Chapter 5
Data Mining Approaches for Sentiment Analysis in Online Social Networks (OSNs)

Praneeth Gunti
National Institute of Technology, Kurukshetra, India

Brij B. Gupta
National Institute of Technology, Kurukshetra, India

Elhadj Benkhelifa
Staffordshire University, UK

ABSTRACT

IoT technology and the widespread usage of public networking platforms and apps also made it possible to use data mining in extracting useful perspectives from unorganised knowledge. In the age of big data, opinion mining may be applied as a valuable way in order to classify views into various sentiment and in general to determine the attitude of the population. Other methods to OSA have been established over the years in various datasets and evaluated in varying conditions. In this respect, this chapter highlights the scope of OMSA strategies and forms of implementing OMSA principles. Besides technological issues of OMSA, this chapter also outlined both technical problems regarding its production and non-technical issues regarding its use. There are obstacles for potential study.

DOI: 10.4018/978-1-7998-8413-2.ch005

INTRODUCTION

Nowadays, there is a massive volume of data coming in, and every day it rises. The rapid advances in data analytics technology have contributed to the production of big data (Addo-Tenkorang, 2016). Furthermore, extensive data has been rendered possible by technologies. People are technologically smart of wearable technology systems, networking devices, and Internet or social network apps. For instance, businesses are profiting from collecting sales data because of the growing amount of transactional data. Various networked sensors are installed in cell phones, smart energy metres, cars, and manufacturing robots. Breakthroughs of Internet technology and sensors have rendered the Internet of Things (IoT) feasible. With the advent of social networking and mobile phones, people can communicate to build an extensive digital archive. For example, Twitter has many users producing 175 million messages every day(Yasin, 2014). Correspondingly, the space used to hold one second of a Video capture is 2000 times greater than a simple document. Besides, International Data Company's information stated that the world produced around 14ZB of knowledge in 2014. By 2022, the volume of data we make is projected to hit 50ZB; nearly half of it will be textual data produced by social networking, such as Facebook, Twitter, and smartphone immediate communication applications like Facebook and Whatsapp. Five hundred million messages are being delivered each day, and 40 million are posted every day. About 4.3 billion updates and 5.75 billion likes are shared on Facebook every day. Information will begin to arrive regardless of the rise in emerging technology.

The unrestricted use of computational information and enormous exchange of information added significantly to big data analysis. Big data and market analytics apply to two parts: big data and analytics. Big Data is valuable because it is entirely beyond human capacity to store, maintain, and interpret it. Big data is distinguished by a large number, range, and velocity. Volume reflects the amount of information that uses a significant range of documents. For e.g., some instances are Wal-Mart generates data around 2.5 Petabytes of data, Tesco generates data of about 1.5 billion pieces of information each month, and Dell maintains information storage that can handle data around 1.5 million records a day. Knowledge is a significant source of variation (Davenport, 2013).

The origins may involve sensors, social media platforms, online applications, smartphones, and other portable devices. The data may be in either structured data such as SQL or unstructured such as XML. Velocity shows how knowledge is readily gathered and provided through various channels. The data may be produced occasional, regular, and in real-time. Analytics relates to a firm's willingness to utilise mathematics, finance, calculus, chance, and optimisation to help them. Companies in most industries are highly data-oriented and gather data from a