



2

Exploring the Use of Big Data in Business Context

Almost all organizations collect and collate relevant data in various forms, such as customers' feedback, inputs from retailers and suppliers, current market trends, etc. The information, thus derived, is used by the management to take major organizational decisions. An organization generally has to spend huge amounts to collect data and information. For example, customer surveys or market research reports require a significant amount of investment by an organization. The cost of collecting information goes on escalating as an organization keeps on collecting more information. The continuously increasing cost decreases the value of the collected information. In other words, collecting and maintaining a pool of data and information is just a waste of resources unless any logical conclusions and business insights can be derived from it. This is where Big Data analytics comes into the picture.

This chapter provides an in-depth look at how Big Data influences businesses in today's world. The key is to understand how Big Data and different methods of data analytics are used in real time and why. How can an organization make optimum use of Big Data? How can large volumes of data be used to get better insights? How can the data obtained be used to form better business strategies, and thereby help in scalability and profitability? The key to implementing a Big Data solution is to manage Big Data to meet business requirements. Business insights gained from Big Data analytics help organizations to reduce their cycle time, fulfil orders quickly, cut excess inventory; and improve forecast accuracy and customer services by exchanging information, such as inventory, levels, forecast data, and sales trends. These insights can be applied to almost all the core domains of an organization and shared with partners, suppliers, customers, and other stakeholders.

In this chapter, you will learn how the data is actually used in the real world to get better business insights and explore future business expansions and also detect and prevent frauds in different industries.

SCENARIO

Mr. Smith, the data analyst of Argon Technology, plans to do some research on the use and effects of Big Data analytics on the business sector. This research requires the extraction and analyzation of data from different data sources to derive meaningful insights. He knew that all types of businesses give due importance to their customers' feedback, and customers generally express their opinions on social media platforms. So Mr. Smith decides to collect data for his research from social networking sites.

Use of Big Data in Social Networking

Human beings are considered social animals and prefer to live in a society. A human being lives in a social environment and gains knowledge and experience through communication. Today, communication is not restricted to meeting in person. The affordable and handy use of mobile phones and the Internet have made communication and sharing data of all kinds possible across the globe. Some popular social networking sites are Twitter, Facebook, and LinkedIn. These social networking sites are also called social media.

In this section, we analyze the effects of Big Data generated from the social media on different industries. Let's first understand the meaning of social network data.

Social network data refers to the data generated from people socializing on social media. On a social networking site, you will find different people constantly adding and updating comments, statuses, preferences, etc. All these activities generate large amounts of data. Analyzing and mining such large

volumes of data show business trends with respect to wants and preferences, and likes and dislikes of a wide audience.

This data can be segregated on the basis of different age groups, locations, and genders for the purpose of analysis. Based on the information extracted, organizations design products and services specific to people's need.

Figure 2.1 shows the social network data generated daily through various social media:

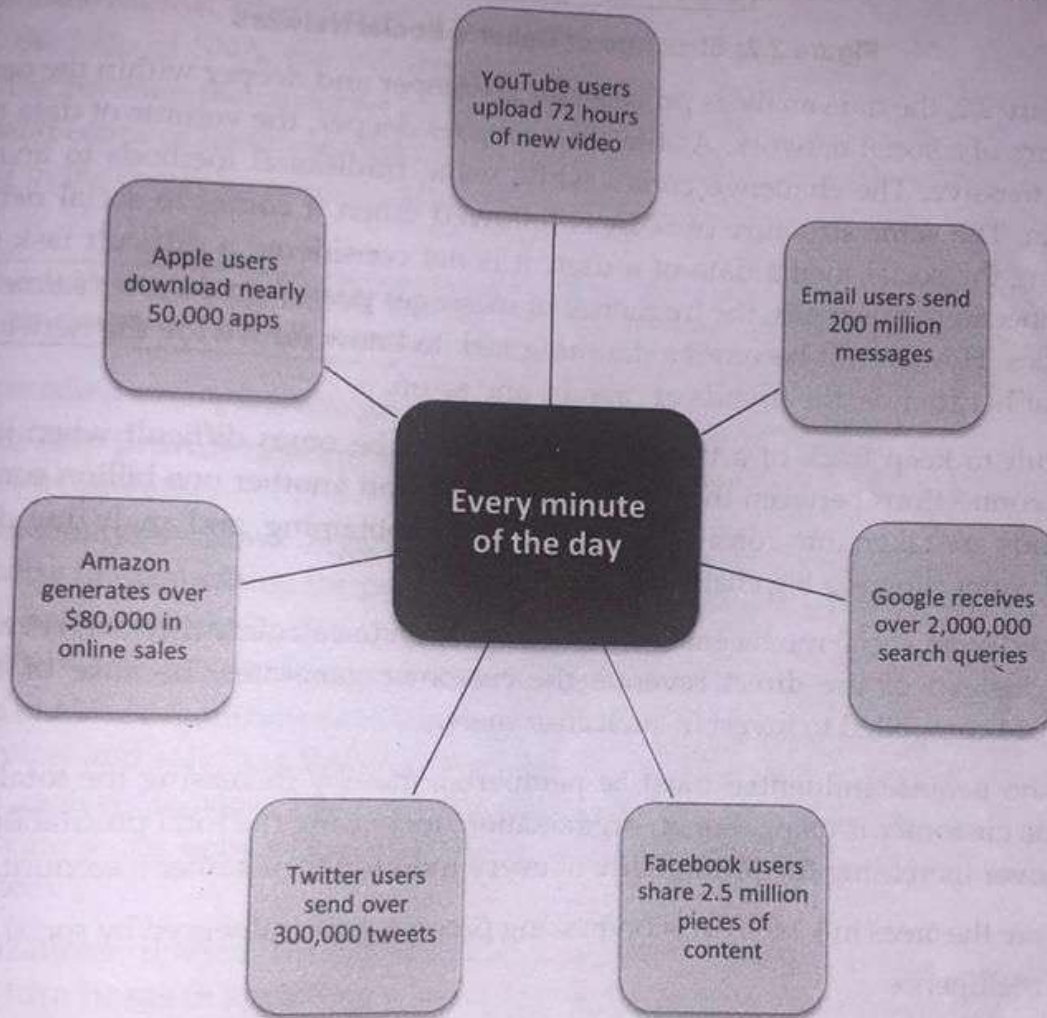


Figure 2.1: Social Network Data Generated Every Minute of the Day

Social Network Analysis (SNA) is the analysis performed on the data obtained from social media. As the data generated is huge in volume, it results in the formation of a Big Data pool.

Let's understand the importance of social network data with the help of an example of a Mobile Network Operator (MNO). The data captured by an MNO in a day, such as the cell phone calls, text messages, and other related details of all its customers is very huge in volume. This type of data is used daily for different purposes.

An MNO does not simply need to record and analyze the calls of a customer but the entire network calls related to that customer. The company must study the data of the people whom the customer called and also of the people in the customer's network who called back the customer. Such a network is called a social network.

Figure 2.2 shows the graphical view of how the caller's social network structure is created:

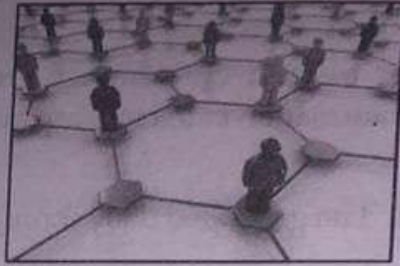


Figure 2.2: Structure of Caller's Social Network

As seen in Figure 2.2, the data analysis process can go deeper and deeper within the network to get a complete picture of a social network. As the analysis goes deeper, the volume of data to be analyzed also becomes massive. The challenge comes while using traditional methods to analyze the huge volume of data. The same structure of SNA is followed when it comes to social networking sites. While analyzing the social media data of a user, it is not considered a difficult task to identify the number of connections a user has, the frequency of messages posted on the user's timeline, and other standard metrics. However, it becomes a daunting task to know how wide the network of a user is, including his or her friends, the friends of friends, and so on.

It is not difficult to keep track of a thousand users, but it becomes difficult when it comes to one million direct connections between these thousand users, and another one billion connections when friends of friends are taken into consideration. Extracting, obtaining, and analyzing data from every single point of connection is a big challenge faced by SNA.

The data derived from social media enables an organization to calculate the total revenue a customer can influence instead of the direct revenue the customer generates. Because of this advantage, organizations are compelled to invest in such customers.

A customer who is quite influential must be pampered; thereby increasing the total profitability of the network the customer is using. For an organization, increasing the total profitability of a network takes priority over increasing the profitability of every individual customer's account.

The following are the areas in which decision-making processes are influenced by social network data:

- Business intelligence
- Marketing
- Product design and development

Business Intelligence

The data generated from different social media is analyzed to gain important business insights. Social Customer Relationship Management (CRM) data is the latest catchphrase used these days to describe this type of data. Such a data analysis helps in changing the perspective of an organization while valuing its customers. Instead of valuing a single customer, organizations can now calculate the value of the entire network that is influenced by that customer.

Consider an example of a mobile service provider that has a low-value customer. The customer has subscribed for a simple call plan and fails to provide any additional revenue as profit. As per the traditional method of evaluating the profitability of a customer, if the customer is not satisfied with

the services and if he or she wants to leave, then the company generally has no problems to let the customer go as he/she is providing low revenue.

However, with the help of SNA, the organization can now identify that some connections in the customer's network make a large number calls and text messages and have a large network of friends. With such an analysis, the organization might take an altogether different decision and might start valuing the customer more. This implies that the influence of a customer is very important to the organization. Studies have shown that when a user of a calling network leaves, others in the network often follow suit. With the help of SNA, the influence a customer can create on the network can be identified instead of the revenue the customer can generate directly. This approach helps to form an entirely different perspective of handling customer requirements. The mobile service provider can now think of investing in the customer to retain him/her.

NOTE

SNA can also help in law enforcement and anti-terrorism efforts as it is possible to identify trouble groups or people who are directly or indirectly connected to each other. Such type of analysis is called link analysis.

From the preceding examples, we can derive the following business insights:

- ❑ SNA can help provide new contexts in which decisions are taken on the basis of available data and not on opinions.
- ❑ Big Data analysis allows organizations to shift goals from maximizing individual account profitability to maximizing the profitability of the customer's network.
- ❑ Big Data helps organizations to identify highly connected customers. It also helps in identifying when, where, and how to align and focus marketing efforts in building a better brand image.
- ❑ Big Data enables organizations to lure highly connected customers by offering them free trials of their services and soliciting their feedback for the betterment of their products and services.
- ❑ Big Data analysis assists organizations by encouraging internal customers to become more active on corporate social networking sites and provide comments and opinions on various products and services.

Some organizations reward their influential customers with discounts and offers, and these customers in turn helps in spreading a positive brand image of the organization. Social networking sites such as LinkedIn or Facebook can obtain insights on the advertisements that most users prefer. This is achieved by designing advertisements based on interests, likes, and preferences that customers as well as their circle of friends, contacts, and colleagues have personally opted for.

The gaming industry also makes use of social network data to develop business intelligence. Analysis of social network data or link analysis helps in tracking telemetry data about gaming including the information about who is playing which game, who plays with whom, for how much time they remain online to play the game, changing patterns of playing among people or groups, and much more.

NOTE

The term telemetry is popular in the video gaming industry, and refers to the capture of in-game activities. Telemetry is analogous to a weblog analyzer tool and captures the activities performed by players while they navigate through the game.

Telemetry also helps in finding the preferred partner for a player for a particular game. Players are often segmented on the basis of their individual playing style.

For example:

- One player segment can opt for advancing as fast as possible to beat a level with the purpose of only completing that level.
- Another player segment might be focused at collecting the bonus points available prior to the completion of each level.
- Another player segment might target at exploring every aspect of each level during the play.

A game developer might find this information very valuable while designing new games, updating existing ones, and targeting a particular audience for games. This information also enables the game developer to provide options and suggestions to players on the groups available to them when other players/users log in. Such suggestions and options keep the players engaged and interested in the game, which helps in generating revenue.

Marketing

Today the preferences of consumers have change due to theirs busy schedules. They no longer have the time to read newspapers thoroughly, watch all the TV commercials, or go through all the e-mails they receive in their inbox. Consumers can now make their preferences clear and select the marketing messages they wish to receive—when, where, and from whom. In today's competitive scenario, marketers aim to deliver what consumers want by using interactive communication across digital channels such as e-mail, mobile, social, and the Web.

These channels, in turn, generate the social data required to provide insights based upon the brand preferences of a target audience, the tone of its voice, the other brands it discusses, its interests, and other information. Conducting social network analysis of this data can generate very useful and meaningful business insights that may help organizations to take timely and informed decisions. When a comparison is made between the efforts spent on the social media marketing platform and the e-mail marketing platform, it is found that the marketing efforts spent on the former yield more returns as compared to the latter.

Let's now take a case study to understand the importance of social media marketing.

CASELET

Walmart has acquired a social media analytics company, named Kosmix, and created @WalmartLabs, a division that analyzes media communication to understand retail trends. The director of product management at @WalmartLabs, Tracy Chu, said the following in a blog post:

"@WalmartLabs is working on ways to help Walmart interpret social media to predict trends and learn more about what our customers want. We are mining social media sources like Twitter and Facebook, blogs, search activity, as well as transaction data to find useful insights." One of the key responsibilities of this division is to monitor public domain conversations and then position Walmart products accordingly. @WalmartLabs has been tracking social chats for analyzing trends in various categories such as holiday toys, online games, and mobile commerce.

Recently, @WalmartLabs has also launched Shopycat, a Facebook app. This app analyzes data from social networks and accordingly makes gift recommendations as per your friend's likes and dislikes. Then the app matches these recommendations to the products available in Walmart and some other site catalogs. Not only this, the app also suggests gift cards and the location of the closest Walmart store from where the gifts can be purchased.

Brand popularity is very critical in marketing. There are apps in which customers can rate a brand on a scale of 1-10 according to how likely they are to recommend the product or service to a friend.

Once a brand's rating crosses 300, the application sends out a report that contains information on what customers feel about the product, along with a detailed analysis of the brand's reputation as compared to the competition.

Affiliate marketing is a reward-based marketing structure, where an affiliated company uses its own market effort to trigger off customers for another company and in turn, is rewarded by the benefited company. Today, one can hardly find any major brand without an affiliate program. According to analysts in this field, affiliate marketing has already become a \$3 billion industry. Websites, such as couponmountain.com, earn revenue of multi million dollars a year by doing affiliate marketing for the brands they promote.

Product Design and Development

With the increasing popularity of social media and growing volume of data every second, organizations competing to make it big in the market must not only identify and extract the information relevant for their company, products and services, but also comprehend and respond to the information on a continuous basis. A system that is able to represent a sentiment as data with a high degree of accuracy provides the client a means to access information on a social platform. To be able to measure sentiments more meticulously is of great value while designing a product or service. Brands must understand the importance of the demographic information they receive to devise better target products and programs.

By listening to what consumers want, by understanding where the gap in the offering is, and so on, organizations can make the right decisions in the direction of their product development and offerings. In this way, social network data can help organizations to improve product development and services, making sure that the consumers ultimately get the products and services they want.

Sentiment analysis refers to a computer programming technique to analyze human emotions, attitudes, and views across popular social networks including Facebook, Twitter, and blogs. The technique requires analytic skills as well as advanced computing applications.

Businesses, research organizations, and marketing professionals across the globe use sentiment analysis in one form or the other to identify and measure customers behavior and online trends. However, this technique is still evolving, and the full potential of sentiment analysis is yet to be explored by marketers and other business professionals. Most organizations today simply rely on the number of likes, tweets and comments, instead of actually studying the quality of the sentiments expressed in the conversations.

CASELET

According to a survey conducted by MSN Money, American Airlines has been ranked as one of the most disliked companies in the US.

Studies reveal that American Airlines has about 346,259 "followers" on Twitter and 273,591 "likes" on Facebook. But this cannot be taken as a true indicator of the company's popularity. Deep studies of the sentiments of customers reveal that the trends of online conversation about the company are negative, which indicate that it is allegedly one of the most disliked airlines. Thus, clearly, the company's efforts to engage social media community have not paid any beneficial returns.

To improve their image and ranking, American Airlines should focus more on the sentiment and emotive data and the "correct" types of incoming data sources rather than just counting the numbers of "followers" and "likes".

SCENARIO

While going through the feeds collected from social networking data, Smith notices that almost all types of businesses have suffered on account of fraudulent activities and many organizations are now turning to Big Data analytics to solve this problem.

Use of Big Data in Preventing Fraudulent Activities

A fraud can be defined as the false representation of facts, leading to concealment or distortion of the truth. Frauds can be committed by both words and conduct and is intended to deceive the other party, generally to gain an advantage over it in some manner. Frauds that occur frequently in financial institutions, such as banks or insurance companies, or involve any type of monetary transactions, such as in the retail industry, are called financial frauds. In such fraudulent cases, online retailers, such as Amazon, eBay, and Groupon, tend to incur huge expenses and losses.

The following are some of the most common types of financial frauds:

- **Credit card fraud**—This type of fraud is quite common these days and is related to the use of credit card facilities. In an online shopping transaction, the online retailer cannot see the authentic user of the card and therefore the valid owner of the card cannot be verified. It is quite likely that a fake or a stolen card is used in the transaction.

In an online transaction, in spite of the security checks such as address verification or card security code, fraudsters manage to manipulate the loopholes in the system.

- **Exchange or return policy fraud**—An online retailer always has a policy allowing the exchange and return of goods, and sometimes people take advantage of this policy. These people buy a product online, use it, and then return it back as they are not satisfied with the product. Sometimes, they even report non-delivery of the product and later attempt to sell it online. What leads to such a fraud is that retailers encourage consumers to order products in bulk and later return the ones that they don't require.

Such a fraud can be averted by charging a restocking fee on the returned goods, getting customer's signature on the delivery of the product, and staying cautious of such customers who are known to commit such frauds.

- **Personal information fraud**—In this type of fraud, people obtain the login information of a customer and then log-in to the customer's account, purchase a product online, and then change the delivery address to a different location. The actual customer keeps calling the retailer to refund the amount as he or she has not made the transaction. Once the transaction is proved fraudulent, the retailer has to refund the amount to the customer.

All these frauds can be prevented only by studying the customer's ordering patterns and keeping track of out-of-line orders. Other aspects should also be taken into consideration such as any change in the shipping address, rush orders, sudden huge orders, and suspicious billing addresses. By observing such precautions, the frequency of the occurrence of such frauds can be reduced to a certain extent, but cannot be completely eliminated.