

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

A.A. 2013 – 2014

Titolo della tesi: Modelli per la previsione del traffico organico su drezzy.it

Autore: Irene Luppino

## **Abstract**

*Come si misura la prestazione di un sito? Uno degli indicatori chiave, o KPI, che permette di monitorare le performance di drezzy.it è il numero di visite. In particolare, siamo interessati alle visite che avvengono da ricerca organica.*

*I risultati di questo progetto consistono in un modello per la previsione del traffico giornaliero e un modello per la previsione settimanale. Il primo è utile per operazioni di monitoraggio ed individuazione delle anomalie, il secondo serve da supporto nelle decisioni strategiche come l'allocazione del budget.*

*Per questo lavoro sono stati utilizzati gli strumenti R, Excel e Google Analytics. Dopo una prima esplorazione dei dati abbiamo deciso di dividere in due parti le serie storiche delle visite su drezzy.it e di lavorare solo con la seconda porzione di dati. Le osservazioni sono state divise perché a partire da Luglio 2013 si è presentato un cambiamento troppo marcato nel numero di visite e a noi interessa l'evoluzione di drezzy.it da quando ha iniziato a ricevere traffico in maniera significativa.*

*La scelta dei modelli è ricaduta sulle medie mobili semplici (SMA) e sugli ARIMA (AutoRegressive Integrated Moving Average). Per la verifica di adeguatezza dei modelli sono stati usati il test di Ljung-Box sui residui ed i correlogrammi ACF e PACF; per testare la bontà predittiva dei modelli e confrontarli fra loro abbiamo usato l'errore residuo RMSE.*

*L'obiettivo iniziale era quello di trovare dei modelli per la previsione del traffico organico a livello mensile e giornaliero. Lo storico dei dati a nostra disposizione è troppo esiguo e non ha consentito analisi robuste sulle osservazioni mensili. I dati a livello settimanale hanno rappresentato un buon compromesso per ovviare al problema della carenza di dati e hanno consentito di stabilizzare comportamenti anomali come il numero di visite in giorni festivi tipo Natale o Ferragosto.*