big data per grandi aziende", ma anche in questa era gli small data sono ancora importanti. Le aziende sono interessate ad analizzare flussi di navigazione, opinioni e sentimenti dei social, dati provenienti dai sensori come nell'IoT e informazioni relative alla localizzazione dei clienti – tutti dati "big" –, ma sono anche interessate a combinarli con gli small data, come gli acquisti effettuati dai clienti in precedenza. Non è questione di big o small data, si tratta di *tutti* i dati.

Nel mondo 3.0, gli analytics non sono più soli: sono stati integrati con i processi produttivi e i sistemi – hanno dato luogo a quello che il nostro amico Bill Franks, chief analytics officer presso l'Institute for Analytics, chiama "analytics operativi". Ciò significa che i dati di marketing non si limitano a informare una nuova campagna di marketing, ma sono integrati nelle offerte in tempo reale disponibili sul web. L'ottimizzazione della supply chain non avviene attraverso una sessione autonoma di analytics, ma è invece incorporata in un sistema di management della supply chain, così che nel magazzino vi sia sempre il numero corretto di prodotti.

Attraverso gli analytics, le aziende dell'era 3.0 intendono conseguire una combinazione di obiettivi: dare forma alle decisioni e definire nuovi prodotti e servizi. Esse sono sempre interessate a usare i dati per influire sulle decisioni, ma vogliono farlo su una scala e con una portata più vasta. Non c'è esempio più efficace dell'imponente progetto ORION di UPS, che è anche un eccellente esempio di analytics operativi. Il progetto, che ha richiesto quasi dieci anni di tempo per essere sviluppato e realizzato da UPS, è un'applicazione analitica per la pianificazione e l'ottimizzazione dei percorsi stradali. Invece di seguire ogni giorno lo stesso percorso, ORION stabilisce il percorso in base agli indirizzi delle consegne e dei ritiri. Attualmente, il programma è in grado di fornire un percorso diverso ogni giorno e in futuro potrà modificarlo in tempo reale in base alle condizioni meteorologiche, a una chiamata per un ritiro o al traffico.

L'investimento e i guadagni relativi a ORION sono entrambi impressionanti: UPS spende diverse centinaia di milioni di dollari e produce introiti annuali ancora maggiori. UPS ha calcolato (con il consueto approccio analitico) che ORION farà risparmiare all'azienda mezzo miliardo di dollari all'anno di manodopera e carburante. Questa è la scala su cui possono operare gli analytics nell'era 3.0.

Le decisioni sono importanti per le aziende che sono entrate nel mondo degli Analytics 3.0, ma queste aziende hanno capito che i data e gli analytics non contribuiscono solo al processo decisionale ma servono anche per sviluppare prodotti e servizi. Anche grandi aziende come General Electric (GE), Monsanto e United Healthcare sono in grado di offrire i data products che le start-up offrivano nell'era degli Analytics 2.0. GE ha

un nuovo modello di business "industriale digitale" alimentato dai *sensor data* dei motori a turbina, turbine a gas, pale eoliche e apparecchiature per la risonanza magnetica. Si usano i dati per creare nuovi modelli di servizi basati sulla previsione dei bisogni, che non sono erogati a intervalli regolari. Monsanto ha un prodotto digitale (Climate Pro) di "semina prescrittiva" che usa dati relativi al clima, alla coltura e al terreno per indicare all'agricoltore il periodo ottimale per seminare, irrigare e per il raccolto. United Heathcare ha una business unit denominata Optum che genera un fatturato annuo di 67 miliardi di dollari dalla vendita di dati, analytics e sistemi informativi.

Naturalmente, nell'era degli Analytics 3.0, i dati e gli analytics sono diventati risorse aziendali comuni e hanno acquisito un'importanza cruciale nelle strategie e nei modelli di business di molte aziende. In breve, il concetto di "competizione attraverso gli analytics" è ormai riconosciuto e accettato, il che, ovviamente, non significa che sia facile avere successo grazie alle innovazioni degli strumenti analitici o che le aziende non debbano impegnarsi in un processo di innovazione continua.

Analytics 4.0: l'avvento degli analytics autonomi

Le prime tre ere degli analytics erano accomunate da una caratteristica: questi dati erano generati da analisti e data scientist che, dopo aver raccolto i dati e costruito un'ipotesi, dicevano a un computer come operare. Ma negli ultimi anni, il mondo degli analytics ha vissuto una profonda trasformazione che implica l'abolizione del fattore umano dall'equazione o, più precisamente, la limitazione del suo ruolo.

L'intelligenza artificiale o le tecnologie cognitive sono da molti considerate la forza tecnologica più rivoluzionaria del mondo contemporaneo. È certamente meno noto il fatto che la maggior parte degli strumenti cognitivi faccia riferimento a modelli statistici o analitici. L'area delle tecnologie cognitive comprende una varietà di tecnologie diverse, ma il *machine learning* o *apprendimento automatico* è una delle più diffuse e si tratta di una tecnologia che è in larga misura di natura statistica. Ma nell'apprendimento automatico, la macchina crea i modelli, stabilisce se sono compatibili o meno con i dati e poi genera ulteriori modelli. Nel caso di alcune forme di machine learning, si potrebbe sostenere che sono i dati a creare il modello, il modello riceve istruzioni da un set di dati ed è in grado di adattarsi a nuove versioni dello stesso.

L'avvento dell'apprendimento automatico è in larga parte dovuto alla rapida crescita dei dati, alla disponibilità del software e alla potenza delle odierne architetture di calcolo. Per esempio, le *reti neurali* – una versione

statistica del machine learning – sono in uso dagli anni Cinquanta e, a partire dagli anni Novanta, sono ampiamente utilizzate dalle applicazioni aziendali. Ma le versioni attuali – in taluni casi si parla di *deep learning* (apprendimento profondo) perché vi sono diversi strati di caratteristiche o variabili che consentono di fare previsioni o prendere decisioni – richiedono grandi quantità di dati da cui apprendere e una notevole potenza di calcolo per risolvere i problemi complessi per cui sono impiegate. Per fortuna, la legge di Moore (secondo la quale la potenza di calcolo raddoppia ogni diciotto mesi) ha fornito la potenza necessaria. Il *data labeling* (l'ordinamento e la categorizzazione di dati non strutturati che sono poi impiegati per insegnare alla macchina) è un'attività che presenta maggiori difficoltà. Ma in molti casi, sono disponibili fonti di dati adatte per questo scopo. Per esempio, il database ImageNet – un database gratuito usato per il riconoscimento delle immagini nelle tecnologie cognitive – contiene 14 milioni di immagini che si possono utilizzare per l'apprendimento della macchina.

In termini di software per l'apprendimento delle macchine, sono disponibili diverse opzioni sia sul versante dei prodotti proprietari sia di quelli open-source. Microsoft, Facebook e Yahoo! hanno tutte messo a disposizione librerie open-source per il machine learning. Start-up come DataRobot e Loop AI Labs propongono le loro versioni proprietarie, mentre le maggiori aziende del settore IT stanno integrando la loro offerta di prodotti con applicazioni per il machine learning. Le tecnologie cognitive sono disponibili come software separati o, sempre più spesso, come applicazioni integrate in altri tipi di software. SAS mette a disposizione metodi per l'apprendimento automatico per aumentare le potenzialità del suo software analitico tradizionalmente basato sull'ipotesi. IBM ha scommesso con convinzione su Watson come pacchetto software autonomo oppure come una serie di programmi più piccoli (API) che si connettono ad altri software. Salesforce.com ha recentemente rilasciato Einstein, un set di applicazioni cognitive integrato nei suoi "cloud" per le vendite, il marketing e i servizi. Riteniamo che praticamente tutti i principali produttori di software installeranno applicazioni cognitive nei loro sistemi per le transazioni commerciali.

Per quanto riguarda l'hardware, i computer più importanti operano in remoto. La disponibilità potenzialmente illimitata di capacità di calcolo a prezzi ragionevoli nel cloud significa che i ricercatori e gli sviluppatori di applicazioni possono avere con facilità la potenza che gli serve per macinare i dati con i tool cognitivi, senza nemmeno dover comprare un computer. E vi sono tipi di processori relativamente nuovi come le unità di elaborazione grafica (GPU – Graphics Processing Units) che sono particolarmente adatti per affrontare alcuni problemi cognitivi come il

deep learning. Esistono anche infrastrutture di calcolo che collegano diversi processori in rete che consentono di eseguire uno "stack" completo di complessi algoritmi e tool.

Le aziende leader nel settore dell'elaborazione e analisi dei dati, stanno pertanto rapidamente migrando verso le tecnologie cognitive in generale, e il machine learning in particolare. Questo perché normalmente l'apprendimento automatico diventa l'unica strada percorribile quando è necessario gestire l'enorme mole di dati a loro disposizione e per creare i modelli personalizzati e rapidamente adattabili di cui hanno bisogno.

Queste tecnologie non sono destinate a sostituire gli analisti umani in tempi brevi ma, come minimo, l'apprendimento automatico rappresenta per loro un ausilio molto significativo per la produttività. Queste tecnologie semi-autonome consentono di creare migliaia di modelli nello stesso tempo di cui un analista (umano) aveva bisogno per crearne uno solo. Produrre velocemente molti modelli significa che un'organizzazione può adottare un approccio più granulare nei confronti dei clienti e dei mercati, e che può reagire con prontezza al cambiamento dei dati. I modelli creati mediante il machine learning possono essere molto più accurati di quelli prodotti con metodi *artigianali* (gli analytics formulati e scrupolosamente modellati dagli analisti) perché spesso prendono in considerazione più variabili in diverse combinazioni. Alcuni metodi di apprendimento automatico sono anche in grado di collaudare un "ensemble" di tipi diversi di algoritmi per capire quale spiega meglio il fattore in questione. Il rovescio della medaglia di questo tipo di approcci è che spesso i modelli generati dal machine learning non sono molto trasparenti o facilmente interpretabili dagli utenti.

Se nella vostra azienda esiste già un team che si occupa di analytics e lavora con modelli statistici nelle aree del marketing, della supply chain, delle risorse umane o in qualche altro ambito, come potrebbe effettuare il passaggio al machine learning? I vostri esperti di analytics dovrebbero acquisire alcune nuove competenze. Invece di vagliare, identificare accuratamente variabili e ipotizzare modelli, gli analisti che adottano il machine learning o i data scientist dovrebbero assemblare grandi quantità di dati e monitorare gli output dell'attività di machine learning per verificare pertinenza e attendibilità dei dati generati.

Probabilmente, dovrebbero anche usare qualche nuovo strumento. Come abbiamo visto, molti produttori di software analitici proprietari si sono affrettati ad aggiungere applicazioni di machine learning, ma molti algoritmi efficaci sono disponibili in formato open-source, per quanto possano offrire un supporto minore agli utenti. Questi esperti potrebbero dover operare anche con nuovi componenti hardware. Dato che i modelli di apprendimento automatico operano solitamente con grandi quantità di dati e