

### ***The way of Spotify deriving Value from Big Data***

To maintain academic honesty, many tertiary schools adopt Turnitin to check the similarity between other published academic papers. These “plagiarism checkers” have their algorithms to evaluate the “distances” from the massive amount of essays located at their database and web pages. Mining for resemble items has become an interest for many companies to boost their quality of services and products. In the era of the Internet of Things, data storage is no longer a huge expense for enterprises, and cloud computing improves their flexibility of works. As data is growing faster than ever, the application of big data has derived many profitable opportunities. Spotify takes a fancy to such opportunities and focuses on working with a vast amount of structured data.

Whenever we talk about Spotify, the first thing comes to mind must be the platform of streaming music service on our smart devices. Surprisingly, it is not what we expected, Spotify is a technology company that regards data as the core development of the platform. We claim Spotify as a data-driven company is that it relies on data analysis and algorithms to enhance the user experience and target audiences with advertisements. The prediction engine makes Spotify famous since it has a high success rate in recommending songs that suit users favour. In about a decade ago, Spotify successfully predicted the Grammy Awards winners by using data collected from the users.<sup>1</sup> Up to the current version of Spotify music platform, there is already a Discover Weekly playlist (personalised playlist) for every user. After a couple of weeks in collecting data, Spotify will be able to investigate “taste profile”, known as the listening pattern, and promote some tracks to users that they have not listened before.

Spotify and big data are closely connected and inseparable. Before we dig into the processes of deriving value from data, let us have a view on the data collected by Spotify. The concept of big data is relative to the technologies to store and process data. Although there is no official definition of big data, we can easily understand the concept with its characteristics. There are 248 millions of active users listen to music with the platform of Spotify.<sup>2</sup> As the adoption of cookies, Spotify directly records every action from the users. There is an enormous volume of data contributed. Moreover, Spotify collects a variety of data sources. The duration of tracks (numeric) and the types of the genre (text) are some data types that we can easily to imagine. To ensure the prediction be accurate, Spotify will also accumulate subtle data, such as album artwork (image) and song rating (numeric).

---

<sup>1</sup> <https://datafloq.com/read/big-data-enabled-spotify-change-music-industry/391>

<sup>2</sup> <https://newsroom.spotify.com/company-info/>

One interesting fact about Spotify is that there are 4 data centres around the world to handle the high velocity of data generation. According to Spotify, there are 4 terabytes of data is recorded in their database per day.<sup>3</sup> The quantities of data is astonishing. Furthermore, as mentioned previously, Spotify directly transacts data from users, most of the record has no ambiguity and noise. The veracity of data is also satisfactory. Although there is some user feedback, an example of unstructured data, the majority of data have their format so that they can be quantified and verified with high accuracy.

Time series and regression are common statistical predictive models used in the financial sector. However, Spotify adopts more advanced methods to develop the recommendation engine. Collaborative filtering is one of the techniques used in Spotify to mining the similar “taste profile”.<sup>4</sup> The method groups users into different clusters by the similarity of listening habit, and predicts their potential interests of songs. The method involves much in-depth knowledge, such as clustering along with k-means and k-nearest neighbour method. Euclidean distance, the “distance” between clusters, requires a large amount of data to ensure the prediction. Owning big data, Spotify can derive data into wisdom and make better decisions based on understanding.

Improvement is the ideal result of deriving data. As data represents the behaviour of the users, identifying the segments of the customers and their needs is the principal use of data. Big data allows Spotify to study the insights of data and iteratively improve its service by learning the information from data. From 2015 to 2019, there is a 200% of increment in the number of active users on the Spotify platform. The high diathesis of service makes Spotify more famous around the world, and the company made a total revenue of USD 29 billion in 2019.<sup>5</sup> Without big data, Spotify cannot make such tremendous success. Spotify turns data into value, the last characteristic of big data, in gaining competitive advantages to continuously improve its streaming music service. We do not know what new features will appear on the platform, but we will be watching with our excitement how Spotify derives value from big data.

### ***References***

DiFranza A. (2019) “SPOTIFY: BIG DATA SHOWS BIG RESULTS” *Northeastern University*. 29 January, 2020. <https://www.northeastern.edu/graduate/blog/spotify-big-data/>

Mark R. (2013) “How Big Data Enabled Spotify To Change The Music Industry” *DataFlog*. 29 January, 2020. <https://dataflog.com/read/big-data-enabled-spotify-change-music-industry/391>

Singh P. (2019) “Spotify Statistics 2019 – 2020: What’s the Future of the Media Streaming App By Numbers” *Appinventiv*, 27 January, 2020. <https://appinventiv.com/blog/spotify-statistics-facts/#A4>

Spotify (2019) “Company Info” *Spotify News Room*. 27 January, 2020. <https://newsroom.spotify.com/company-info/>

---

<sup>3</sup> <https://dataflog.com/read/big-data-enabled-spotify-change-music-industry/391>

<sup>4</sup> <https://www.northeastern.edu/graduate/blog/spotify-big-data/>

<sup>5</sup> <https://appinventiv.com/blog/spotify-statistics-facts/#A4>