

Master universitario in Analisi Dati per la Business Intelligence e Data Science

A.A. 2018/2019

Titolo della tesi: Un'applicazione di association-rule mining su prodotti medico-ospedalieri

Autore: Silvio Antonio Pili

Abstract

In questo lavoro abbiamo analizzato dei dati forniti da *TechMed* (nome fittizio), compagnia leader nella vendita di tecnologia medica agli ospedali. Gli obiettivi del lavoro sono due. Il primo consiste nel ricercare co-occorrenze tra i dati forniti dalla compagnia attraverso la derivazione di regole associative, per fornire ai *sales representative* di *TechMed* uno strumento per capire a fondo il comportamento d'acquisto dei propri clienti, gli ospedali, per individuare quali prodotti questi sarebbero più disposti ad acquistare. Per rispondere a questo primo obiettivo abbiamo utilizzato il software *Knime Analytics Platform*, per costruire un *workflow* che individui gli *itemset* frequenti e le regole di associazione nei dati forniti.

Il secondo obiettivo del lavoro, invece, riguarda la costruzione di una *dashboard*, un cruscotto interattivo che consenta sia ai commerciali di *TechMed*, che ai propri analisti, di poter esplorare e navigare in modo rapido ed efficace i propri dati. Con uno strumento di questo tipo, sarà più facile capire, attraverso gli strumenti e le tecniche della *data visualization*, quali *pattern* si presentano tra i dati analizzati. Per questo secondo obiettivo abbiamo fatto ricorso a *Tableau Software*, leader nel campo della *Business Intelligence*. Questo lavoro si compone di due capitoli, suddivisi in diverse sezioni. Nel primo capitolo vengono delineate le principali tendenze contemporanee nel campo del *retail* e del business online come sfide a cui rispondere con gli strumenti del *data mining*. Tratteremo quindi della *Market Basket Analysis*, dell'algoritmo Apriori – il più noto algoritmo

impiegato in questa tecnica – e degli indici utilizzati per interpretare la presenza, la natura e la forza delle regole di associazione trovate a partire dagli *itemset* frequenti.

Nel secondo capitolo ci occuperemo invece del caso studio. Dopo aver descritto il contesto analitico, verrà proposto il disegno metodologico e infine verrà presentata l'analisi vera e propria, facendo attenzione ai diversi punti di applicazione con ciascuno degli strumenti impiegati.