

Chapter 1

Introduction to Big Data Analytics

Dina Darwish

Ahram Canadian University, Egypt

ABSTRACT

Big data refers to data collections that are either too huge or too complex for traditional data-processing application software to manage. The three major concepts initially associated with big data are volume, variety, and velocity. The fourth major concept, veracity, is concerned with the accuracy or believability of the data. Big data analytics is the act of acquiring and analyzing massive volumes of data to discover market trends, insights, and patterns that may help firms in making better business decisions. Across all corporate sectors, improving efficiency results in more shrewd operations overall, more profits, and happy customers. This chapter gives an overview on how to store and manage big data, importance of big data analytics, how to apply big data analytics using different methods and tools to benefit businesses, and big data analytics applications in various fields, as well as challenges facing big data analytics.

INTRODUCTION

The creation of digital data is partially a result of the use of online-connected gadgets. Thus, information about their users is transmitted through cellphones, tablets, and desktops. Connected smart gadgets share data on how consumers utilize commonplace items.

Data originates from a variety of sources besides linked devices, including demographic data, climatic data, scientific and medical data, energy usage data, etc. The location of device users, their travels, their hobbies, their consumption patterns, their pastime activities, their projects, and other information are all provided by these data. But also details on how the tools, equipment, and infrastructure are utilized. The amount of digital data is continually expanding as more people utilize the Internet and mobile devices. We currently reside in an informational society that is transitioning to a knowledge-based society. We require more data in order to derive better information. Information is a key component of the political, cultural, and economic spheres in the society of information (Hilbert & Lopez, 2011).

DOI: 10.4018/979-8-3693-0413-6.ch001

The phrase “Big data” refers to the development and application of technologies that deliver the appropriate information from a vast amount of data that has been expanding exponentially in our society to the appropriate user at the appropriate moment. Along with handling the complexity of managing more varied forms and complicated and interrelated data, the problem is dealing with the continuously growing volumes of data. Its definition changes depending on the communities that are interested in it as a user or provider of services because it is a complicated polymorphic object. Big data, a technology developed by the web’s titans, positions itself as a way to provide everyone real-time access to massive databases.

Big Data is highly challenging to define exactly because different fields have different ideas of what is considered to be a large amount of data. It identifies a class of techniques and technologies rather than a specific set of technologies. This is a new field, and the definition is evolving as we try to figure out how to use this new paradigm and capitalize on the benefits. Big data is a broad term for data that can be stored, processed, and computed more efficiently than traditional databases (Riahi & Riahi, 2015). Big Data as a resource necessitates the use of tools and techniques that may be used to examine and draw patterns from vast amounts of data (Najafabadi et al., 2015).

Because of the variety and velocity of the data manipulation, structured data analysis evolves. As a result, it is no longer sufficient to simply analyze data and generate reports; due to the variety of the data, the systems in place also need to be able to support data analysis. In order to aid in its exploitation, analysis entails automatically identifying, among a variety of quickly changing data, the correlations between the data.

The term “Big Data Analytics” refers to the procedure of gathering, compiling, and analyzing sizable data sets in order to identify various patterns and other pertinent data. Big data analytics is a collection of technologies and approaches that call for novel forms of integration in order to reveal significant hidden values from sizable datasets that are more complicated and huge in scale than typical datasets. It generally focuses on finding better and more efficient solutions to both new and old problems.

This chapter discusses the characteristics and ecosystem of Big data, as well as the types and applications of Big data analytics. The main topics to be covered in this chapter includes the following;

- The meaning of the term “Big data” and its characteristics
- Big data ecosystem
- Storage of Big data
- Big data management technologies
- Big data analytics
- Big data analytics life cycle
- Big data Analytics Processing
- Big data analytics and machine learning
- Key studies and trends in machine learning and Big data analytics
- Current research directions, developments and challenges in machine learning and Big data analytics
- Benefits of Big data analytics to businesses
- Types of Big data analytics
- Tools used in Big data analytics
- Applications of Big data analytics
- Ethical implications of Big Data
- Challenges and Barriers for Big data analytics

46 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/introduction-to-big-data-analytics/336344

Related Content

Business Intelligence Should be Centralized

Brian Johnson (2013). *Principles and Applications of Business Intelligence Research* (pp. 139-152).

www.irma-international.org/chapter/business-intelligence-should-centralized/72567

A Comparison of Simultaneous Confidence Intervals to Identify Handwritten Digits

Nicolle Clements (2014). *International Journal of Business Intelligence Research* (pp. 29-40).

www.irma-international.org/article/a-comparison-of-simultaneous-confidence-intervals-to-identify-handwritten-digits/122450

Data Envelopment Analysis and Analytics Software for Optimizing Building Energy Efficiency

Zinovy Radovilsky, Pallavi Taneja and Payal Sahay (2022). *International Journal of Business Analytics* (pp. 1-17).

www.irma-international.org/article/data-envelopment-analysis-and-analytics-software-for-optimizing-building-energy-efficiency/290404

Transformation of Marketing With Technology: Case Approach for Artificial Intelligence

Fatih Pinarbasi and Habib Mehmet Akpınar (2020). *Handbook of Research on Strategic Fit and Design in Business Ecosystems* (pp. 144-165).

www.irma-international.org/chapter/transformation-of-marketing-with-technology/235572

A Modified Kruskal's Algorithm to Improve Genetic Search for Open Vehicle Routing Problem

Joydeep Dutta, Partha Sarathi Barma, Samarjit Kar and Tanmay De (2019). *International Journal of Business Analytics* (pp. 55-76).

www.irma-international.org/article/a-modified-kruskals-algorithm-to-improve-genetic-search-for-open-vehicle-routing-problem/218835