Impact of Big Data on Organizational Growth

Krishna[1], Kavitha.M[2]
Assistant professor[1, 2]
Department of Computer Science
Acharya Bangalore B-School, Andrahalli Bangalore, India

Abstract:
Big Data impacts in the financial, commercial, educational, historical, health, scientific sectors, in which data is not utilized in an efficient manner to grow an organization due to lack of knowledge on the Big Data .Proper analysis of historical data can improve the quality of an organization. In fact, 70% of companies fear that they risk becoming irrelevant or uncompetitive if they don't embrace it. Companies struggle to make use of the data, hereby some of the computer tools is been introduced to maintain the historical data in efficient and effective manner. And to create awareness among an individual basic level of knowledge to be provided in an educational level.

1. INTRODUCTION

Problems facing on big data.

1.   Data Accuracy.
2.   Growth of technologies.
3.   Lack of Skill.
4.   Server maintenance.
5.   Data Storage Compatibility.

•   Accuracy is the measure of whether a part of information reflects reality; it is also the major issue in the growth of organization based on big data analysis. Inaccuracy costs the organization in wasting operating costs, increased capital expenditure, reduced customer service levels and an overall it affects on profits. We do have tools which helps us to organize the historical data which reduces the inaccuracy effect on data.

•   Technologies are fast growing, which can also be the major issue to maintain the historical data in efficient manner without compatibility. Updating with upgraded latest technologies with whole data is risk factor for company.

•   According to the survey on big data it was a shocking result that only 65% of Industrial people know about the impact of data and its analysis on their business growth. Lack of knowledge about historical data maintenance. When it reaches the saturation point it will be problem to the people those who are running the organization without the knowledge of big data.

•   Server maintenance is also one of the issues because for a small organizations and startups cannot afford the maintenance cost, trained manpower, tool cost etc. but its mandatory to maintain data with secure for analysis process to growth of organization. It is the biggest challenge to the small organizations According   to the survey only 34% organization big data initiative is successful. This we can conclude that there is a huge gap in use of theoretical knowledge putting it into practical way.

Historical Data collection based on 2017 survey, Big data and data analysis is more important to build the smart city and growth of beautiful Global condition, in future if we share the proper knowledge and scenario of big data maintenance and data analysis, we can come across with the global problems like Global financial and economic condition, and also job opportunity increases in global market on data scientist, new tools will introduce, with a chain link global development is possible only with awareness on the importance of big data and data analysis.

Illustrating on "Why big data has become so important?"

•   Most data collected now   is unstructured and requires different storage and processing than that found in traditional relational databases.

•   Available computational power is sky-rocketing, meaning there are more opportunities to process big data.

•   The Internet has democratized data, steadily increasing the data available while also producing more and more raw data. Data in its raw form has no value. Data needs to be processed in order to be of valuable. However, herein lays the inherent problem of big data. Is processing data from native object format to a usable insight worth the massive capital cost of doing so? Or is there just too much data with unknown values to justify the gamble of processing it with big data tools? Most of us would agree that being able to predict the weather would have value; the question is whether that value could outweigh the costs of

crunching all the real-time data into a weather report that could be counted on. Big data refers to a process that is used when traditional data mining and handling techniques cannot uncover the insights and meaning of the underlying data. Data that is unstructured or time sensitive or simply very large cannot be processed by relational database engines. This type of data requires a different processing approach called big data, which uses massive parallelism on readily-available hardware. According to the IBM investor briefing, it's a just beginning for us as shown in statistical analysis uses of data maintenance require much more awareness among us . when we talk about data analysis it is the big trend of enhance the market and marketing in a unique way like big data is used to better understand customers and their behaviours and preferences. Companies are keen to expand their traditional data sets with social media data, browser logs as well as text analytics and sensor data to get a more complete picture of their customers. After that company starts to enhance their organization and send promotional add's for publicity like face book contains the adds depends on search history. IRCTS login page contains the adds, Some University and educational sites contains adds Suggesting products, courses, items purely based on search history. Like this the historical data has to be analyzed based on customers search history and their interest. This is the new technique and use of big data which is been used currently by the Google, yahoo, flipchart, face book, Amazon such companies are doing to grow in a unique way. ETL tools combine three important functions (extract, transform, and load) required to get data from one big data environment and put it into another data environment. Traditionally, ETL has been used with batch processing in data warehouse environments. Data warehouses provide business users with a way to consolidate information to analyze and report on data relevant to their business focus. ETL tools are used to transform data into the format required by data warehouses. The transformation is actually done in an intermediate location before the data is loaded into the data warehouse. Many software vendors, including IBM, Informatica, Pervasive, Talend, and Pentaho, provide ETL software tools.

**ETL provides the underlying infrastructure for integration by performing three important functions:**

- Extract: Read data from the source database.

- **Transform:** Convert the format of the extracted data so that it conforms to the requirements of the target database. Transformation is done by using rules or merging data with other data.

- **Load:** Write data to the target database.
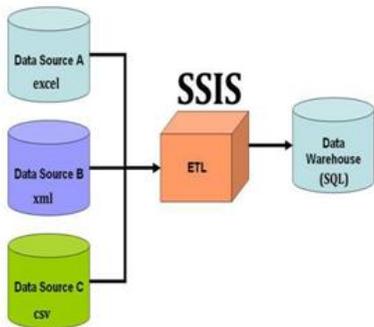
However, ETL is evolving to support integration across much more than traditional data warehouses. ETL can support integration across transactional systems, operational data stores,

BI platforms, MDM hubs, the cloud, and Hadoop platforms. ETL software vendors are extending their solutions to provide big data extraction, transformation, and loading between Hadoop and traditional data management platforms. ETL and software tools for other data integration processes like data cleansing, profiling, and auditing all work on different aspects of the data to ensure that the data will be deemed trustworthy. ETL tools integrate with data quality tools, and many incorporate tools for data cleansing, data mapping, and identifying data lineage. With ETL, you only extract the data you will need for the integration. ETL tools are needed for the loading and conversion of structured and unstructured data into Hadoop. Advanced ETL tools can read and write multiple files in parallel from and to Hadoop to simplify how data is merged into a common transformation process. Some solutions incorporate libraries of prebuilt ETL transformations for both the transaction and interaction data that run on Hadoop or a traditional grid infrastructure. Data transformation is the process of changing the format of data so that it can be used by different applications. This may mean a change from the format the data is stored in into the format needed by the application that will use the data. This process also includes mapping instructions so that applications are told how to get the data they need to process. The process of data transformation is made far more complex because of the staggering growth in the amount of unstructured data. A business application such as a customer relationship management has specific requirements for how data should be stored. The data is likely to be structured in the organized rows and columns of a relational database. Data is semi-structured or unstructured if it does not follow rigid format requirements. The information contained in an e-mail message is considered unstructured, for example. Some of a company's most important information is in unstructured and semi-structured forms such as documents, e-mail messages, complex messaging formats, customer support interactions, transactions, and information coming from packaged applications like ERP and CRM.
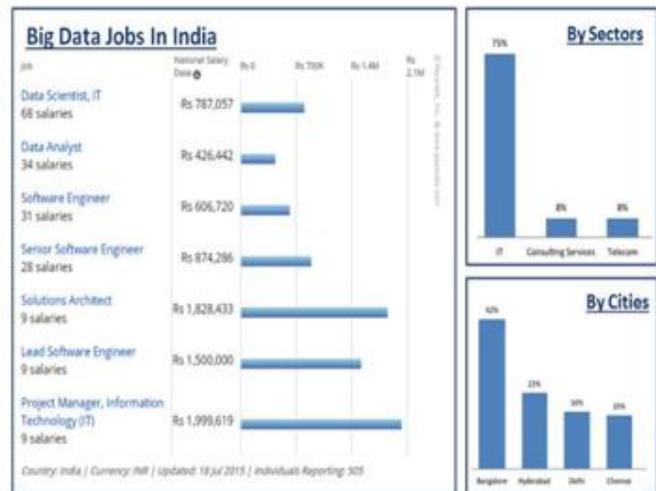
Big data awareness can erase a term called unemployment in India, with high paid designations such as Data Scientists, Data Analyst, Software Engineer, senior engineer's etc., as shown in the above statistical analysis. And it is also been noted that 40% of IT Industry and other sectors has been covered with data analysis job, which has a huge demand in current market.
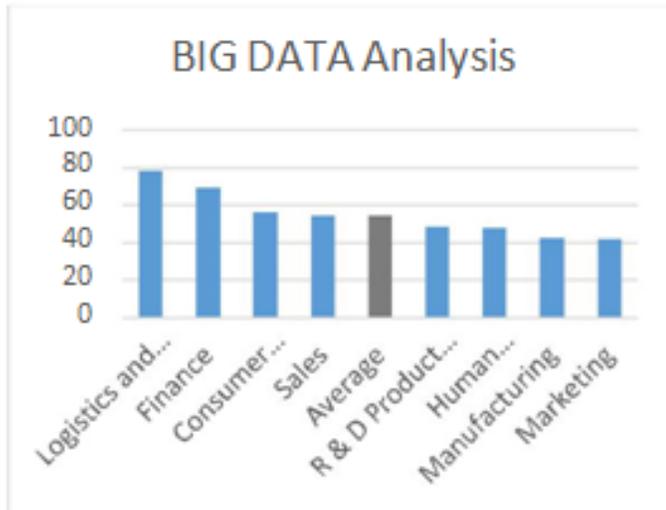
BIG DATA Analysis

Banks are betting on big data analytics and real time execution to better engage with customers.

• 81% of large banks put customer centricity as a top priority.

• 1 out of every 2 executives believes they do not have mature capabilities to support their customer strategies.

• Technology plays a key role in a customer focused strategy such as

• 59% of Real-time processing of data analytics

• 59% of predictive analytics.

• 54% of Data Visualization.

II. CONCLUSION:

Data to be maintained and analyzed in a structured manner in-order to improvise the organization in which Data represents both the positive and negatives. Basic awareness of the Big-Data to be known by every individual, where government can include it in the basic education level. Utilization of the data can be efficient if it is used in an effective manner. Unnecessary usage memory can be avoided by applying the business logic to the data while storing in a compressed memory.

III. REFERENCES:

[1]. Mining data Streams with Concept Drift in Massive Online Analysis Framework is published in WSEAS transactions on computers. Article is in the press. ISSN/E-ISSN: 1109-2750 /2224-2872, Volume 15, March 2016, Art #14, pp.133-142. The journal is Scopus (Elsevier. B. V.) Indexed.

[2]. Challenges and Issues in Handling Big Data published in International Journal of Innovations and advancement in Computer Science. Vol 4, pp: 620-625, Special Issue, March 2015, ISSN 2347-8616.