Chapter 41
Predictive Analysis for Digital Marketing Using Big Data:
Big Data for Predictive Analysis

Balamurugan Balusamy
VIT University, India

Priya Jha
VIT University, India

Tamizh Arasi
VIT University, India

Malathi Velu
VIT University, India

ABSTRACT

Big data analytics in recent years had developed lightning fast applications that deal with predictive analysis of huge volumes of data in domains of finance, health, weather, travel, marketing and more. Business analysts take their decisions using the statistical analysis of the available data pulled in from social media, user surveys, blogs and internet resources. Customer sentiment has to be taken into account for designing, launching and pricing a product to be inducted into the market and the emotions of the consumers changes and is influenced by several tangible and intangible factors. The possibility of using Big data analytics to present data in a quickly viewable format giving different perspectives of the same data is appreciated in the field of finance and health, where the advent of decision support system is possible in all aspects of their working. Cognitive computing and artificial intelligence are making big data analytical algorithms to think more on their own, leading to come out with Big data agents with their own functionalities.

INTRODUCTION

Due to the substantial scale evolution of structured, semi-structured, unstructured data, it has become a challenge for traditional storage systems and analytics tools. To unlock the great potential of the heterogeneous natured big data, advanced predictive analytics tools are the need of the era (Lohr, 2012) and

DOI: 10.4018/978-1-5225-7501-6.ch041

(Waller, 2013). The data sources are highly scattered and are continuously generating huge data sets through transactions, clickstreams, surveillance, sensors and communication technologies (Tsai et.al, 2015). Predictive analytics for big data is an exponentially growing area of research today and the need of the hour for efficient analysis, management and faster real time utilization of both the stored data as well as real time data that are generated (Lohr, 2012). Predictive analysis is the advanced branch of data engineering that is concerned with predicting the outcome of events based on the analysis of historical data by applying sophisticated techniques of machine learning and regression (Ratner B., 2011) and (Mishra N., & Silakari S., 2012). Prediction is closely related to probability, the futuristic outcomes of certain events can be determined based on the analysis of historical and present data sets (Zaman and Mukhles, 2005). The current trends and patterns in the data sets are explored and after thorough analysis; the results are used to predict the outcomes of the future events.

Due to the vastness of the data sets, several important characteristics of information associated with the data, that could have been leveraged for productive applications often tend to remain unexplored and go to waste (Cao et.al, 2014), to prevent such loss, predictive analytics has a great role to play in the process. Analyzing data, discovering familiar patterns, deriving meaning from the patterns, formulating decisions and ultimately responding to the needs intelligently is the is the ideal goal of predictive analytics for big data. Today, businesses apply predictive analytics (forecasting, hypothesis testing, risk modeling, propensity modeling) in their working environment to increase involvement with their customers, reduce operational costs, and optimize processes involved and other numerous advantages that it offers (Tsai et.al, 2015). Figure 1 depicts the overall organization of the chapter.

The predictor variable is the measurable variable for the entity. In other words, a predictor is a feature selected to be used as an input to a classification model (Jain et.al, 1999). A predictor can take four type of values-continuous, which consists of floating point or numerical magnitude values, categorical, for example Boolean values, word-like or text-like. Combining and integrating several predictor variables can result in a predictive model that can be used to make the required future forecasts within a certain level of reliability.

Predictive analytics has a heavy application in business environments (Banjade & Maharjan, 2011), (Chen et.al, 2012) and (Waller & Fawcett, 2013), as it associations commercial awareness with statistical and analytics techniques to better understand customers, products (Banjade & Maharjan, 2011), and partners and to identify potential opportunities and decrease the chances of risks for the company.

Figure 1. Predictive analysis process overview