Chapter 58
Big Data in Operation Management

Arushi Jain
Ambedkar Institute of Advanced Communication Technology and Research, India

Vishal Bhatnagar
Ambedkar Institute of Advanced Communication Technologies and Research, India

ABSTRACT

The word big data analytics have been increased substantially these days, one of the most prominent reasons is to predict the behavior of the customer purchase. This analysis helps to understand what customer wants to purchase, where they want to go, what they want to eat etc. So that valuable insights can be converted into actions. The knowledge thus gained helps in understanding the needs of every customer individually so that it becomes easier to do the business with them. This is the revolutionary change to build a customer-centric business. To build a customer centric business an organization must be observant about what customer is doing, must keep a record about what customer is purchasing and lastly should discover the insights to maximum the profit for customer. In this chapter we discussed about various approaches to big data management and the use cases where these approaches can be applied successfully.

INTRODUCTION

Big data is the revolutionary world in the field of information technology because of its enormous influence on different domain. Big data is the voluminous and complex collection of data that comes from different sources such as sensors, content posted on social media website, sale purchase transaction etc. Such voluminous data becomes tough to process using ancient processing application. By 2020, International data corporation (IDC) predicts the number will have reached 40 Zettabytes (ZB) that is the world will produce 50 times the amount of information. A huge surge in the amount of data being generated that needs to be stored and analyzed quickly has been witnessed in the recent years. For example walmart handles millions of sale purchase transactions per hour, Facebook handles 40 billion photos uploaded

DOI: 10.4018/978-1-5225-7501-6.ch058

by its users each day. Organizations are using big data to analyze and find insights to build an optimal information system. Big Data can be defined using five V's. These are:

- **Volume:** This refers to the amount of data been generated from different sources such as data logs from twitter, click streams of web pages and mobile apps, sensor-enabled equipment capturing data, etc.
- **Velocity:** This refers to the rate at which data is generated and received. For example for an effective marketing offer to a consumer, ecommerce applications combine mobile GPS location and personal preferences.
- **Variety:** This refers to various types of structured, unstructured and semi- structured data types. Unstructured data consist of files such as audio and video. Unstructured data has many of the requirements similar to that of structured data, such as summarization, audit ability, and privacy.
- **Value:** This refers to the intrinsic value that the data may possess, and must be discovered. There is wide variety of techniques to derive value from data. The advancement in the recent years have led to exponential decrease in the cost of storage and processing of data, thus providing statistical analysis on the entire data possible, unlike the past where random samples were analyzed to draw inferences.
- **Veracity:** This refers to the abnormality in data. Veracity in data analysis is one of the biggest challenges. This is dealt with by properly defining the problem statement before analysis, finding relevant data and using proven techniques for analysis so that the result is trustworthy and useful. There are various tools and techniques in the market for big data analytics. Hadoop is Java-based programming framework that supports processing of large data sets. It was started out as a project by Yahoo to analyze its data but now it is part of the Apache project.

LITERATURE SURVEY

To manage the growing demands, there is a need to increase the capacity and performance of tools and methods employed for analysis of data. Chen et al. (2014), in their work "Big data: A survey" focused on big data and reviewed related technologies and examined the application of big data in various fiels. Al-Jarrah et al. (2015), in their work "Efficient Machine Learning for Big Data: A Review" reviewed the data modeling in large scale data intensive field relating to model efficiency and new algorithm approaches. Hoffmann and Birnbrich (2012) to protect their customer from third party fraud proposed a conceptual link between retail bank activities in "The impact of fraud prevention on bank-customer relationships: An empirical investigation in retail banking". Srivastava and Gopalkrishnan (2015) revealed some of the best techniques which are used by the banks across the globe and can be used by the Indian banks to enhance their services offerings to the customers in "Impact of Big Data Analytics on Banking Sector: Learning for Indian Banks". Azar and Hassanien (2014) for dimensionality reduction presented a linguistic hedges neuro-fuzzy classifier with selected features (LHNFCSF). In this paper author compared the new classifier with the other classifiers for various classification problems in "Dimensionality reduction of medical big data using neural-fuzzy classifier". Hassanien et al. (2015) focused on application, challenges and opportunities of big data in "Big Data in Complex Systems: Challenges