Modern Big Data Analysis for Business and It's Difficulties

Dharmendra Ambani[a], Dr.Kishor Atkotiya[b]

[a] Lecturer at Department of Computer Science, Harivandana College, Rajkot, India
[b] Associated Professor & Head of the Department of Statistics, Saurashtra University, Rajkot, India

Corresponding author: [a] dharmendraambani@gmail.com, [b]atkishor@yahoo.co.in

Abstract

In this paper we had tried to identify challenges to simplify Big Data for businesses. With the introduction of Big Data and performance business perspective for organization. The growth of big data and analytics has changed the entire appearance and item management. In our era day by day business will improve so raw data and other information like employee mobile data recorded for security reason, email, blog and many more information will store so day by day data will increases when business is going to growth. For example MNC e commerce store (amazon) require to store all information like warehouse, employee data, user data in this type of business comes big data and we need to handle them for improve business. The purpose of this study is to summaries applications of big data for business, importance, tools, and difficulties when handling big data and idea as well as competitive benefits of business. The determined part is to get insight from the raw data and makes to refine. Data are generally identified in the form of easy, semi complex and complex. We have used secondary data to identify various opportunities in the field of big data. The big data mostly comes unstructured and any sources require to analyze best business analytics technique..

Keywords: Big Data, Big Data Analysis, Business Analysis, Difficulties of Big Data, Literature review, Modern Large Data, Varieties of Big Data

1. Introduction

Data deviates with respect to time and its values are also changes from time to time. The bifurcation of massive data by using judgment decision process is known as Knowledge Discovery [2]. Just imagine the world without any data storage; then we not able to perform transaction, business, and many more difficulties are come so data is require. If data is very large big data again difficulties then we need to analysis to get meaningful information means extract usable information from unstructured data.

The various phases involved in big data analyzing [5]

1)   Data Acquisition

2)   Data Storage

3)   Data Analysis

The evolution of big data comes from evolution of growth of internet and latest technology. Having inbuilt storage competence with various method of data gathering, Very gigantic amount of data become easily accessible within the minutes. Because in every single second world store millions of record over the internet. In every micro second more data is being created over the internet as well as offline and need to be store and analyze in order to get or get extract values from stored data. So organization get proper data as possible from

large amount of stored data.

2. Literature Review

Big data is defined as the availability of a large volume of data that is difficult to store, segment, and emphasize in a broad way [2]. The primary suppliers of this type of data are larger e-commerce corporations and social media networks.

The purpose of retaining huge data is to improve the application of available knowledge, to reduce the time required to emphasize the same, and to facilitate the decision-making process. [7].

Techniques, treatment, process, and technology, as well as a method to evaluate data, are all referred to as business analytics. It also employs professionals who turn data into insights to assist all businesses in comprehending the company and making better decisions. Furthermore, not only does the business analytics concept aid in reporting, but it also aids in obtaining quick business responses or results. Corporate intelligence is defined as allowing business employees to readily report, search, query, and do predictive analytics using data stored in a data warehouse. The treatment of decision-making necessitates the use of a wide range of data. Many perspectives and predictive analytics and approaches from statistics, Machine Learning, and data mining are included in business analytics.

The definition states to the instruments and technology utilized to maintain big data. Over the past years, data generated by learning environment have started to tools and technology to handle them [1].

Big Data: The Characteristics Six Vs. [6]

The big data bunch of features commonly denoted to as the six Vs.: Volume, Variety, Velocity, Veracity, Variability and value.
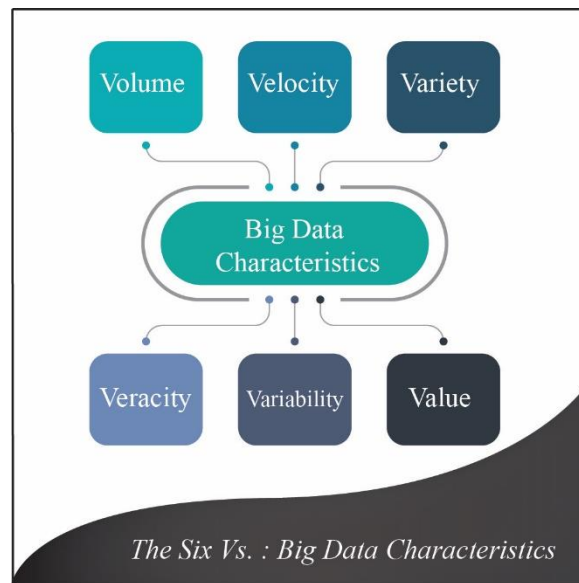

Fig. 1. Identified Characteristics

A. Volume

Now a days data will grow continue. We can say grow of data in large volume of big data. For example, before few years we have a small cellphone with approx. 256KB of storage space, now a days we have 256GB of storage space in our smartphone and still we have to require more space is called volume of data.

B. Variety

Data having many format. It can be structured, semi-structured or as unstructured. For example, mobile sales report, mobile call recorded report are structured. JSON, Web Page, XML are semi-structured. When we have collection of sentiment, smiley and natural language are as unstructured.

C. Velocity

Whenever require or say about speed of data known as velocity. For example. When we search particular file from smartphone and get within microsecond, second or minute called velocity main focus on streaming of data.

D. Veracity

When we retrieve meaningful information from data via any

Techniques of analysis, may be people or organization not trust. Means people having doubt in our analysis data called veracity.

E. Volatility

When we talk about time periods, for example, we might save a banking password that will expire in a few days and then be forgotten about.

F. Value

This feature is based on the relevance of value. For example, if a business focuses on profit, profit can be considered value; another company value is a focus on customer involvement, satisfaction, and relationship.

3. Comparative Analysis Of Education Data Mining Approaches

Larger firms, in particular, require a large amount of data to be processed in a timely and cost-effective manner, which necessitates extensive resources. In most circumstances, 50% of all businesses are unable to recognize and evaluate these data. Its adoption has also been cited as a stumbling block for businesses [6]. Analysis is the process of breaking down all forms of barriers into their component elements so that they can be thoroughly evaluated analytically. [4].

Rich consumer knowledge gained through the use of accessible data is required at the decision-making stage. It will take unique skills to effectively study this data, and it will have an impact on the decision-making process [9].

Stages for the process

• Judgments norms are tied on decision aspects which are social aspects, technology aspects and economic aspects.

• Candidate situations, There are various situations which organizations can select for big data e.g. big demand and carefully hopeful.

• Candidate Technology are the data repository, cloud analytics, rooted analytics and big data conceptualization.

• Technology evaluation pointers are global market size, enterprise acceptance ratio, entrance constrain and asset of industry.

• Technology Planning Implications, there are two variants of inferences are here which is technology development implication for situation of huge demand and technology planning inferences for situations cautiously optimistic.

Organizations are measuring for ideas in their products and routes to increase return of reserves. Product or process innovation approach requires extensive research for a predictive model that can ensure more growth and profit. [11].

4. Sources Of Big Data

Several sources of massive data peers are through social media such as Facebook, Google forums, and search engines.

These data are produced from online transactions, blog, emails, posts, videos, search queries where many people also generates data by means of Internet of Things applications and devices.

The data majorly comes from people's likes, their posts, and shared writings, uploaded contents etc.

5. Application

It includes the information generated by different types of applications. Some of the below form the require and

use application of big data [6].

1. *Health Data*: gather lots of data related to step count, heart fit, temperature etc.

2. ***Search Engine Data:*** Search engine retrieve all information just based on character and keyword.

3. ***Stock Exchange Data:*** It is a data regarding purchase and sells share and measuring profit they get.

4. ***Social Media:*** Social media app like Facebook manage images, post and video etc. While linked in manage data related to job and twitter, Instagram as well as snap chat manage large data.

6. Impotance Of Bigdata In Business

Difficulty in Business Intelligence (BI) refers to a technology-based method for investigating and analyzing data as well as presenting predictive information to assist scientists, business leaders, business managers, and other users in making more informed business decisions [10]. BI employs a variety of tools, applications, methods, and technology to assist businesses in gathering data from internal and external sources, preparing it for analysis, creating and executing data in order to gain effective ideation from the data, and producing reports and charts for data possible solutions so that the analysis results generated will assist businesses in making accurate and timely decisions.

The opportunities associated with data analysis in many firms have sparked a significant interest in business intelligence, which refers to strategies and technologies that aid in producing a better understanding of market power and making accurate decisions in real time [15].

7. How It Useful For Business (Process)

The data mining procedure cannot be completed quickly. Aspects such as ideation velocity, vigor, and ease of emphasis are required at this level.

Progress of data in rapports of scope of operators has augmented. We had only four data experts working at the time of Twitter's launch, but now the company employs thousands of people, primarily in its Hadoop cluster node data centers. Twitter boosted analytics power ahead of schedule, and if businesses don't do the same, these issues will become nightmares.

Big data is becoming increasingly crucial to decision-makers. Huge amounts of highly specific data from a variety of sources, including scanners, cell phones, devotion cards, credit cards, the web, and social media platforms, provide the opportunity to provide significant assistance to businesses. This is only feasible if the data can be analyzed effectively to reveal useful insights, allowing decision-makers to take advantage of the opportunities presented by the abundance of historical and real-time data generated by supply chains, manufacturing processes, customer behaviors, and other sources.

8. Innovation In Big Data Technology Till 2020

Before 2005 Google launched the map reduce technique for structure the data. In 2005 yahoo also introduce the Hadoop for map reduce for handle upcoming big data. So that searching is easy and very fast retrival.in current era many business are using Hadoop to handle data in HDFS (Hadoop distributed file system).

The below tools which are used in Big data are [5].

Fig 2 Tools for Big Data

According to many literature survey from various journal and books, the data created in 2011 whole year is same data created 1 days in 2020. So many businesses shifted in big data. Now current era Multi National Company are trying to learn big data and preparing themselves to shift big data. Now a days no one is use hadoop, maximum company shifting into the spark the reason behind that spark is 99 times quicker than map reduce and spark code is written in Java, python and R Programming.

9. Results And Discussions

Every evolution comes with different hitches and so is the big data in business. As turf linking utilization of big data are multiplying, the difficulties are also stands up [5]. Many difficulties while managing big data are as below:

Storage Problems, Sentiment analysis [8]: Predicting customer churn, Security, Time Reduction, Cost Bargaining, Support in house business choices, identifying new Big Data-Based analysis, safety and privacy, mishandling of Big Data, Data managing, Data Storage and Analysis, largely business model entails of a series of phases to be executed in sequence within deadlines, knowledge detection and computational complications, scalability and visualization of data, information safety, risk management and fraud detection [13], absence of modest indulgence.

1. Data growth problems

One of the most well-known challenges of Big Data is storing such a large amount of data in an appropriate manner. The amount of data kept in corporate data centers is rapidly increasing. It becomes increasingly challenging to preserve large data sets as they grow in size over time.

2. Uncertainty for selecting tool for analyzing

Selecting the tools for data analysis is very crucial as it plays vital to get desired and optimum output with respect to time. To identify optimum tool for data analysis needs various aspects to check.

These queries make barrier for businesses and also incapable of getting the solutions. At the end very poor choices and choosing a useless and improper technology. As a result, money, time, efforts and work hours are lost.

3. Lack of data specialists

To process huge amount of data with latest technologies needs skilled data specialists who can analyze the data with help of latest tools and making output worth. Unavailability of data specialists is due to the data handling tools has changed rapidly with respect to time.

4. Fortifying data

Safeguarding huge amount of data is quite difficult. Often companies are so engaged in accepting, loading and processing data sets which improves data security for later stages. But, this is not an innovative step as unsecured data sources can develop breeding stages for malicious hackers.

5. Integrating data from a various sources

Data in business originates from a various sources, such as social media pages, websites, ERP applications, customer logs, financial reports, blogs, e-mails, presentations and reports created by employees. Conjoining all these data to create reports is a tough task.

6. Storage

The regular bases capacity of external hard disk is terabytes, the daily volume of data produced through online in order of Exabyte and in future day as well as future data growing is rapidly boosting. The ordinary RDBMS tools will be incapable to store, handle and analyze of big data. We are not able to use SQL based query for handle big data.

7. Analysis

The data are generated mostly in social media inform of say from internet. Data are comes in variety means structure, semi-structured and unstructured. Data also spilt in small pieces so analysis of big data is again difficulties.

10. Limitation Of Big Data

A recent hiring policy that no correlation between academic achievement and work performance means big

data can be misused and unstructured. When data transmission can occur confusion and chances the complexity of data handling and management. User level outcome cannot be present directly. And user level data is subject to more noise and May not transferable.

11 Discussions

The data generation is increasing day by day. Analyzing these data is really a difficult assignment. It should combined with business's structure, also the companies do have their identified and huge corporates. Countries in the world, IT companies, E-Commerce organizations and the relevant departments have started encountering on big data. Big corporates are betting and grouping with data processes and combining advanced technologies

However, a precise understanding of analogous business costs and benefits, as well as a precise understanding of the impact on businesses, was not establish, while impacts such as reduction of costs, improved step to decision-making process, increase customer satisfaction, or operating logicality are known. Hence, our findings indicate that there is not the concept for big data and business analysis.

As a result, this paper represents a first step toward a thorough understanding by highlighting the essential antecedents clustered in a general and technical overview of BIG DATA as well as the numerous obstacles encountered when dealing with Big Data. Big data is defined as a large amount of data that is beyond the analyzing capabilities of traditional data set systems to arrange and analyse the data within a specified time period. Advanced methods are required for analyzing and storing enormous amounts of data at any location at any time. This study investigates the phases of data size, the definition of big data and its issues, and solutions for the processing of big data in the actual world, including an explanation of how big data is processed using the Hadoop Map-reduce V2 YARN framework

References

[1]   Dursun Delen & Sudha Ram, "Research challenges and opportunities in business analytics", Journal of Business Analytics, Vol.1, No. 1, pp. 2-12, August 2018.

[2]   Pradeep S and Jagadish S Kallimani , The Different Tools and Technique to Handle Challenges in Big Data , ICOEI, 2019.

[3]   J. Alberto Espinosa, Stephen Kaisler, Frank Armour, William H. Money, "Big Data Redux: New Issues and Challenges Moving Forward", Hawaii International Conference on System Sciences, 2019.

[4]   Azzah Al Ghamdi, Prof. Thomas Thomson, "Big Data Storage and its Future", IEEE , 2018.

[5]   Deepali Arora, Piyush Malik, "Analytics: Key to go from generating big data to deriving business value", IEEE First International Conference on Big Data Computing Service and Applications, 2015.

[6]   Jafar Raza Alam, Asma Sajid, Ramzan Talib, Muneeb Niaz, "A Review on the Role of Big Data in Business", Vol. 3, No. 4, pg.446 – 453, April 2014.

[7]   Mudasir Ahmad Wani, "Big Data: Issues, Challenges, and Techniques in Business Intelligence", Big Data Analytics, Advances in Intelligent Systems and Computing. Springer Singapore, January 2018.

[8]   Rashi Chaudhary, Prakhar Pandey, JaJaj Ranjan Pandey, "Business model innovation through Big Data", IEEE, 2015.

[9]   D. P. Acharjya, Kauser Ahmed P, "A Survey on Big Data Analytics: Challenges, Open Research Issues and Tools", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 2, pp. 511-518, 2016.

[10]   Nada Elgendy, Ahmed Elragal, "Big Data Analytics: A Literature Review Paper", Springer International Publishing Switzerland 2014, pp. 214–227, August 2014.

[11]   Shankar Ganes h Manikandan, Siddart h Ravi, "Big Data Analysis using Apache Hadoop", IEEE, 2014.

[12]   Hiba Alsghaier, Mohammed Akour, Issa Shehabat, Samah Aldiabat, "The Importance of Big Data Analytics in Business: A Case Study", pp. 111-115, 2017.

[13]   X. Wu, X. Zhu, G. Q. Wu, et al., "Data mining with big data", IEEE Trans. on Knowledge and Data Engineering, Vol. 26, No. 1, pp. 97 – 107, January 2012.

[14]   S. Salloum, R. Dautov, X. Chen, P. X. Peng, and J. Z. Huang, "Big data analytics on apache spark", International Journal of Data Science and Analytics, Vol. 1, No. 3, pp. 145–164, 2016.

[15] K. Sravanthi, T.S. Reddy, "Applications of Big data in Various Fields", International Journal of Computer Science and Information Technologies, Vol. 6, No. 5, pp.4629- 4632, 2015.

[16] Faroukhi, A.Z., El Alaoui, I., Gahi, Y. et al., "Big data monetization throughout Big Data Value Chain: a comprehensive review", J Big Data 7, 3 (2020).

[17] Jain, P., Gyanchandani, M. & Khare, N., "Big data privacy: a technological perspective and review", J Big Data 3, 25 (2016).

[18] Hasan, M.M., Popp, J. & Oláh, J., "Current landscape and influence of big data on finance". J Big Data 7, 21 (2020).