

Master Universitario in: "ANALISI DATI PER LA BUSINESS INTELLIGENCE"

A.A. 2016-2017

Titolo della tesi: Previsione dei fatturati giornalieri per una catena di negozi
Autore: Nicola Giorcelli

Abstract

Scopo del progetto di tirocinio è stato quello di costruire un modello di forecast per una catena di negozi che fosse in grado, date alcune informazioni anagrafiche e temporali, di prevedere i corrispettivi del giorno successivo all'ultimo record per ogni negozio della catena.

La cliente è un'azienda leader nel mercato italiano per la vendita al dettaglio di prodotti per la casa, è molto interessata alle nuove tecnologie, in particolare all'analisi dati. Ha a disposizione un DWH in cui sono archiviate le informazioni dei singoli negozi e sullo scontrinato che generano giornalmente. Attualmente il compito di prevedere le vendite giornaliere di un singolo negozio è affidato al manager di competenza.

In regime metodologico CRISP-DM sono stati esplorati i dati e si è costruito un primo modello che, però, non ha soddisfatto le aspettative del cliente.

Spronato da un paio di professori della Bocconi, consulenti del cliente, e dal tutor aziendale è stata esplorata la teoria delle serie storiche, soprattutto i metodi ARIMA, ed è stato sviluppato un algoritmo in grado di creare previsioni con un orizzonte settimanale. Questo processo però, data la complessità dell'ARIMA e la numerosità dei modelli, si è rivelato molto oneroso.

Si è quindi deciso di fare un passo indietro ed usare i metodi di machine learning per la regressione, mantenendo il flusso di pulizia dati e feature creation usato per le serie storiche.

La fase di preparazione dei dati e il primo algoritmo sono stati sviluppati sfruttando il linguaggio R. Il successivo modello e il suo deploy con i servizi web sono stati creati tramite un tool cloud messo a disposizione gratuitamente: Azure Machine Learning Studio.