BIG DATA-TRENDS, ANALYSIS AND CHALLENGES

[1]Apoorva Verma , [2] Dr. Monika Rathore
[1]Student, [2]Associate Professor
International School of Informatics and Management,
Jaipur, India

Abstract:Welcome to the age of Big Data! A report by the forum "Big Data, Big Impact" declared big data as a new class of economic asset, like currency or gold. Big Data is data set that are so voluminous and complex that traditional data processing application software are inadequate to deal with them. This paper describes about growing impact of big data on industries and about challenges being imposed by big data and about benefits of big data.

Index Terms: Big Data, Hadoop, Map Reduce, Big data analytics

I. INTRODUCTION

People lives on-demand, on-command Digital universe with data prolife ring by Institutions, Individuals and Machines at a very high rate. "Big Data" is appearing is many contexts from meteorological, genomics, complex physics simulations, biological and environmental research, finance and environmental research, finance and business to health care.As the name implies "**Big Data literally means large collection of data sets containing information."** [4]The volume and the heterogeneity of data with the speed it is generated, makes it difficult for the present computing infrastructure to manage Big Data. Traditional data management, warehousing and analysis systems fall short of tools to analyze this data.Due to its specific nature of Big Data, it is stored in distributed file system architectures. Big Data is a heterogeneous mix of data both structured (traditional datasets –in rows and columns like DBMS tables) and unstructured data like e-mail attachments, manuals, images, PDF documents, medical records such as x-rays, ECG and MRI images, forms, rich media like graphics, video and audio, contacts, forms and documents. Businesses are primarily concerned with managing unstructured data, because over 80 percent of enterprise data is unstructured and require significant storage space and effort to manage.

Big data analytics is the area where advanced analytic techniques operate on big data sets. It is really about two things, Big data and Analytics and how the two have teamed up to create one of the most profound trends in business intelligence (BI).

Section 2 describes about **Growing Impact of Big Data in Industry**: -This section is all about significance of big data on modern industries.
Section 3 describes about **Character tics and challenges posed by Big Data: -** This section mainly describes what are the challenges which are faced due to big data and some of the solutions to overcome through them.
Section 4 describes about **Benefits of Big Data: -** This section tells about merits of big data in practical life.
Section 5 describes about **Future Research and Trends in Big Data: -** This section speaks about what are the possible fields and research which can take place in big data along with trends in it.

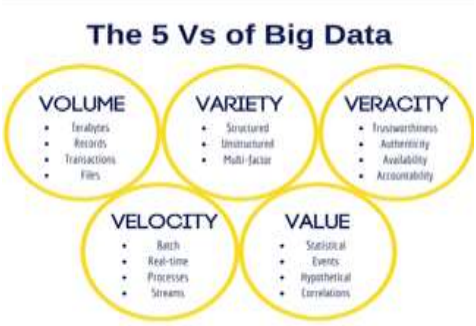
II GROWING IMPACT OF BIG DATA IN INDUSTRY

Big data is being used in numerous no. of fields like Govt., International development, manufacturing, Health care, Education, Media, Internet of Things, Information Technology. For modern industry, data generated by machines and devices, cloud-based solutions, business management, etc.,has reached a total volume of more than 1000 Exabyte annually and is expected to increase20-fold in the next ten years.From an industry point of view, big data is going to play an important role in the fourth industrial revolution. The first industrial revolution (from the end of 18th to the start of 20th centuries) depended on water and steam power, the second (from the start of 20th century to early seventies) depended on mass production, based on division of labor and electrical energy, and the third (from early seventies to the present day) dependedon electronics and on IT for further automation of production. The fuel of the fourth industrial revolution, dubbed as ''Industries 4.0'' by the German Government, will be big data to be made available through Cyber- Physical Systems (CPS). The primary objective behind the use of big data in industrial applications is to achieve a fault-free and cost-efficient running of the process, while realizing the desired performance levels, especially with respect to quality.

III CHARACTERTICS AND CHALLENGES POSED BY BIG DATA

Big data challenges include capturing data, data storage, data analysis, searching, sharing, transfer, visualization, querying, updating information, privacy and data source. Database technology (including parallel databases) was considered for the task but was found to be neither well-suited nor cost-effective for solving problems with big data.

For solving these problems techniques like Hadoop and Map reduce were developed.IBM data scientists break big data into four dimensions: volume, variety, velocity, veracity and value.



1.Volume: -Volume refers to the incredible amounts of data generated each second from social media, cell phones, cars, M2M sensors, photographs, video etc. We now use distributed systems where parts of the data are stored in different locations and brought together by software. With just Facebook alone there are 10 billion messages, 4.5 billion times that 'like' button is pressed and over 350 million pictures are uploaded every day. Collecting and analyzing this data is clearly an engineering challenge of immensely vast proportions.

2.Variety: - Variety is defined as the different types of data we can use. Data of today looks very different from the data of the past. We no longer have structured data that just fits easily and nicely into the table but it is un-structured also. New and innovative big datatechnology is now allowing structured and un-structured data to be harvested, stored and used.

3.Veracity: -It is the quality of trustworthinessData. For ex. Use of GPS data. Often GPS data will "drift" off course as you pursue throughan urban area. Satellites signals are lost asthey bounce off tall buildings and otherstructures. When this happens location,data has to be fused with another data sourceto provide accurate data.

4.Velocity: - Velocity refers to the speed at which vast amounts of data are being generated, collected and analyzed. Every Day the no. of emails twitter messages, photos, video clips etc. increases at lighting speed around the world. Big data technology allows us now to analyze the data while it is being generated, without ever putting it into the database.

5.Value: - When we talk about value we talk about the worth of data being extracted. Having endless amount of data is one thing but till can be turned into value it is useless. The most important part of big data is to understand cost and benefits of collecting and analyzing and ensuring that ultimately data is reaped and can be monetized.Ignoring big data won't make it go away, and while it may not immediately kill your business it should not be ignored for very long. The results of big data can be measured making it easy to determine a return on investment. Big data is definitely a tool worth looking into it.

IVBENEFITS OF BIG DATA

The most important five benefits of big data are:

1 Cost saving-Some tools of big data like Hadoop and cloud-based analytics can bring cost benefits to the business when large amount of data is to be stored and those tools also help in identifying more efficient ways of doing business.

2 Time reduction- The high speed of tools helps in memory analytics which help in business analyzing and make quick decisions thus saves lot of time

3 New products development- By knowing customer needs through analytics you can create new products according to the wants of the customer.

4 Understand the market conditions- By analyzing big data you can understand current market situations.

5 Control online reputation- Big data tools can do sentiments analysis which help to get feedback and helps us to know who is saying what about the company.

V FUTURE RESEARCH AND TRENDS

Big data systems can be decomposedinto four sequential modules; namely,data generation, data acquisition, datastorage, and data analytics.Some possiblefuture trends of big data for modern industry include, but are not limited to:

• Novel techniques and improvementsfor big data analysis and mining.
• Cloud based solution related to big data storage and transmission.
• Big data solution focused on control and monitoring.
• Big-data based plant-wide optimizationand prognosis.
• Big data solution for supply chain and risk management systems.
• Big data solution for smart grids andclean power systems.

VIREVIEW OF LITERATURE

1Big Data for modern industry: Challenges and Trends by Shen Yin and Okyay Kaynak

- Published in the year 2015 by IEEE- This paper is all about the use of big data technology in the modern industry, what are the challenges coming and what are the new hopes, trends and applications of big data in modern industry.
- Big data is going to play an important role in the fourth industrial revolution.

2Big Data Analysis: Challenges and Solutions by Puneet Singh Duggal and Sanchita Paul

- Published in the year 2013 by Research gate – In this research paper the authors suggest various methods for catering to the problems in hand through Map Reduce framework over Hadoop Distributed File System (HDFS).
- Big Data Analysis-Heterogeneity, scale, timeliness, complexity, and privacy problems with Big Data hamper the progress at all phases of the process that can create value from data.
- Writers have also given and explained various architecture of Hadoop HDFS clusters, Map Reduce etc.

3 The Age of Big Data by Steve Lohr

- Published in the year 2012 by the New York Times. The author has written about how big data is used in various fields like social networks, industries, economic forecasting etc.

4 Data-intensive applications, challenges, techniques and technologies:A survey on Big Data by C.L Philip Chen and Chun-Yang Zhan

- Published in the year 2014 by Information Sciences. The author has written about principles used for designing the big data systems. It speaks about opportunities and the Challenges with big data.
- A large number of fields and sectors, ranging from economic and business activities to public administration, from national security to scientific researches in many areas, involve with Big Data problems.

5 Business Intelligence and Analytics: From Big Data to Big Impact by Hsinchun Chen, Roger H. L. Chiang and Veda C. Storey

- Published in the year 2014 by JSTOR. This paper is basically about how business intelligence and decision making in the business is affected and how the decisions are taken using big data.

6 Big data analytics by Philip Russom

- Published in the year 2011 by TWDI.org. In this the author has done the analysis of various industries and surveys and has given his analytical study about how big data is important and beneficial for industrial sector.

7 Challenges of Big Data Analysis by Jianquing Fan, Fang Han, Han Liu.

- Published on 5 Feb 2014 by National Science Review. This paper is all about various challenges industries face with big data.

S No	Author's Name	**Study Carried**	**Magazine Name**	**Results**	**Year**

1.	Shen Yin and Okyay Kaynak	Industrial view point Challenges and Trends	IEEE	Trends used in industries and challenges the are facing	2015
2.	Puneet Singh Duggal and Sanchita Paul	Big data Technology, map reduce and HDFS	Research Gate	What is big data? Architecture of Map Reduce and HDFS and their usage	2013
3.	Steve Lohr	Use of big data in various fields	New York Times	How big data is used in various fields.	2012
4.	C.L Philip Chen and Chun-Yang Zhang	Survey on usage of big data	Information Sciences	Result Of survey	2014
5.	Hsinchun Chen, Roger H. L. Chiang and Veda C. Storey	Business Intelligence and how Big Data Impacts business analytics	JSTOR	How the big data effects analysis process of industries.	2014
6.	Phillip Russom	Big data analytics	TWDI	It was a survey-based analysis	2011
7.	Jianquing Fan, Fang Han, Han Liu	Challenges Of big data analysis	Science Review	This paper was all about various challenges industries face with big data.	2014

VIIDiscussions

- This paper is a detail study and research of big data in industries.
- It has tried to explain about various aspects of big data, its character tics and other features.
- This also explained about 5 V's of big data, Challenges faced by industries with big data, how all those can be solved using various technologies.
- This paper explains about the impact of big data on industries and last what is the future of big data.

VIIIConclusion

This research paper concludes that big data is large volume of data generated from various sources and using what all techniques we can manage that big data. The various challenges like managing structured and unstructured big data which can be done using techniques like Hadoop and map reduce. Also, we have written about impacts which big data make on business intelligence(BI). Decision making is the most important part of business which is now completely undertaken by big data. Thus, we can conclude from this paper is that big data is the future of industries and no doubt fourth industrial revolution will be due to big data.

References

[1] Shen Yin and Okyay Kaynak," Big data for modern industry, challenges and trends" in IEEE in 2015.
[2] Puneet Singh Duggal and Sanchita Paul," Big data analysis, Challenges and solutions" in Research gate in 2013.
[3] Steve Lohr," Age of big data" in New York Times in 2012.
[4] C.L Philip Chen and Chun-Yang Zhang," Data-intensive applications, challenges, techniques and technologies: A survey on Big Data" in Information Sciences in 2014.
[5] Hsinchun Chen, Roger H. L. Chiang and Veda C. Storey," Business Intelligence and Analytics: From Big Data to Big Impact" in JSTOR in 2014.
[6] Phillip Russom," Big data Analytics" in TWDI in 2011.
[7] Jianquing Fan, Fang Han, Han Liu," Challenges of Big Data Analysis" in Science Review in 2014.
[8]https://digitaleconomyforum.org/chapter-2-how-your-digital-footprint-generates-big-data/
[9] https://www.google.co.in/benefits+of+ big data/