# Big Data Analytics

*Shabeena.R,Oviya.G
Adhiyamaan College of engineering, Hosur, Tamil Nadu, India.
*Corresponding author Email:shabeenariyaz1617@gmail.com

**Abstract.**Big data analytics describes the process of incovering trends, patterns&correlations in large amount of raw data to help ma data-information descision .these process use familiar statistical analytics technique-like clustering and regression and apply them to more extensive data sets with the help of power tools. Big data is a buzzword or catch phrase used to describe a massive volume of both structured and unstructured data that is so large that it's difficult to process using traditional data base and software techniques. The world's data collection is researching a tipping point for major technological Changes that can bring new ways in division making, managing our health, cities, finance and education. When dealing with larger datasets,organizations face difficulties in being able to create, manipulate and manage bid data. Our professional front we may think about the data growing in our company. Giga bytes-> terabyte->peta byte->exa byte->zetta byte->yotta byte. For example: it the data grow then it will move to data ware house and pushed to SQL analysis services cube for analytics services cube for analytics. However think about one day someone ask to analysis/mining the whole data or what else can predict with adding more data to data mining which we use to ignore earlier. The question is how we will handle the situation where no available technology (in company) is capable of handling such huge data?. We have to think about what value I am going to get .1 mean the outcome or ROI of the implementation of BIG data platform ,think about what outcomes/benefit we can provide with this platform ,how it's going to be different,what challenges we are trying to solve. That is called the big data concept. Next question is how to handle it. Hadoop is the technology which will not only help with storage but also in querying the data.

## 1. Big Data

Big data is nothing but an assortment of such a huge and complex data that it become very retrieve to capture,store,process,retrieve and analyze it with the help of on- hand database management tool or traditional data processing technique.There are many real life examples of big data! Facebook is generating 500 + terabyte by of data per day,NYSE(New York stock exchange) generates about 1 terabyte of new trade data per day, a Jet airline collects 10 bite of censor. Data for every 30 minutes of flying time. All these are day today examples of big data!As of December 31, 2012 there are 1.06 billion monthly active users on Facebook and 680 million mobile users' .on average 3.2 billion likes and comments or posted every day on Facebook 72 percentage of web audience is on Facebook. There are so many activities going on Facebook from wall post, sharing images video writing comments and liking post etc. In fact Facebook started using Hadoop in mid-2009 and was one of the initial user of Hadoop
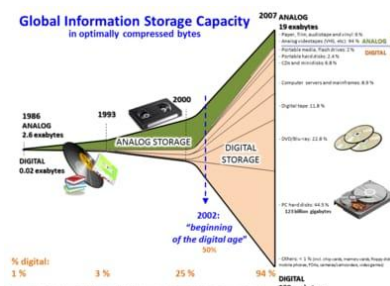


**FIGURE 1.** Global Information Storage Capacity

## 2. Characteristic

According to IBM the three characteristics of big data are:Volume: Facebook generating 500 + terabyte of data per day, By analyzing large-scale social media and browsing behavior,businesses can create a more complete profile of customers and stream them into narrow segments of preferences.With this level of specificity and insight,businesses can make more informed marketing decisions to promote their product or service to those more likely to want it.80% of your Netflix stream is influenced by its algorithm-powered recommendation system.Within this market,the data science platforms,tools and service segment is forecasted remain the fastest growing, with a CAGR of 17.82% through 2026 from $12.8 billion by 2026 from $5.6 billion this year.The size of data determine the value and potential insights, and whether it can be considered big data or not. The size of big data is usually larger than terabyte and petabyte.Velocity: analyzing 2 million records each day to identify the reason for losses.The type and nature of data. The early technology's like RDBMS where capable to handle structure data efficiently and effectively.Effective usage of hidden insight.

Exposed from the data collected via social media, log files, sensor etc.Big data draws from, text images, audio, video plus IT complete missing pixes through data fusion.Variety: images, audio, video, sensor data, log files etc. with time, data volume is growing exponentially. Earlier we used to talk about megabytes or gigabytes. But times have arrived when we talk about data volume in terms of Terabyte, petabytes and also zettabytes! Global data volume was around in 2011 and is expected to be in 2015. It is also known that global information doubles in every 2 years!



**FIGURE 2.** Huge amount data

## 3. Analysis

Effective analysis of big data provide a lot of business advantage as organisation will learn which area to focus on and which area or less important.Big data analysis provides some early indicator that can prevent the company from a huge loss or help in grasping a great opportunity with open hands! A precise analysis of big data help in decision making for instance, now a day'speople relay so much on Facebook and Twitter before buying any product or service.All thanks to big data explosion.
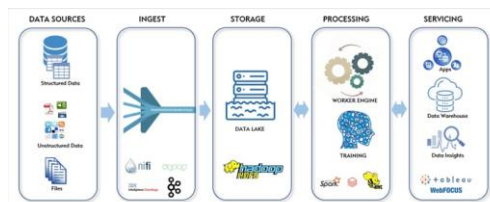


**FIGURE 3**. Data processing

Data scientist or soon replacing business analystaur data analyst. Data scientist or experts who find solution to analyse data.Just as web analysis, we have data scientist who have good business inside as to how to handle your business challenge. Shark data scientist or not only involved in dealing business problem, but also choosing the relevant issue that can bring value addition to the organisationThe term big data has been used since the yearly 1990 all though it's not exactly known who first wed the term most of people credit John R. Mashey. Hadoop does not have any expanding version like oops. The charming yellow elephant you see in basic named after Doug's son's toy elephant!Everyday a large amount of unstructured data is getting dumbed into our machine. The major challenge is not to store large data set in our system but to retrieve and analyse the big data in the organisation. That two data present in different machine are different location. In this situation unnecessity for hadop arises. Hadoop has the ability to analyse the data present in different machine at different locality very quickly and in a very cost effective way. It uses the concept of mapReduce which enables it to divide the query in two small part and process them in parallel. This is also known as parallel computing.

## 4. What is the future of big data?

In the future, big data analytics will increasingly focus on data fresheners with the ultimate goal of real time analysis, enabling better informed decision and increased competitiveness.Be ready for the future of big data analytics: While many large companies are already edging closer to, if not already fully embracing, all of these trends, giving them an edge over their competitors, the future of big data analytics is no longer locked behind a wall of price barriers. Data engineers and scientists are developing innovative ways to uncover insights hidden beneath the heap of data without requiring the budget of a Fortune 500.We're going to see a lot more small and mid-size companies incorporating big data analytics into their business strategies.The future is bright for those who take action to understand and embrace it.To learn more about how data observability – and other trends in the analytics space – can level up your business, schedule time with the form below.

## 5. Application

Big data has increased the demand of information management specialists so much so that,SoftwareAG,Oracle Corporation,IBM,Microsoft,SAP,EMC,HP and Dell have spent more than $15 billion on software firms specializing in

data management and analytics. In 2010, this industry was worth more than $100 billion and was growing at almost 10 percent a year, about twice as fast as the software business as a whole.



**FIGURE 4**. Data applications

Developed economies increasingly use data-intensive technologies. There are 4.6 billion mobile-phone subscriptions worldwide, and between 1 billion and 2 billion people accessing the internet. Between 1990 and 2005, more than 1 billion people worldwide entered the middle class, which means more people became more literate, which in turn led to information growth. The world's effective capacity to exchange information through telecommunication networks was 281 petabytes in 1986, 47 petabytes in 1993, 2.2 exabytes in 2000, 65 Exabyte in 2007and predictions put the amount of internet traffic at 667 exabytes annually by 2014. According to one estimate, one-third of the globally stored information is in the form of alphanumeric text and still image data, which is the format most useful for most big data applications. This also shows the potential of yet unused data (i.e. in the form of video and audio content).While many vendors offer off-the-shelf products for big data, experts promote the development of in-house custom-tailored systems if the company has sufficient technical capabilities.

## 6. Conclusion

Big data can also be leveraged to forecast weather, natural disasters, urban and community planning, traffic management, logistics and machine efficiency, personalized healthcare, customized learning, autonomous vehicles, fraud detention, translation, smart homes, robotics, etc.The amount of big data is already massive, but it is expected to grow exponentially as new technologies such as the more pervasive IoT devices, drones and wearables will jump into the fray.

## References

[1]. Figure:https://d1jnx9ba8s6j9r.cloudfront.net/blog/wp-content/uploads/2018/06/Five-Vs-of-Big-Data-What-is-Big-Data-Edureka-768x432.png
[2]. https://www.slideshare.net/nandhiniarumugam619/big-data-abstract-45974848
[3]. https://en.wikipedia.org/wiki/Big_data
[4]. https://www.montecarlodata.com/blog-the-future-of-big-data-analytics-and-data-science/