Chapter 12
Recent Trends in Big Data: Challenges and Opportunities

Kannadhasan Suriyan
https://orcid.org/0000-0001-6443-9993
Study World College of Engineering, India

Kanagaraj Venusamy
Rajalakshmi Engineering College, India

R. Nagarajan
https://orcid.org/0000-0002-4990-5869
Gnanamani College of Technology, India

ABSTRACT

Using big data and algorithmic techniques, not all of the data gathered in this way is relevant for analysis or decision-making. To be more specific, the chapter addresses issues that arise during the fine tuning of large data sets, presenting open research questions that can aid in the processing of large data sets and the extraction of valuable information from them, as well as providing an overview of big data tools and techniques that can be used to address these issues. Healthcare, public administration, retail, and other multidisciplinary scientific inquiries are only few of the areas where boundaries might be blurred. Big data is mostly derived from social computing, internet text and document storage, and internet search indexing. Online communities, recommendation systems, reputation systems, and prediction markets are all examples of social computing. Internet search indexing includes ISI, IEEE Xplorer, Scopus, Thomson Reuters, etc.

1. INTRODUCTION

Fast digital technology transitions provide the data source, resulting in an explosion of large amounts of data. Traditional database administration systems and data

DOI: 10.4018/979-8-3693-1131-8.ch012

processing programs are unable to handle the enormous datasets generated by these devices. Petabytes and beyond are not uncommon for various types of data. Structured, unstructured, or semi-structured are all possible options. Such data are technically defined by the 3Vs (Volume, Velocity, and Variety). To put it another way: volume refers to an enormous amount of information that is being generated every day; velocity measures how quickly that information can be gathered and processed. There is a lot of variety in the types of data that may be found in a variety. V stands for truthfulness and refers to the fact that it is both accessible and accountable. For big data analysis, the primary goal is to process large amounts of data in a variety of traditional and non-traditional ways. Even though photos and video can be stored and displayed, they are not semantically annotated or searchable because they are not naturally in a structured format like tweets and blogs. converting this information into a logically organized format Any subsequent analysis will be extremely difficult to perform effectively. The next generation of information technology industries will be well-served by the strength of big data. There are a number of IT companies that relate to big data, cloud computing, the internet of things, and social media as a means of solving deliberate problems.

Despite the fact that data warehouses are used to manage datasets, obtaining knowledge data from the available big data is a time-consuming procedure. When dealing with extremely huge datasets, it is impossible to use data mining techniques. This lack of coordination between database systems and analysis tools like data mining and statistical analysis is a major issue in the study of large amounts of data. It's difficult to discover new information in these scenarios. It is imperative that data be presented in a way that can be used in real-world scenarios. In explaining the data revolution, there is a need to consider its epistemological implications (Kitchin, 2014). As a result, the research on complexity theory of big data will help us understand the key characteristics and development of complex patterns in big data, simplify its representation, gain better knowledge abstraction, and lead the design of computer models. There isn't a single field in which big data is associated. It is highly likely that 4G will not be able to meet the increasing demands of manufacturing and industrial automation technology in the age of M2M, IoT, Big Data, and Smart Factories. Based on cyber-physical manufacturing systems, "smart manufacturing" has become a global development trend (CPMS). Automation, collaboration, real-time monitoring, and smart linked control are all hallmarks of Industrial IoT (IIoT), which is one of the significant difficulties resulting from the CPSM trend's development. The use of advanced manufacturing technology has resulted in the creation of a vast volume of data. There is a barrier to CPMS advancement because current mobile technologies, including 3G and 4G, can't deliver on the high dependability, fast data transfer and low latency required by the CPMS standard. Innovation in diverse industrial processes is required to ensure efficiency