

Il Data Scientist: la professione del futuro

La società attuale sta attraversando profondi cambiamenti. La grande crisi finanziaria iniziata nel 2008 ha segnato profondamente l'economia dei paesi occidentali, con forti ripercussioni anche su quella dei paesi emergenti. La società, intesa nella sua globalità, si sta rendendo sempre più conto da un lato di non poter reggere per molto il paradigma dello sviluppo fine a se stesso, dall'altro di non riuscire a trovare un nuovo modello di crescita sostenibile.

La crisi che ha colpito gran parte della popolazione mondiale costringe, infatti, gli acquirenti a modificare le strategie di spesa praticando un modello di consumo selettivo e intermittente. Inoltre, grazie alla diffusione dei nuovi media quali *social network*, *blog*, *web community*, i consumatori sono ora molto più informati e più competenti rispetto al passato, credono meno nei messaggi promozionali e preferiscono scambiarsi tra loro idee e giudizi su prodotti e servizi.

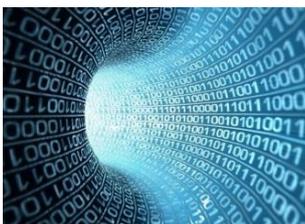


Questo porta le imprese commerciali a orientarsi verso un nuovo modello di business focalizzando l'attenzione sui bisogni "reali" dell'acquirente, cercando di assecondare le sue scelte in tema etico e ambientale, allo scopo di fornire un miglior servizio, riconoscendone il valore e sviluppando così la relazione a lungo termine.

In vista di quest'obiettivo le imprese devono compiere un grande sforzo per strutturarsi diversamente rispetto al passato sul versante della tecnologia informativa, attraverso l'acquisizione e il corretto utilizzo d'informazioni strategiche volte a determinare le differenti propensioni, esigenze e aspettative dei potenziali destinatari delle nuove proposte commerciali.

Il recente sviluppo della tecnologia dell'informazione è orientato a questa necessità ed è volto all'utilizzo di *database* sempre più capaci ed efficienti per la registrazione di enormi quantità di dati e di potenti strumenti d'interrogazione per la loro elaborazione al fine di reperire velocemente le informazioni richieste.

Per questo le aziende che investiranno risorse per governare efficacemente tali fonti d'informazione, sia dal punto di vista della gestione, sia dal punto di vista analitico, godranno nei prossimi anni di un notevole vantaggio competitivo.



Tale tecnologia prende il nome di **Big Data**: enormi database d'informazioni che vengono raccolte e analizzate per fornire risposte ai problemi della vita reale. Queste informazioni, a differenza del passato, non sono necessariamente "strutturate" (come i tradizionali *dataset* costituiti da tabelle di dati formate da righe e colonne), ma sono spesso composte di dati "semi-strutturati" che possono comprendere anche documenti di vario genere come file di testi, *email*, *form* compilati nei portali web, commenti e messaggi inviati su siti specializzati o sui social network.

Per gestire queste grandi quantità d'informazioni è richiesta una forza lavoro molto esperta, i **Data Scientist** (scienziati dei dati), ovvero professionisti capaci di trasformare enormi masse di dati in informazioni comprensibili e fruibili per l'adozione di decisioni importanti per le strategie aziendali.



La figura professionale del Data Scientist richiede una particolare combinazione di competenze che coinvolgono l'informatica, la statistica e la capacità d'interpretare i risultati ottenuti, come afferma il Prof. Dino Pedreschi, docente di informatica all'Università di Pisa, in un'intervista del 2014 al quotidiano [Il Sole 24 Ore](#):

“Il Data Scientist è una figura che deve avere più competenze. La prima è sapere gestire, acquisire, organizzare ed elaborare dati. La seconda competenza è di tipo statistico, ovvero il sapere come e quali dati estrarre, la terza capacità è una forma di storytelling, il sapere comunicare a tutti, con diverse forme di rappresentazione, cosa suggeriscono i dati.”

La conoscenza informatica è fondamentale per l'attività di preparazione e consolidamento dei dati ed è destinata a occupare circa l'80% del tempo dell'analista, il quale sa bene che nessuna decisione può essere presa correttamente in presenza di dati incompleti o imprecisi. In questa fase sono importanti la padronanza di linguaggi di programmazione per l'interrogazione dei database e la conoscenza dei dati delle tabelle in essi contenuti.



Sui dati resi fruibili per l'analisi vengono utilizzati modelli statistico-matematici miranti a ottenere le informazioni richieste. Uno dei modelli più utilizzati è quello che consente di classificare i clienti in funzione del loro comportamento d'acquisto, in modo da migliorare la relazione con i clienti rivolgendo a loro offerte mirate.

La presentazione dell'analisi è l'ultima fase, ma non meno importante, del lavoro dell'analista: l'aspetto visuale, sotto forma d'immagini o grafici, consente una migliore comprensione dei risultati e dei benefici che si possono ottenere dall'impiego delle strategie proposte.



Questo lavoro, come si può capire, è una miscela di scienza e arte, nella quale è richiesta una grande esperienza paragonabile, se si vuole, a quella di un grande chef che prepara le proprie specialità raccogliendo con cura gli ingredienti, pulendoli, tagliandoli, cuocendoli e, alla fine, propone il piatto accompagnato da una presentazione accurata per renderlo il più attraente possibile.

Non fa meraviglia che questa esperienza sia tra le più ricercate negli anni a venire come sostiene Pedreschi:

“Quella del data scientist sarà nei prossimi anni tra le figure più ricercate nel mondo del lavoro e sono sempre di più le università che preparano percorsi post-universitari aperti a tutti i curriculum.”

Tale previsione è in linea con quella degli analisti di settore come dimostra uno studio del 2011 della [McKinsey](#) in cui si stimava che entro il 2018 i soli Stati Uniti potranno affrontare una carenza di 140.000 a 190.000 persone con una grande esperienza in questo settore e di 1,5 milioni di manager e analisti che siano in grado di prendere decisioni in base alle analisi prodotte: **“By 2018, the United States alone faces a shortage of 140,000 to 190,000 people with deep analytical skills as well as 1.5 million managers and analysts to analyze big data and make decisions based on their findings.”**

In un articolo del 2015 della [Commissione Europea](#) si prevede che, nonostante l'alto tasso di disoccupazione, ci sarà entro il 2020 una carenza di oltre 825.000 professionisti del settore dell'Information e Communication Technology: *“Although millions of Europeans are currently without a job companies have a hard time finding skilled digital technology experts. As a result, there could be up to 825,000 unfilled vacancies for ICT (Information and Communications Technology) professionals by 2020”*.

Da questi studi è facile capire che **si tratta di una professione non circoscritta alle grandi aziende ma riguardante tutte le organizzazioni, dalla pubblica amministrazione alle multinazionali.**

Come riportato da alcuni progetti della [Commissione Europea](#), l'utilizzo di questa tecnologia potrà avere applicazioni anche nella vita quotidiana supportando, ad esempio, la Sanità Pubblica, per salvare vite con diagnosi più mirate, i Trasporti Pubblici, per ridurre gli incidenti e gli ingorghi stradali, coadiuvando chi opera nell'Ambiente, per ridurre il consumo di energia, chi opera in Agricoltura, per un cibo più sano e una miglior produttività.

Nel mondo attuale che sta esplodendo di dati (soprattutto quelli derivanti dai social media), la vera sfida per le grandi organizzazioni, inserite in tale ininterrotto flusso informativo, non consiste solo nella loro capacità di acquisire e memorizzare informazioni, ma di trovare in questo settore – ancora emergente– i professionisti che siano in grado di combinare competenze tecniche e analitiche e, soprattutto, che siano mossi da curiosità e creatività: i Data Scientist.

“L'azione va incontro all'insuccesso anche perché non di rado le conoscenze radunate con fervore di zelo non erano guidate da un filo conduttore. Non conosce chi cerca, bensì colui che sa cercare.” Luigi Einaudi

Riferimenti

McKinsey Global Institute (2011), *Big Data: the next frontier for innovation, competition, and productivity*. Si veda: <http://www.mckinsey.com/business-functions/business-technology/our-insights/big-data-the-next-frontier-for-innovation>

Coffetti E., Pasini P. (2015), *Data scientist focus e trends, Economia e Management*, n. 1/2015. Si veda: http://www.sdabocconi.it/sites/default/files/numbers_coffetti_pasini.pdf

Contatti

studio roccato 
Data Science Training & Consulting

e-mail: alfredo.roccato@fastwebnet.it

www.alfredoroccatto.it