

Università di Bologna

di Furio Camillo



Che farò da grande? Risponde il data mining

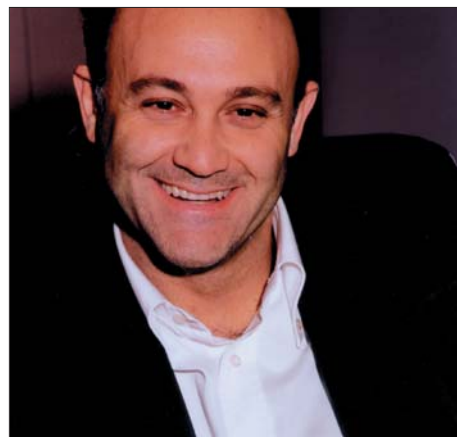
Fidarsi di Eraclito (“Avere una grande cultura non significa essere più intelligenti”) o di Henry James (“La vera scienza insegna a dubitare e ad essere ignoranti”)? E se si sceglie l’università per verificare la fondatezza di questi aforismi, quali criteri utilizzare per scegliere la facoltà più adeguata alle proprie inclinazioni? Ammettiamolo, non è facile. Anche perché la scelta della facoltà rappresenta un momento cruciale nel percorso formativo e lavorativo di un giovane, che comincia a disegnare la propria vita futura. Per aiutare i giovani studenti della scuola secondaria che intendono proseguire gli studi universitari, il Consorzio Interuniversitario AlmaLaurea ha realizzato nei primi mesi del 2006 un innovativo strumento di autoorientamento. **Avvalendosi degli strumenti di elaborazione del Dipartimento di Scienze Statistiche dell’Università di Bologna e mediante un utilizzo intensivo della tecnologia SAS, AlmaLaurea ha cercato di offrire ai diplomandi e ai neodiplomati un originale strumento di autovalutazione che li aiuta a compiere a ragion veduta**

una delle scelte più importanti della loro vita. Lo strumento è reso disponibile per mezzo di un canale agile e adatto a raggiungere un gran numero di giovani: la pagina “Scuola & Giovani” del sito web di Repubblica.it.

Un percorso di autovalutazione

Di che cosa si tratta, nel dettaglio, e come si usa? Il sistema di autovalutazione ha innanzitutto lo scopo di risolvere un quesito: l’università si sceglie con il cuore o con la ragione? E non è una questione da poco, visto che un quinto dei nuovi iscritti abbandona dopo il primo anno. Sono molte le variabili che contribuiscono alla decisione, compiuta a volte senza la giusta consapevolezza o sopravvalutando alcuni elementi (come gli interessi personali) a scapito di altri (come gli sbocchi occupazionali). Il percorso del test aiuta lo studente a verificare come “ha navigato” finora nell’arcipelago della formazione e che cosa occorre migliorare per non incappare in brutte sorprese. Il test si dipana attraverso tre sezioni: la conoscenza delle caratteristiche del sistema universitario e del mercato del lavoro; il percorso scolastico

e i punti di forza individuali; il futuro lavorativo. L’analisi delle risposte è costruita associando ai diversi profili un logo (nella fattispecie un “animaletto”) che indica la lontananza o la vicinanza al livello di soddisfazione/insoddisfazione espresso dai giovani laureati dopo alcuni anni di impegno lavorativo. E i simboli degli animaletti



Furio Camillo, Professore associato di Statistica aziendale, Università di Bologna



Formica ambiziosa



Cane da guardia



Lupo d'appartamento



Ornitorinco



Delfino mediterraneo



Leone rampante



Gatto sornione



Aquilotto alpino



Tartaruga da giardino



Cavallo di Zorro

Sopra sono rappresentati i dieci "animaletti" a cui associare i diversi profili emersi dalle risposte al test

sono l'unico elemento ludico di un test realizzato da un team di psicologi, sociologi, statistici e informatici.

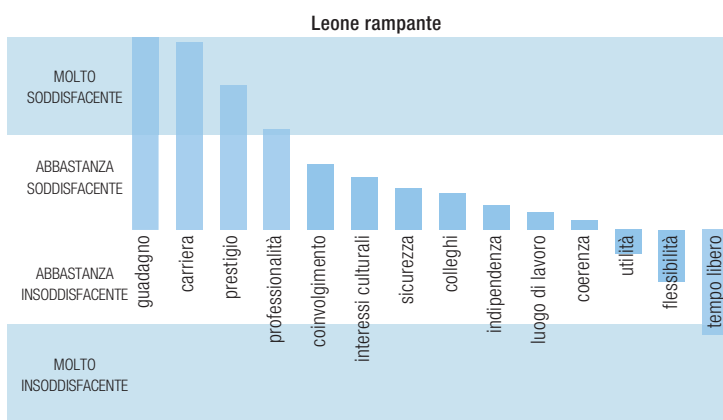
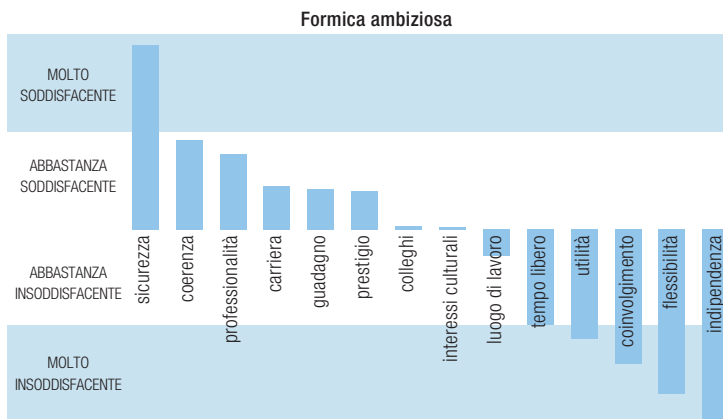
Cluster e "animaletti"

Ovviamente, il test ha un valore orientativo, perché nessuno può dire con assoluta certezza qual è la scelta migliore per un aspirante universitario. Ma avere le idee più precise e poter chiarire quali sono le proprie aspettative è un passo decisivo per una scelta consapevole. Sono oltre 45.000 i giovani che in poco più di quattro mesi si sono sottoposti al test e hanno potuto ottenere informazioni importanti sul proprio profilo individuale. Come? Nel corso degli anni, AlmaLaurea ha svolto un'indagine su circa 80.000 laureati rilevando status, opinioni e valutazioni sulla condizione occupazionale dopo uno, tre e cinque anni dalla laurea.

Con le tecnologie SAS di data mining, i giovani laureati sono stati segmentati in 10 cluster in base ai possibili motivi di soddisfazione-insoddisfazione per il lavoro svolto e a ciascun cluster è stato assegnato il simbolo di "un animaletto". Per mezzo di strategie combinate di analisi discriminante ("Proc discrim" e funzionalità rese disponibili da Enterprise Miner) è stato possibile costruire un modello, la cui simulazione indica in tempo reale al giovane studente la probabilità di appartenere a un dato cluster e di assomigliare all'animaletto corrispondente.

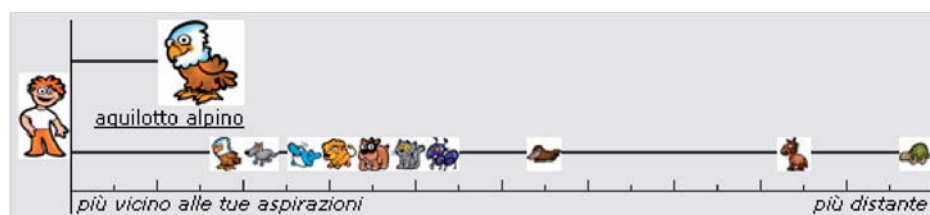
Data mining e domanda formativa

In un secondo tempo, in fase di assessment del modello, vengono simulate le scelte del futuro universitario e viene stimata la probabilità che l'individuo medio si iscriva a una facoltà piuttosto che a un'altra. Peraltro, data l'ampiezza del campione, le simulazioni consentono di valutare a livello nazionale l'orientamento



della domanda formativa universitaria espressa dai giovani che si iscrivono all'università e i risultati sono al vaglio di operatori e di analisti che si occupano a vario titolo del sistema universitario, oltre che dei Ministeri coinvolti (Istruzione e Ricerca-Università). Da una prima analisi dei dati emergono alcune interessanti indicazioni. Ad esempio, si è manifestata una netta differenza tra la distribuzione dei giovani rispondenti e quella relativa ai laureati di 5 anni fa. Come è emerso anche in altre indagini, i giovani sono molto sensibili al tema della sicurezza contrattuale del lavoro, item considerato di gran lunga più importante dai diciannovenni di oggi rispetto ai giovani lavoratori laureati da 5 anni. Del resto, i due animaletti meno frequentemente vicini ai desideri dei ragazzi odierni sono il gatto, animale indipendente e flessibile, e l'aquilotto, animale coinvolto e indipendente. L'attuale generazione sembra quindi penalizzare un'idea di lavoro correlata semanticamente

con l'instabilità: flessibile non è bello e indipendente neanche.



Nel caso esposto in figura il test ha dato origine a una somiglianza con "l'aquilotto alpino" e una forte dissomiglianza con la "tartaruga da giardino". Il test restituiva in tempo reale anche le schede dei due "animaletti" in questione, ossia quello più vicino e quello più lontano dal rispondente.