Comparative Analysis on Basic Big Data Analytics: Challenges, Issues, Trends and Tools

[1]A.Ragavisri, [2]Dr.P.Periyasamy
[1]Research Scholar, [2]Associate Professor
[1]Department of Computer Science,
[1]Sree Saraswathi Thyagaraja College (Autonomous), Pollachi, India

Abstract: Essentially, the term Big Data relates to the datasets which have a combination of velocity, volume, and variety that can't be handles by a simple traditional data. In order to improve the availability of data, it handles in three setouts like Structure, Unstructured and Semi formats. The main objective of Big Data analytics is to obtain very able values, bear all conclusion and holdup decision making. This paper supplies an intense analysis of Big Data Challenges along with different types of Tools, Trends and Issues accompanied by prominent attributes which have an expose research issues.

Watchwords: **Big Data Analysis, Applications, Challenges, Technologies, Hadoop, Map reduce, cloud computing, Benchmark, Mobile networking, IoT**

I. INTRODUCTION

Big Data handle the vast majority of data because of working in a digital environment, people and the devices produce a static data which is created while brooking plentiful facts provided where Big Data provides a potential revolutionize art of management to take exact conclusion on time which results in performing with massive data sets to analyze calibration all to make patterns, trends and an organization from unstructured data into structured data to bring out the perfect solution. Now a day's Big Data is grooming extremely due to the valuable information's scattered in a business organization support where the companies start adopting the intention of building workable information from Big Data.

While using this platform there is a worth full solution for various problems hence it is a new technology it plays an important role in the upcoming applications/projects. From this survey, it is clear that Big Data is a robust impulse to the next upcoming generation of information technology industries.

The significant of big data does not loop around how much data a company has but how a company implements the collected data. Every company uses data in its own way; the more qualified company uses its data, the more probable it has to grow. The company can extract data from any source and survey it to notice answers which will enable:

A. Cost Savings: Some tools of Big Data like Hadoop and Cloud-Based Analytics can bring cost advantages to business when large amounts of data are to be stored and these tools also assist to pinpoint more logical ways of performing business.

B. Time Reductions: The high speed of tools like Hadoop and in-memory analytics can easily identify new sources of data which helps businesses analyzing data immediately and make quick decisions based on the learning.

C. New Product Development: By knowing the trends of customer needs and satisfaction through analytics you can create products according to the wants of customers.

D. Understand the market conditions: By analyzing big data you can get a better understanding of current market conditions. For reference, it can analyze customers securable behaviors, a company can pick out the products that are sold the most and build products according to this trend. As it can obtain a leading challenger.

E.Control online reputation: Big data tools can do sentiment analysis. Hence, you can get a suggestion about who is saying found in your company. If you want to detect and enhance the online existence of your business, then big data tools can support in all this.

The basic target of this paper is to travel the possible influence of Big Data Challenges, about its issues, trends and diverse tools peer with it. As an upshot of this study furnishes a stage to investigate Big Data at deviant stages. Moreover, we footing open research issues in Big Data. This paper is divided into further sections.

II. CHALLENGES FACED BY BIG DATA

A data volume is ready to survive the growth and perform their work with massive data. Yet the organization has to perform the processes that satisfy the needs of consumers, products as well as services too. Among the 85% of companies using Big Data, only 37% have been successful in data-driven intuitions. And only a 10% increase in the accessibility of the data can lead to an increase of $65Mn in the net income of a company.

Since Big Data offers tons of welfare along with its own set of issues. As it is a new set of complex technologies, runs still in the growing stages of development and evolution.

More commonly faced issues include deficient knowledge about the technologies involved, data privacy, and deficient analytical capabilities of organizations. A lot of enterprises also face the same issues like lack of skills for handling Big Data technologies. Not everyone is exactly trained to work under Big Data, which result in major risk.
Actually, this is not the only challenge or problem though there are many challenges too, in that some are identified after organizations begin to move into the Big Data space and some on the surface of a roadmap.

Fig 1: Major Challenges

A. Handling a Large Amount of Data

Everyone knows that there is a huge bang of increase in data that are available when we compare it with last years along with today's data. For clarification just assume that a wonderful residency opened in Italy the last month.
It is clear that there is a massive amount of data to store, compute and additionally retrieve. The significant challenges are not between availability, but the management can do that. Now relating it with entropy the data may be around 7.6 times the extent in the middle of the earth and the moon might be 2022, which is assuredly a challenge.
Accompanying the rise of unstructured data leads to an increase in data formats for instance video, audio, social media, smart device data etc.
Some of the latest procedures to manage this type of data results in the hybrid of relational databases blend with NoSQL databases. By way of illustration of this is MongoDB, whichever built-in part of MEAN stack. Other distributed computing systems like Hadoop to assist Big Data volumes.
Netflix is to satisfy the cascade based on Node.js. With the amount of content and the complex formats available on the policy, in order to handle that stack helps in storage and retrieval of the data. Others use MEAN stack, and with a relational database model, they could, in fact, manage the data.

B. Real-time can be Complex

Considering the data, it is clear that there is nothing to limit the standing available data at similar ejection. Leads to update the data at every second along with organizations that required being aware. Specimen about Retail Company which analyze about customer bearing, real-time data begin from owned purchases to help. Big Data Analysis guides us in veracity and velocity brings ETL engines, computation engines, visualization, frameworks, and other important inputs.
In order to update the data, the business maintains the data itself with the combination of standing as well as available data. This leads to focus on better perceptions to enlarge the decision-making capacities.
Although every organization does not handle the real-time data due to the lack of updating with their natural tools, techniques that are available. At present, only a few definitive tools which result in a lack of mandatory enlightened.

C. Data Security

The massive organization lacks in handling the enormous amount of data as well as security which result in larger dare for data related issues which are combined with enterprise leads to an enormous range of sources may be entrusted for securing along with organizational merits.
In order to decrease these issues, strangers should deal with a variety of data to handle the availability of data which results in conflicts data along with their respective solutions. Basically, when we are in the business background it is must to manage, regulate according to the valid/interest with positive directions of answering.
Performing with a huge source of data as well as security trends to the unknown selection of data which is agreement along with the security of the organization leads to hacker's performance.
Hence it is just to import the data security in order to perform effective operations for secure data storage, collections, and retrieval of data.

D. Shortage of Skilled People

At present, there is a lack of experienced people working under the skilled Big Data. Where these lines are quoted by many enterprises because of their utilization as well as working in effective data analysis systems. But there is the biggest lack to build certified Data Scientist.
Repeatedly training people at the initial state is also costly while dealing with new technologies. For this reason many are performing under automation solutions likes Machine learning and artificial intelligence to develop an awareness which also leads to the well-trained people or expandable skill developers.

The Tremendous Seven: big data challenges and ways to solve them

1) Insufficient understanding and acceptance of big data
2) A confusing variety of big data technologies
3) Paying loads of money
4) The complexity of managing data quality
5) Dangerous big data security holes
6) A sensitive activity of translate big data into valuable insights
7) Troubles of upscaling

III. ISSUES FOUND IN BIG DATA

The dominant Utility of the Big Data leads to point the black hat for their personal devil purposes. It reveals that a known pair of issues with big data. In this, we guide you to handle that type of issues and direct you to avoid the villain those who procuring the accessible big data and triggers all kind of perversity.

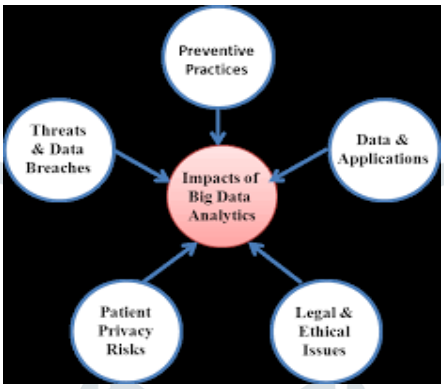

Fig 2: Common Bid data issue

Following are some of the Big Data issues which you can control it:

A. PRIVACY BREACH

Privacy breach plays vital issues in Big Data. It response like transferring of others private information to access without their notice it may be either purposely/wrongly. It only performs when there is a lack of security measures. In order to save the data from the hackers, they should handle proper protocols.

The act of privacy breach is to infuse branded anti-malware software which prefers the standard security to the data storage system also from the collected data.

B. IT'D BE IMPOSSIBLE TO BE ANONYMOUS

Now it is possible to radiate discrete with unknown data in public datasets either by performing by a computer, or an internet connection to probably select the unknown data. This task is not easy but still the opportunity for identifying the larceny. This must be practiced by massive people to process the impossible private presence of online data.

C. ANALYTICS ISN'T 100% ACCURATE

In this research, there is a thing called the margin of error, miscalculations, exceptions, and another point that causes a relatively negative amount of error. While in an analysis of big data it is not exposed, mainly because of the action for the huge amount of data. As it is very critical to handle the analysis manually, your best task is to guard your analytics which is not available heavily as an inaccurate data to use a trusted data analysis tool that covenant the highest level of precision.

D. E-DISCOVERY PROBLEMS

The word E-discovery means finding the electronic data to satisfy the evidence in a legal proceeding. Sometimes in Court or even in the Government can preside E-discovery in the format of hacking leads to support the search of critical evidence. Electronic evidence is much easier to search and collect. Nevertheless, now a day it is more difficult to search for electronic evidence because there are tons and tons of data to obey the legal limits. Furthermore, at present E-discovery is more expensive than ever.

E. YOUR PERSONAL INFORMATION IS AT RISK

Due to Data breaches the revelation of customer information to people who do not have access to sensitive information. With every breach there is some major critical process like Social Security numbers, home addresses, email addresses, contact information, debit, and credit card numbers, your personal information are at risk.

When performing in a business which contains profusion data so make clear that lead to privacy and set the security for their own personal and diplomatic information.

F. DISCRIMINATION

On utilizing the big data the discrimination is more prevalent that is capable to access electronic information (like habitual online activities, online preferences) that directs us to negativity which affects the opportunity of a person so stay secure on a loan and without that person's potential to justify such type of information are said to be increasingly unfair. It is good for decisions plainly out of electronic data, especially those that have a gigantic smash on someone.

G. LEGAL PROTECTIONS ARE ONLY A FEW FOR INVOLVED INDIVIDUALS

There is only minimum valid security that secures privacy on the analysis of big data. The Government has already conceded risks to privacy, but there is a lack of effort taken to address with the help of virtue about passing laws. We have to wait for our representatives to do their responsibilities, but it is always clear to express how vital this is.

H. PATENT AND COPYRIGHT MAY BE IRRELEVANT IN THE FUTURE

Patent and copyright may be unrelated in the future because it will be more difficult to prove whether vaguely is unique or if it damages copyright because there are tons to verify them against some data storehouses and some of these data storehouses will be unrevealed or preserved with thick implicit walls that block the access.

I. INTERPRETATIONS CAN BECOME AN ETHICAL ISSUE

The wider usage of analyzing big data is to help in upcoming decisions. Although relying only on electronic data, without many disturbances on its bang on people or on the environment it can become ethical issues. We must learn that big data is here to help us and not to order us. Our resolution should always be based on how it will influence everyone involved, and not exclusively on numbers on a spreadsheet.

J. BIG DATA IS HERE TO STAY

Repeatedly, big data has always been everywhere along with us. It has an advanced as well as enlightenment, so it is still mature. So now we have learned to value it, big data will not be extracted for permit anymore because of risks and issues involved, but it is not exact. The value of big data may be in your hand, and make the best out of it. But make clear about your control on handling your big data as private and secure.

IV. TRENDS USED IN BIG DATA

The quantity of data generated at present from every industrial domain also refers as big data which may be huge, enclosing data gathering, data analysis, and data implementation process. For a year backs, big data analytics trends are dynamic from a departmental approach to business-driven data approach hold agile technologies and an increased focus on advanced analytics. The business enterprises need to execute the right data-driven and big data analytics trends should remain leading in the competition.

Fig 3: Big data Trends

Once big data was fundamentally established by big businesses, who could offer the technology and channels used to collect and analyze the information. Now the range of big data is changed to lead the business enterprises of large and small depended on big data for intelligent business insights which results in evolving big data at an incredibly high rate. The exact sample of the growth is big data in the cloud which also helps even small businesses to proceeds the advantage of the fresh technology trends.
The incessant brook of information is precious to the business, but it can also be a challenge to draw applicable awareness from a large data pool of data which may be unstructured. Even with these roadblocks, there's no denying the fact that big data offers enormous business opportunities for growth. The following are some of the trends that will be the mass of the technologies in the world in 2019 and beyond.

A. Fast Growing IoT Networks

As the Internet of Things (IoT) will be the trend now a day, which will originate more than $350 billion per annum by 2020. Affording to the current industry trends and research reports, the global IoT market will rise at a CAGR of 29.5%. Business houses will depend on more data points to gather information for largely detailed business imaginations.

B. Predictive Analytics

Predictive Analytics suggest a customized perception that promotes organizations to originate recent customer responses or purchases and promote cross-sell occasion. Predictive Analytics assist technology to merge into multiple domains like finance, healthcare, automotive, aerospace, retailing, hospitality, pharmaceuticals, and manufacturing industries.

C. Dark Data

Dark data technology is the digital information that is currently not in use for business analysis. This data is gained through various computer network operations which are not used in a manner to derive imagination or for decision-making. As analytics and data become daily characteristic of organizations, there is a massive need to understand that any data left unknown opportunity to lose and may be directed to the potential security risk.

D. CDOs in Demand

The contour of the Chief Data Officer (CDO) has enlarged and human resource personals are explored for professionals who can fill this trendy job role. Regardless in demand, CDO is still a comparatively new concept to many companies. Organizations have a register that they need to engage a CDO, so if you are a data leader managing enterprise-wide data cleaning, analysis, visualization and studying intelligent insights, CDO may be the work profile for you.

E. Quantum Computing

Tech giants like IBM, Microsoft, and Google, as well as Intel, rush against each other to work for accurate in an effort to build the first quantum computer. Quantum Computing permits defective data encryption, weather prediction, solving complex medical problems, real conversations, and better financial modeling to build organizations develop quantum computing components, algorithms, applications and software tools on qubit or quantum bit cloud services

F. Open Source

In 2020 we will observe plenty of free data and software tools to become obtainable in the cloud. Even small organizations and start-ups similar will benefit the most of this data trend in 2020. Open source analytical languages like R, a GNU project associated with statistical computing and graphics has seen a cosmic assuming credit to the open source wave.

G. Edge Computing

Edge Computing has introduced in the technological space flowing network showing for quite a while now. All credits to edge computing that data analytics is half needful on the network bandwidth to save data provincial close to the data source. Edge Computing builds data to handle and store away from the silo process closer to the end users with action taking place either in the device itself or in the cloud surface or in the border data center.

V. TOOLS IN BIG DATA

At present, the market is submerged with an array of Big Data tools. It guides cost efficiency, better time management into the data analytical tasks. The following lists are some of the best big data tools with their key Attributes and logo.

Fig 4: Collections of Tools in Big Data

A. Hadoop:

The Apache Hadoop software library referred as big data framework. It permits distributed processing of large data sets beyond clusters of computers. It is aimed to scale up from single servers to multiple machines.

Attributes:

- Authentication improvements when using HTTP proxy server and allow Specification for Hadoop Compatible File system effort, also support for POSIX-style file system extended attributes
- It offers a robust ecosystem that is well fixed to meet the analytical needs of a developer
- It brings Flexibility and allows faster data in Data Processing

B. HPCC:

HPCC is a big data tool originates by LexisNexis Risk Solution. It conveys on a single platform, a single architecture and a single programming language for data processing.

Attributes:

- Highly efficient accomplish big data tasks with far less code and offers high redundancy and availability
- It can be used for both complex data processing on a Thor cluster
- Graphical IDE for simplifies development, testing and debugging and it automatically optimizes code for parallel processing
- Provide enhance scalability and performance and also ECL code compiles into optimized C++, and it can also extend using C++ libraries

C. Storm:

The storm is a free and open source big data computation system. It process distributed real-time, fault-tolerant processing system with real-time computation potentiality.

Attributes:

- It benchmarked as processing one million 100 byte messages per second per node and uses parallel calculations that run across a cluster of machines
- It automatically restarts in a case node dies so the worker will be restarted on another node
- Storm guarantees that each unit of data will be processed at least once or exactly once, the deployed Storm is surely the easiest tool for Big Data analysis

D. Qubole:

Qubole Data is an independent Big Data management platform. It is a self-managed, self-optimizing tool which allows the data team to focus on business outcomes.

Attributes:

- Single Platform for every use case. It is an open-source engine, optimized for the Cloud and Comprehensive Security, Governance, and Compliance.
- Provides actionable Alerts, Insights, and Recommendations to optimize reliability, performance, and costs and also automatically enacts policies to avoid performing repetitive manual actions

E. Cassandra:

The Apache Cassandra database is widely used today to provide successful management of large amounts of data.

Attributes:

- Support for replicating across multiple data centers by providing lower latency for users was data is automatically replicated to multiple nodes for fault-tolerance.
- It is most suitable for applications that can't afford to lose data, even when an entire data center is down, it offers support contracts and services are available from third parties.

F. Statwing:

Statwing is an easy-to-use statistical tool. It was made for big data analysts. In modern interface chooses statistical tests automatically.

Attributes:

- Explore any data in seconds and helps to clean data, explore relationships, and create charts in minutes
- It allows creating histograms, scatters plots, heat maps, and bar charts that transport to Excel or PowerPoint it also interpret outcome in plain English, so analysts are strangers with statistical analysis

G. CouchDB:

CouchDB stores data in JSON documents that can be accessed by web or query using JavaScript. It supports distributed scaling with fault-tolerant storage. It enables gaining data by defining the Couch Replication Protocol.

Attributes:

- CouchDB is a single-node database that works like any other database also allows running a single logical database server on any number of servers
- It makes use of the ubiquitous HTTP protocol and JSON data format to easy replication of a database across multiple server instances
- Easy interface for document insertion, updates, retrieval, and deletion and JSON-based document format can be translatable across different languages.

H. Pentaho:

Pentaho provides big data tools to withdraw prepare and blend data. It offers visualizations and analytics that convert the way to run any business. This Big data tool allows turning big data into big imaginations.

Attributes:

- Data access and integration for effective data visualization to empower users to architect big data at the source and streams them for accurate analytics.
- Seamlessly switch or combine data processing within-cluster execution to get maximum processing that allows checking data with easy access to analytics, including charts, visualizations, and reporting
- Supports a wide spectrum of big data sources by offering unique capabilities

I. Flink:

Apache Flink deals with an open-source stream processing Big Data tool. It is easy to distribute, handle high-performing, and found to be always-available also perform accurate data streaming applications.

Attributes:

- Provides results that are accurate, even for out-of-order or late-arriving data and it is stateful and fault-tolerant and can recover from failures.
- It can perform at a large scale, running on thousands of nodes that has good throughput and latency characteristics
- This big data tool supports stream processing and windowing with event time semantics to supports flexible windowing based on time, count, or sessions to data-driven windows and a wide scale of bond to third-parties systems for data sources and descend.

J. Cloudera:

Cloudera is the quick, easiest and highly secure modern big data platform. It allows anyone to get any data beyond any environment within a single or scalable platform.

Attributes:

- High-performance analytics to offers provision for multi-cloud
- Establish and control Cloudera Enterprise across AWS, Microsoft Azure and Google Cloud Platform.
- Spin up and terminate clusters, and only pay for what is needed when need it and developing and training data models for reporting, exploring, and self-servicing business intelligence.
- Delivering real-time insights for monitoring and detection and conducting accurate model scoring and serving.

K. Open Refine:

Open Refine is a strong big data tool. It guides to work with difficult data, cleaning it and transforming it from one format to another. It also performs extending it with web services and external data.

Attributes:

- Open Refine tool help you explore large data sets with ease and it can be used to link and extend your dataset with various web services it also imports data in various formats to explore data sets in a matter of seconds.
- Apply basic and advanced cell transformations that allow dealing with cells that contain multiple values to create instantaneous links between datasets.
- Use named-entity extraction on text fields to automatically identify topics to perform advanced data operations with the help of Refine Expression Language.

L. RapidMiner:

RapidMiner is an open source big data tool. It is used for data preparation, machine learning, and model deployment. It deals with a suite of products to build new data mining processes and setup predictive analysis.

Attributes:

- Allow multiple data management methods, GUI or batch processing to integrates with in-house databases for interactive, shareable dashboards
- Big Data predictive analytics and remote analysis processing
- Data filtering, merging, joining and aggregating and build, train and validate predictive models
- Store streaming data to numerous databases to reports and triggered notifications.

M. Data Cleaner:

Data Cleaner refers to the data quality analysis utilization and a solution platform. It has tough data profiling engine. It is elasticity and thereby adds data cleansing, transformations, matching, and merging.

Feature:

- Interactive and explorative data profiling and fuzzy duplicate record detection with data transformation and standardization, data validation and reporting
- Use of reference data to cleanse data and master the data ingestion pipeline in Hadoop data lake to ensure that rules about the data are correct before user spends their time on the processing
- Discover the outliers and other immoral details to either exclude or suit the incorrect data

N. Kaggle:

Kaggle is the world's largest big data community. It uses organizations and researchers to report their data & statistics. It is the exact place to survey data expert.

Attributes:

- The best place to discover and seamlessly analyze open data that allows the search box to find open datasets and also contribute to the open data movement and connect with other data enthusiasts.

O. Hive:

Hive is considered as an open-source-software big data tool. It guides the programmers to analyze massive data sets on Hadoop. It supports querying and managing huge data sets in a real and also faster.

Attributes:

- It assists SQL like query language for interaction and Data modeling, and also it compiles language with two main tasks such as map and reducer
- It allows defining these tasks using Java or Python and it is designed for managing and querying only structured data
- Hive's SQL-inspired language separates the user from the complexity of Map Reduce programming to offers Java Database Connectivity (JDBC) interface.

VI. CONCLUSION

As wrapping up this paper I conclude that the Big Data can grasp and handle variety of research challenges, issues, trends and different tools for elevate an effective solution for gained problems that arise from the infinite data. According to this situation it is plainly accompany to handle the characterized design of batch processing and the real-time processing. On each manifesto that is based on Big Data has an obvious functionality. Modified techniques that are meant to adapt about analysis and survey the statistical analysis, cloud computing, data stream processing, data mining, quantum computing, and intelligent analysis and at last machine learning. We intern that in the future many people will concentrate on these types of techniques that helps to find out the problems arises from the Big Data results in inadequate as well as efficiently.

REFERENCES

[1]K.Kambatla, G. Kollias, V. Kumar, and A. Gram, "Trends in Big Data Analytics", Journal of Parallel and distributed computing, 74(7) (2014), pp.2561-2573

[2] Ferguson, Mike, "Architecting a big data platform for analytics", A Whitepaper prepared for IBM (2012)

[3]Stephen Kaisleri_SW, Alberto Espinosa, Big data: Issues and challenges moving forward Stephen Kaisleri_SW, AlbertoEspinosa 013 46th Hawaii International Conference on System Sciences

[4] Neelam Singh, Neha Garg, Varsha Mittal, Data-insights, motivation and challenges, Volume 4, Issue 12, December-2013, 2172, ISSN 2229-5518 2013

[5] Karthik Kambatlaa, Giorgos Kollias b, Vipin Kumar, Ananth Grama, Trends in big data analytics, (2014) 74 2561-2573

[6]Big Data Analytics what it is and why it matters, SAS, accessed 14 May 2018, https://www,sas,com/en_us/insights/analytics/big-data-analytics.html

[7]Vangie Beal, (2018), Big Data Analytics, accessed 10 May 2018, https://www.webopedia.com/TERM/B/big_data_analytics.html

[8] Ankita S.Tiwarkhede, Prof. Vinit Kakde, A Review Paper on Big Data Analytics, Volume 4, Issue 4, April 2015, ISSN 2319-7064

[9] Pritee Chunarkar-Patil, Akshanda Bhosale, Big data analytics, Volume 2, Issue 5, September 2018, DOI: 10.15406/oajs.2018.02.00095

[10]Dr. S. Vijayarani, Ms. S. Sharmila, Research in the big data-An overview, Volume 4, Issue 3, September 2016, DOI: 10.5121/ieij.2016.4301

[11]Prof. Raj Jain, Survey of Recent Research Progress and Issues in Big Data, accessed 20 December 2013, https://www.cse.wustl.edu/~jain/cse570- 13/ftp/bigdata2/index.html

[12]Althaf Rahaman. SK, Sai Rajesh. K, Girija Rani.k, Challenging tools on Research Issues in Big Data Analytics, Volume 6, Issue 1, (2018), ISSN: 2321-9939

[13]Dr. Venkatesh Naganathan, Comparative Analysis of Big Data, Big Data Analytics: Challenges and Trends, e-ISSN: 2395-0056, p-ISSN: 2395-0072, (2018), pp.1948-1964

[14] Hadoop. https://hadoop.apache.org/.

[15] Hadoop MapReduce Tutorial. https://hadoop.apache.org/common/docs/r0.20.2/mapred_tutorial.html

[16]Danyang Dua, Aihua Lia*, Survey on the Applications of Big Data in Chinese Real Estate Enterprise 1st International Conference on Data Science, 2014

[17] https://www.pgs-soft.com/harnessing-big-data/

[18] LITERATURE SURVEY (BIG DATA ANALYTICS)

[19] Tole, Alexandru Adrian, "Big data challenges", Database Systems Journal 4.3 (2013): 31-40
[20] Jaseena, K.U., and Julie M.David, "Issues, challenges, and Solutions: Big Data Mining", Netcom, CSIT, GRAPH-HOC, SPTM-2014 (2014): 131-140
[21] Zicari, Roberto V, "Big data: Challenges and opportunities", Big data computing (2014): 564
[22] Ammu, Nrusimham, and Mohd Irfanuddin, "Big Data Challenges", International Journal of Advanced Trends in Computer Science and Engineering 2.1 (2013): 613-615
[23] Debi Prasanna Acharya, Kauser Ahmed P. (2016), "A Survey on BigDataAnalytics: Challenges, Open Research Issues, and Tools", accessed 20 April 2018, https://www.researchgate.net/publication/296550027_A_Survey_on_Big_Data_Analytics_Challenges_Open_Research_Issues_and_tools
[24] Nasser, T, and R. S. Tariq, "Big data challenges", J Comput Eng Inf Technol 4: 3, DOI: http://dx.doi.org/10.4172/23249307 (2015): 2
[25]https://www.prnewswire.com/news-releases/study-finds-marketers-struggle-with-the-big-data-and-digital-tools-of-today-142312475.html
[26] Anurag Agrahari, Prof D.T.V. Dharmaji Rao, "A Review Paper on Big Data: Technologies, Tools, and Trends" IRJET- e-ISSN: 2395-0056 p-ISSN: 2395-0072 volume: 4, Issues: 10, Oct-2017