Chapter 1
Data Analytics:
An Overview

Anu Sayal
Jain University (Deemed), India

ABSTRACT

This chapter provides a comprehensive and unified view of data analytics. Data analytics is the process of analyzing the raw data in order to draw inferences about the information in hand. Data analysis techniques are primarily used to get an insight which further facilitates enhancement of the sector under consideration. These techniques are beneficial for optimizing a process under consideration and also for increasing the overall efficiency of a system. These techniques also act as performance boosters as their implementation in the business models help in reduction of costs by considerable amount. It is the most important for any organization as it facilitates better decision-making approaches and also provides an analysis of customer trends as well as satisfaction which further leads to improved products as well as services. It also helps in effective marketing of the products and services. Data analytics has widespread application in various sectors. Various tools are used for carrying out data analytics jobs. All this is discussed in the chapter.

INTRODUCTION: WHAT IS DATA?

Data is basically referred to as individual facts, statistics or any type of information usually in the form of numerical values. If considered in technical sense the data consists of set of values in the form of quantitative or qualitative values which may be about one or more person or object. Data and information are used in an interchangeable manner. Considering some specific cases, the data is converted into

DOI: 10.4018/978-1-6684-5722-1.ch001

information when it is being observed in some context or in any kind of post-analysis whereas in academic sense data are purely the elements of information. Data are used in various fields like finance, scientific research, business management, governance and in every other form of the human organizational activities.

Data are also referred to as the atoms of decision making, they are the smallest units related to factual information that can be utilised discussion, reasoning or calculation purpose. Data can vary from abstract ideas to concrete measurements and even statistics. Data are collected, classified, tabulated and analysed for specific purposes. They are also used for the creation of data visualizations like tables, graphs or even images. Data processing usually takes places in different stages. Data is transformed into some suitable information for making decisions once it has been analysed.

The main objective of this chapter is that it deals with the process of data analytics. The concepts related to the types of data, types of data analytics, data mining process, IoT and digital twins have been explained in a lucid manner. The techniques and the methodologies required for each type of process have been clearly explained in the chapter. In this chapter challenges in the field of data analytics have been discussed and also the solutions related to these challenges have been proposed. At the end of the chapter a comprehensive conclusion has been provided to sum up all the detailed concepts included in the chapter.

Literature review

Grubbs and Frank (1950) framed the criteria for a sample testing the outlying observations. Chakravarti et. al. (1967) prepared a handbook related to the methods in the field of applied statistics. Anscombe (1973) conducted research on the graphs related to the statistical analysis. Box et. al. (1978) compiled their work related to the data analysis, design and also model building. Draper and Smith (1981) copiled their work related to the applied regression analysis. Snedecor et. al. (1989) compiled their work in the form of statistical methods. Barnett and Lewis (1994) conducted their research on the outliers and compiled their work in the form of outliers in statistical data. Parnas and Madey (1995) made a compilation of their work for functional documents for computer systems. Berry and Linoff (2000) concentrated their research on data mining process and compiled it. Giorgini et al. (2005) conducted research on the analysis of goal-oriented form for designing the data warehousing. Supakkul and Chung (2009) directed their research to deal with the problems related to the stakeholders. Barone et al. (2010) conducted a study on the enterprise modelling for business intelligence. Hopkins and Shockley (2011) focussed their research on data analytics, big data considering the practical aspects of the related problems. Horkoff et. al. (2014) conducted their research on strategic