Chapter 26
Big Data and Advance Analytics: Architecture, Techniques, Applications, and Challenges

Surabhi Verma

National Institute of Industrial Engineering, Mumbai, India

ABSTRACT

The insights that firms gain from big data analytics (BDA) in real time is used to direct, automate and optimize the decision making to successfully achieve their organizational goals. Data management (DM) and advance analytics (AA) tools and techniques are some of the key contributors to making BDA possible. This paper aims to investigate the characteristics of BD, processes of data management, AA techniques, applications across sectors and issues that are related to their effective implementation and management within broader context of BDA. A range of recently published literature on the characteristics of BD, DM processes, AA techniques are reviewed to explore their current state, applications, issues and challenges learned from their practice. The finding discusses different characteristics of BD, a framework for BDA using data management processes and AA techniques. It also discusses the opportunities/applications and challenges managers dealing with these technologies face for gaining competitive advantages in businesses. The study findings are intended to assist academicians and managers in effectively quantifying the data available in an organization into BD by understanding its properties, understanding the emerging technologies, applications and issues behind BDA implementation.

INTRODUCTION

In today's highly competitive markets enterprises are differentiated by their ability to make timely, accurate, and effective decisions at all levels – strategic, tactical, and operational – to address different business processes and performance (Tien, 2013). Increasingly, companies in almost every industries around the world are analyzing Big Data (BD) (both structured and unstructured) to create and capture value through real-time decision-making (Amankwah-Amoah, 2015). BD is defined as high volume, high velocity or/and high variety information assets which require new forms of processing to enable

enhanced insight discovery, decision-making, and process optimization (Gartner, 2001; Kshetri, 2014). More specifically, a data set is called BD if it is difficult to capture, store, analyze and visualize using existing technologies in enterprises (Halaweh, 2015). To capitalize advantages from BD, a firm needs Advanced Analytics (AA) techniques and technology to capture, curate, analyze and visualize ever-growing data sets (Gandomi and Haider, 2015). This integration of BD with AA is collectively known as Big Data Analytics (BDA) (Mcneely and Hahm, 2014). BDA is used to describe the data sets that is so large (from terabytes to exabytes) and complex (from social media data to the sensor and mobile data) that they require advanced data storage, management, analytical and visualization techniques and technologies (Chen et al. 2012). It helps in combining historical data with present events and projected future actions (Chen and Zhang, 2014). To uncover the value from BD, it needs to be tapped, analyzed and used for decision making (Kimble and Milolidakis, 2015).

BDA is the subject of attention for business managers, researchers, and government, and to some extent, challenge (Bihl et al., 2016). The exponential rise of BDA in very short period left firms unprepared to handle BD (Bendler et al., 2014). In the past, the rise of new technologies concepts was first discussed in technical and academic publications. The fast evolution of BDA technologies and extremely competitive business environment left little time for the discourse to develop and mature in the academic domain (Bhimani, 2015). There are several articles, industrial reports and books on BDA, including Big Data@ work, but not enough fundamental discourse in academic publications (Kwon et al., 2015). This lagging of academic literature in BDA context implies that a coherent understanding of this emerging technology and its applications in different areas is yet to be developed. For instance, there is little consensus on the fundamental question of what quantify BD (Gu and Zhang, 2015), which type of data and AA techniques is suitable in a particular context (Tan et al., 2015) and how the information extracted after analysis of BD could be used across industries and for public welfare by the government (Vera-Baquero et al., 2015). Thus, there is need to extend the academic literature of BD and AA concepts, techniques, benefits, and challenges.

Many academic studies on BDA have discussed different tools and techniques to analyze BD (Gandomi and Haider, 2015; Goth, 2015), none of them have discussed which type of structured and unstructured data is suitable in a particular situation and how this tremendous data could be leveraged to create value across industries (Vera-Baquero, 2015). Therefore, this paper aims to: investigate the current state of big data, review different process and advance analytics tools and techniques utilized for leveraging big data; provide a concise summary of applications of big data across sectors; and identify challenges which need to overcome before implementing big data. In this process, a wide range of articles is reviewed which discussed various tools and techniques, applications and challenges of big data. In order to understand the process of leveraging big data, a wide range of tools and techniques are reviewed which can manipulate big data. This paper also presents the review of applications of big data differentiated by sectors (i.e., retail, oil and gas, manufacturing and more).

The reminder of this research paper is as follows. This study starts with understanding the current state of BD. In section 3, a holistic framework of BDA is discussed which covers all the steps involved in leveraging the BDA (i.e. from data acquisition to decision making based on information extracted from various analysis). In the same section, different techniques and tools of AA are discussed. In section 4 various challenges faced by an organization to adopt BDA are discussed in detail. The discussion has remained focused on usage of BD at various level and factors that are used to evaluate BDA software suite.

Related Content

EMG-Based Mobile Assessment System for Neck and Shoulder Fatigue
Pei Lun Lai, Hsiu-Sen Chiangand Qi-An Huang (2017). *International Journal of Big Data and Analytics in Healthcare (pp. 39-50).*
www.irma-international.org/article/emg-based-mobile-assessment-system-for-neck-and-shoulder-fatigue/204447

Cardiovascular Risk Detection Through Big Data Analysis
Miguel A. Sánchez-Acevedo, Zaydi Anaí Acosta-Chiand Ma. del Rocío Morales-Salgado (2020). *International Journal of Big Data and Analytics in Healthcare (pp. 1-11).*
www.irma-international.org/article/cardiovascular-risk-detection-through-big-data-analysis/259985

The Strategic Adoption of Big Data in Organizations
Nabeel Al-Qirim, Kamel Rouibah, Mohamad Adel Serhani, Ali Tarhini, Ashraf Khalil, Mahmoud Maqablehand Marton Gergely (2019). *Managerial Perspectives on Intelligent Big Data Analytics (pp. 43-54).*
www.irma-international.org/chapter/the-strategic-adoption-of-big-data-in-organizations/224330

Data Visualization of Big Data for Predictive and Descriptive Analytics for Stroke, COVID-19, and Diabetes
Richard S. Segalland Soichiro Takashashi (2023). *International Journal of Big Data and Analytics in Healthcare (pp. 1-31).*
www.irma-international.org/article/data-visualization-of-big-data-for-predictive-and-descriptive-analytics-for-stroke-covid-19-and-diabetes/331996

Applying Qualitative Matrix Coding Queries and Qualitative Crosstab Matrices for Explorations of Online Survey Data
 (2019). *Online Survey Design and Data Analytics: Emerging Research and Opportunities (pp. 181-204).*
www.irma-international.org/chapter/applying-qualitative-matrix-coding-queries-and-qualitative-crosstab-matrices-for-explorations-of-online-survey-data/227256