

BIG DATA

Mr. Amit Sharma

Abstract

Nowadays data is plays an important role in everyone's life, it is growing rapidly and it is common with all. As we know, just like us the organizations, companies, institutes, etc. are dealing with large amount of Data; this large amount of data can be said as "Big Data". It is not easy to deal with these types of large data. The operations of Retrieving, Searching, etc. become very difficult. Large Data Is Very Useful too in different IT and Computer fields like: - IOT's, Machine Learning, Artificial Intelligence, etc. This Big Data is not only useful in IT or Computer Fields But also in crypto-currency Mining, predictions of a user's choices, also useful in decision making. Big Data which contains variety of data, with different-different data types or even dataset of non-similar data. It is basically 'Large Data' which is rapidly growing day-by-day or even in every Second. This Data is very difficult to Process because of its Complex structure. A lot of Raw data or Even already Processed Data can be found here, just difficult to find as it lead to wrong combinations or incorrect Information. In this Paper will cover some details on Big Data.

Keywords :- [Big Data, Large Data.]

Introduction

Big Data is an Emerging Trend in fields of Information Technology. There are many features of 'Big Data' which are commonly called V's. there are different numbers of V's like 3V, 5V, 7V. Here we will look over all the 7V's namely: - volume ,velocity ,variability ,veracity ,variety ,value ,Visualization. We will cover all these further. Big Data is not a very new term but it started coming in eyes of people in around 2005, at that time people realised that how much data is generated at big platforms by them, as a user you can also see how much data is generated by you for different - different big platforms like - "Social Media", it is one of the biggest and common example for the big data generators, millions and billions of user use social media daily for hours and post a large amount of data daily. This is just one example of 'Big Data', but we can see a lot more like - "Share market" same as social media millions of user invest in share market and the data is not easy to handle as not only people but even companies, organization, government, institutes etc. do investments and keep eyes for hoping of best results. These shares depends of big data analysis. So as we can see that Big Data is useful here too. Big Data also help us by providing complete and accurate answers, as it is collection of large amount of information that can be used further and even to get more accurate information. Big Data is also useful to get new Knowledge and new Facts that maybe not have been discovered yet. This is how the Big Data useful to us.

Definition of Big Data According to Wikipedia :-

" The term *big data* has been in use since the 1990s, with

some giving credit to "John Mashey" for popularizing the term. Big data usually includes data sets with sizes beyond the ability of commonly used software tools to capture, curate, manage, and process data within a tolerable elapsed time."

John R. Mashey (born 1946) is an American computer scientist, director and entrepreneur.

Features OR Characters of Big Data :- As we have already read about the names of features or characters of Big Data. They are 7V's :- volume ,velocity ,variability ,veracity ,variety ,value ,Visualization. Here we will lookat these in detail.

- Volume :- Simple refers to amount. Here in Big Data the Volume is amount of Data stored, it can also said as capacity of data, like how much Data is generated by a user every day that can be considered as volume of data generated by user. Big Data Grows rapidly as people generate large amount of data every second and the size of that data is called 'Volume of data'
- Velocity :- The speed of data generation by users at a time and the systems' handling speed of receiving or sending of data can be called as 'Velocity'. It can also be the sufficient capacity to handle the data, delivery or action time.
- Variability :- it is a process where the data is cleaned , to remove incomplete or misinformation from data which can create wrong information or wrong facts or wrong knowledge , it simply means to validate the data can be called as 'Variability'
- Veracity :- This can be said as a process of finding accurate data, basically Organization, Government,

Tech giants, Institutes etc. need to focus on the data they required, so this process is to be followed to find accurate data

- Variety :- This means the forms or types of data, as we know all the data available in Big Data is not generated from single source. It is generated from different sources and it is in different types like :- Numeric Data, Graphical Data, Pie Charts, Images, Videos, Text, Documents Files etc.
- Value :- The Data which have some sort of meaning can be simply said as 'Value'. The data must implies meaningful and useful information so that the user can get best knowledge from it
- Visualization :- This is a process of Visualizing or showing the data in an understandable form as we know that Big Data is a large amount of data which is not easy to understand, after this process of visualizing the data will be showed in a understandable format like :- presenting some data in graphs, pie charts, tables etc.



Figure 1. Taken From Shutterstock

Application of Big Data :- Big Data is applicable in many fields so here we learn about some application areas.

- Data of Customers are analysed to know there preferences this help to predicte the customers choices and prediction of their futur purchases, this is way the big companies who are in sales are always in demand of customer data.
- This allows companies to predict the next travel or tours of customers too, this is why the travel and tourism companies always in demand of customers data
- The Big data is applicable in IOT's as we knows that nowadays we use smart products which are based on

IOT(Internet of Things)these products are like smart lights, smart T.V. , etc. which are basically smart home appliances which can connect to a network and can operate on demand of user even without touching. The big data analysis allows companies to get as useful data to make these types of products.

- Big Data is useful in Education sector as well, because of its famous property of getting new knowledge or new facts helps in education fields too.
- It is also helpful in media or entertainment fields as the companies analysis user data to understand the preferences or likes of people, so that they can plan the future projects accordingly and can get as much as more users they want.



Figure 2. Self Made

Challenges in Big Data :- As Big AS Data is there it have some challenges to handle it or in a lot other ways, so here we will have look on those challenges about Big Data.

- Security Issues :- Big Data is inclusive of a lager amount of Data as know it includes a lot information about different - different things,these Data may include some sensitive Data of someone(Any Individual or of a group), some personal details or other related data. Which is a big security concern for that person. To solve this type of problem regular security analysis should run to remove these types of data and make it secure for everyone.
- Analyses issues :- Because of the large amount of data it is very difficult to analysis data, as it needs some powerful tools and technology, without these it is not possible to deal with these types data.
- Tools :- It is one of the biggest issue as choose right tools foe data analysis.if someone choose wrong tool it can decrease the quality of data required. "Hadoop" is one of the famous tool which can be used for data

analysis.

- Quality of Data :- It is one of the basic requirement for any firm, organization, companies, etc. as the real data or information which is required can be found by analysing the Big Data.

Conclusion :- Big Data is large set of data which is inclusive of different types of data, which can be useful in different-different ways. The data is first generated by users or people around the world. Then this data is analysed and reduced to understandable form, here

Reference :-

Definition of Big Data From Wikipedia https://en.wikipedia.org/wiki/Big_data on 19/10/2023

the data can be understandable and informative. This data is required by the tech giants to gain user information and data. To understand users preferences so that they can make products accordingly. The quality of data is must ensured before using it. The right tools must be used , ensuring of privacy and security of each and every individuals personal data. After Ensuring all this the data can be used. All these things concluded that Big Data is very Useful but also required a lot of efforts to get correct information or data.

Figure 1. is taken from Shutterstock <https://tinyurl.com/35admby5> on 19/10/2023