Cristian Pintor

ITMD 413-V01

05/04/2021

What is Big Data?

Big Data is the collection of large volumes of data that grows exponentially with time. This data is so large and complex that traditional data management tools are not capable to storing it or processing it. These massive volumes of data can be used to solve business problems that wouldn’t have been able to be solved before.

Big Data is important because it has a lot of uses and benefits. It can be used for various business activities, from customer experience to analytics. One of them is product development, where companies like Netflix use big data to create models for future customer demand. These predictive models are built for new services and products based on past products with the data that was gathered from them. Another way in which big data can be used is for customer experience, where data is gathered from different sources such as social media, web visits, call logs, etc. This data is used in order to improve customer experience and provide a better service and experience.

There are three types in which big data can be classified as: structured data, unstructured data, and semi-structured data. Structured data is organized and easy to work with. It is set up in a similar fashion such as spreadsheets where information is grouped into columns and rows. Unstructured data is simply data that is nor organized. It can be things such as un-transcribed phone calls or poorly formatted tweets. Lastly, semi-structured data is unstructured data with metadata attached to it. This can be inherent data collected, such as time, location, device ID stamp or email address, or it can be a semantic tag attached to the data later.

According to Oracle, the three Vs stand for volume, velocity and variety. Volume stands for the large volumes of low-density unstructured data that is gathered. It can be anything from Twitter data deeds, to clickstreams on a website or app. With velocity, it is the speed at which new data is coming in. Some data may come in real-time, while other may come in different batches at different times. Finally, variety is the different formats in which data being received. It may come in traditional formats such as excel sheets or CSV files, but a lot of data is also coming in as PDFs, images, Tweets, videos, graphics etc. All of these types of data may require more analytics skills to review it and more time to go through it.

References

Allen, Richard. “Types of Big Data: Understanding & Interacting With Key Types.” *SelectHub Raquo*, 12 Oct. 2020, www.selecthub.com/big-data-analytics/types-of-big-data-analytics/.

“Big Data: The 3 Vs Explained.” *Big Data LDN*, 12 Apr. 2019, bigdataldn.com/intelligence/big-data-the-3-vs-explained/.

“Big Data: What It Is and Why It Matters.” *SAS*, www.sas.com/en\_us/insights/big-data/what-is-big-data.html.

“What Is Big Data?” *Oracle*, www.oracle.com/big-data/what-is-big-data/.