

Big Data Tools and Techniques: A Roadmap for Predictive Analytics



Ritu Ratra, Preeti Gulia

Abstract: Nowadays, large volume of data is generated in the form of text, voice, video, images and sound. It is very challenging job to handle and to get process these different types of data. It is very laborious process to analysis big data by using the traditional data processing applications. Due to huge scattered file systems, a big data analysis is a difficult task. So, to analyses the big data, a number of tools and techniques are required. Some of the techniques of data mining are used to analyze the big data such as clustering, prediction, and classification and decision tree etc. Apache Hadoop, Apache spark, Apache Storm, MongoDB, NOSQL, HPCC are the tools used to handle big data. This paper presents a review and comparative study of these tools and techniques which are basically used for Big Data analytics. A brief summary of tools and techniques is represented here.

Keywords: Big data, Clustering, Hadoop, Spark, MongoDB, HDFS

I. INTRODUCTION

Big data is a term that can be defined as a set of techniques that evoke a type of integration that is used to uncover huge unknown values. These values are of very complex nature, and of an immense scale. The term “Big Data” was first documented in research paper by scientist at NASA in 1997. Big data is the term that is related to enormous velocity, variety and volume of data. Big data is an adage, operates to designate a large amount of data. This data can be structured and unstructured. And it is difficult to practice by using the traditional tools and techniques. In IT industries there is big amount of big Data that is common to many departments i.e. very huge amount of data lies in the store of industries and no tool is exist to handle that data before big data comes into picture [2].

A. Parameters of Big Data:

Big Data Analytics deals with storing and processing of the different, difficult, complex and enormous datasets. In general, Big Data is characterized by ten parameters [20]. These are also known as V’s of big data. These are shown in the fig 1. These are

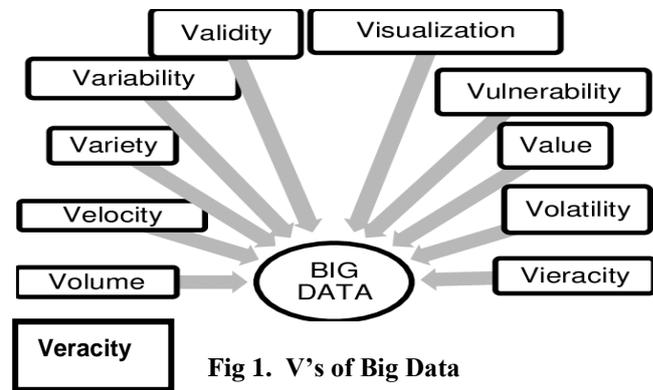


Fig 1. V’s of Big Data

Value is the one of the foremost property of big data. It is a significant for infrastructure system of IT, to save enormous amount of values in their databases. **Velocity** is a term that is concern with the high speed of data generation .The rate at which speed data is generated, influences the potential of data. The flow of data is massive and continuous. **Variety** states to various types of sources and the data is presented in both forms that are structured and unstructured. Now data in the form of videos files, e-mail, audios files, world processing files etc. is also being considered. **Volume** is the term of “Big Data” that is concerned with extremely large data. It is volume that decides that whether a data is big data or not. So, “Volume” is the important parameter out of various parameters that should be considered while dealing with ‘Big Data’. **Variability** is a term that deals with the inconsistency of data. Data is not 100% correct when dealing with large volume of data. **Validity** deals with accuracy and correctness of data. **Vulnerability** is concern with the security feature of data. After all, a data breaches with big data is a big breach. **Volatility** is a parameter of Big Data that is concern with statistical measure of the dispersion for a given set of returns. **Visualization** is a current characteristic of big data that is deal with visualization of data. **Variability** in big data’s context is the inconsistency of speed at which the data is stored in to our system There are a number of tools and techniques used to analyse the big data. To discuss these tools and techniques this paper is organized as follows: The section 2 explains the techniques used in data mining for Big data, The section 3 provides the overview of tools of Big Data Analytics and comparative study of Hadoop, Spark and MongoDB , Section 4 concludes the work .

II DATA MINING TECHNIQUES FOR BIG DATA

The techniques for big data mining are as follows:

Revised Manuscript Received on December 30, 2019.

* Correspondence Author

Ms Ritu Ratra, Research Scholar, Department of Computer Science & Applications, Maharshi Dayanand University, Rohtak, Haryana

Dr. Preeti Gulia, Assistant Professor, Department of Computer Science & Applications, Maharshi Dayanand University, Rohtak, Haryana

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

A. Classification Trees:

It is a most common technique that is used in data-mining to classify a categorical variable which is dependent and this is based on the measurements. This measurement is of forecast variables which can be one or many. The output is in the tree format with nodes and there is an association in the nodes. And this can be read to form if-then rules [21], [22]. It builds classification in the form of trees structure. As shown in fig 2, it divides the data nodes into small subgroup. These methods can be used when the data mining task has predictions or classification of outcomes.

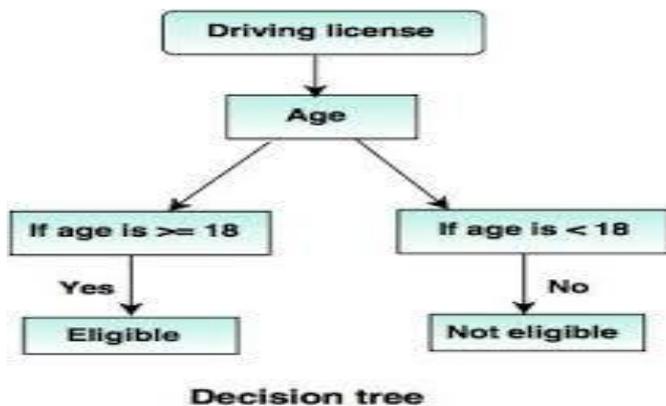


Fig 2. Classification Trees [23]

B. Logistic Regression:

It is statistical technique that expands the conception to work with classification and provides a method that is used to predict the probability of the occurrence of the independent variables as a function.

C. Neural Networks:

It is a software algorithm that is a set of comparable architecture of brains of animals which are parallel in nature. In this the association is consider as network and consists of the following elements: 1. input layers, 2. hidden layers, and 3. output layers as shown in the fig 3. Particular weight is assigned to each unit. Different varieties of data is inputted to the input node, and by a technique of trial and error, various algorithms are used to adjust the weights until it meets a certain halting criteria.

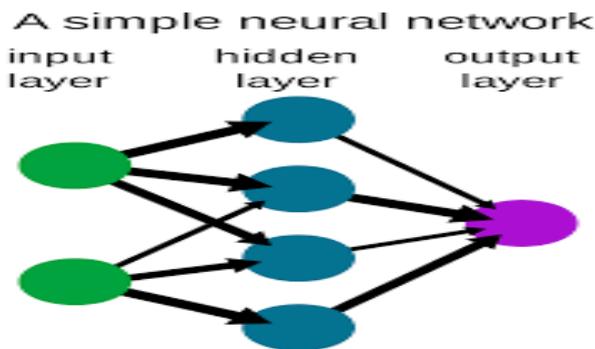


Fig 3. Neural Networks [23]

D. Clustering Techniques:

A technique that is used to partitioning a set of data into meaningful subclasses which is knows as cluster. The K-

nearest neighbor technique can be used for this. It finds out the gap between the actual record and the training data. After finding that gap it assigns the particular record to the class which is closest neighbor in the whole data set.

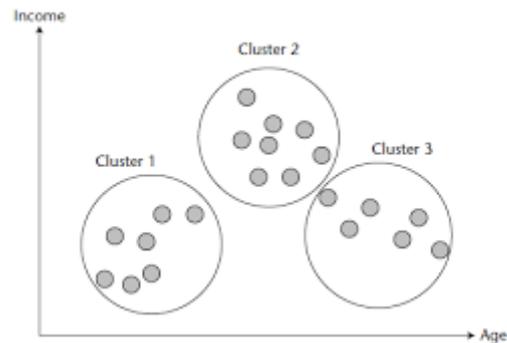


Fig 4. Clustering Technique [23]

E. Association Rule Mining:

At the very basic level, it is observed that **Association rule mining technique** involves the use of Machine learning models. The main purpose is to analyze data for patterns, in a database. It is used to identify the frequent if-then associations, and these are called *association rules*. The antecedent (if) and the consequent (then) are the two parts of an association rule. An antecedent (if) is concerned as an item that is discovered within the data and a consequent (then) is concerned as an item that is discovered with the antecedent.

F. Machine Learning

In Machine learning various software can be included that can learn from the data. It provides computers that ability which can be used to learn without being explicitly programmed, and the main focus is on making predictions which are based on known properties. Machine learning, nowadays is being used to help to find the difference between spam and non-spam messages and to learn preferences of users and after than make recommendations which are based on that information etc.

II. BIG DATA ANALYTICS TOOLS

The word “Big data” can be applied to a dataset which increases at very intense rate. And it becomes difficult to store and process that data. Big data constantly increasing from a few TB of data to many PB of data, so there are some problems related to storage, searching, sharing, visualizing and analytics. Hence big data analytics is where the use some advance techniques which are applied on big data sets; there is a variety of tools that are used for Analytics of Big Data. The tools used for the same purpose can be categorized into different stages of lifecycle of big data as shown in Table I. That is based on their usage and implementation [22].

Table I: Types of Big Data Analytics Tools

	Data Collection tools	Data Storage tools and frameworks	Data filtering and extraction tools	Data cleaning and validation tools
1	Semantria	Apache HBase (Hadoop database)	Scraper	DataCleaner
2	Opinion Crawl	CouchDB	OctoParse	MapReduce
3	OpenText	MangoDB	ParseHub	Rapidminer
4	Trackur	Apache spark	Mozenda	OpenRefine
5	SAS Sentiment Analysis	Oracle, NoSQL Database	Content Grabber	Talend

A. Data collection tools:

There is no doubt that today that there are a number of Big Data tools that are present in market. Semantria, Opinion Crawl, OpenText, Trackur are some of them which are commonly used.

Semantria:

Semantria provides us with a unique service which is proceeded towards by collecting various information from various clients. And after then the process of analyzing that information is applied diligently to produce the most valuable and desirable insights. It is very helpful to find trends and identifies various patterns which are useful to get success. Semantria is a tool that powerfully combines various text analytics. With this, users can enjoy the following benefits: collect more reliable, actionable insights. Some features of Semantria are fast customizable, comprehensive system, extraction of entity, classification, clustering, visualization, 10+ languages etc.

Open text:

The Open Text is Sentiment Analysis module. It is a special type of engine that is used in classification to find out various subjective patterns. It is also used to evaluate the expressions of sentiment that is present in text form. First of all the analysis work is done at the topic level, sentence level, and document level. Its prime function is to acknowledge whether parts of text are realistic.

Trackur:

Trackur is a tool that is used to collect the information. It uses its automated sentiment analysis to look at the specific keywords that the users are supervising and after then decisions are carried out. The sentiment may be positive, negative or may be neutral with the related document. In Trackur algorithm, it could be used to observe the social sites and can outline news, to collect information through the trends and automated sentiment analysis.

SAS Sentiment Analysis

SAS is also a sentiment analysis tool that automatically extracts sentiments in real time. It performs this task with the help of combination of various statistical modeling techniques. These processing techniques are based on rule of natural language. There are some automated reports which are built-in show the various patterns and show their reactions in detail. With the help of evaluations that are ongoing, user can properly refine their models and make different adjustments related to classifications.

Opinion Crawl

Opinion Crawl, an online tool, used as a sentiment analysis for current affairs. It permits various visitors to evaluate the Web sentiment based on a particular topic like a company. User can input a topic that he want to assess and can get an ad-hoc sentiment assessment related to the topic. User will get a pie chart for every individual topic that shows the real-time sentiment. There is also a list of the current news headlines, some images related to the topic. All these ideas allow user to check that what type of issues are derived the sentiment i.e.in a positive way or negative way.

B. Data Storage and frameworks tools:

The captured data that may be structured or unstructured need to be stored in databases. There is need of some databases to accommodate Big Data. A lot of frameworks have been developed by organizations like Apache, Oracle etc. that are used as analytics tools to fetch and process data which is stored on these repositories. Some of these are as follows:

Apache Hadoop

Apache Hadoop is the one of the technology designed to process Big Data, which is unification of structured and unstructured data huge volume. Apache Hadoop is an open source platform and processing framework that exclusively provides batch processing. Hadoop was firstly influenced by Google's Map Reduce. In Map Reduce software framework the whole program is divided into a number of parts these are small is size. These small parts also called as fragments. These fragments can be executed on any system in the cluster [5].

Components of Hadoop

There are a lot of components which are used in composition of Hadoop. These all worked together to execute batch data. Main components are as:

HDFS:

The Hadoop Distributed File System (HDFS) is the main component of the Hadoop software framework. It is the file system of Hadoop. HDFS is configured to save large volume of data. It uses low-cost hardware that is distributed in nature [4]. It is a fault-tolerant storage system that stores large size files from TB to PB. There are two types of nodes in HDFS Name node and Data node.

Name Node, it works as the master node. It contains the all information related to the all data node. It has the information of free space, addresses of nodes, all the data that they store, active node, passive node. It also keeps the information of task tracker and job tracker.

Data Node: Data node is also known as slave node. Data node in Hadoop is used to store the data. And it is the duty of TaskTracker to keep the track of on-going job which resides on the data node and it also take care of the jobs coming from name node.

MapReduce: MapReduce is a framework that helps developers to jot down programs to method massive volume of unstructured knowledge parallel over a distributed design. MapReduce consists of many elements like JobTracker, TaskTracker, JobHistoryServer etc. it's additionally referred to as the Hadoop's native instruction execution engine. It was introduced to process the huge amount of data and to store these huge data on commodity hardware. For processing the large volume data it uses clusters to store records. Map function and Reduce function are two functions that are the base of the Map Reduce programming model. In master node the Map function works. And it accepts the input. And after then divide that accepted input into sub modules and then distribute it into slave nodes.

YARN YARN (Yet another Resource Negotiator) is the core Hadoop services that supports two major Services: World resource management (ResourceManager) and per-application management (ApplicationMaster). It is the cluster coordinating element of the Hadoop stack. YARN makes it attainable to execute [7] [19]. It is the MapReduce engine that is responsible for practicality of Hadoop.

MapReduce is a framework that run on hardware that are less costly. It doesn't conceive to save anything in memory. MapReduce has unimaginable measurability potential. It has been employed in creation of thousands of nodes. Different additions to the Hadoop scheme will scale back the effect of this to variable degrees; however it will always be an element in quickly implementation of an inspiration on a cluster of Hadoop.

Working of Hadoop

In the architecture of Hadoop there is only one master node ,works as master server known as JobTracker. There are several slave node servers known as TaskTracker's. Keeping the track of the slave nodes is the central job of JobTracker. It established an interface infrastructure for various job. Users input the MR(MapReduce) jobs to the JobTracker, where the pending jobs are reside in queue. The order of access is FIFO. In Fig 5 working of Hadoop is given. It is the responsibility of JobTracker to coordinate the mapper's execution and reducer's execution [5]. When the Map Task is completed, the JobTracker starts its functionality by initiating the reduce task. Now it is the duty of JobTracker to give proper instruction to TaskTracker. After then TaskTracker starts the downloading files and mainly concatenate the various files into a single unit (entity).

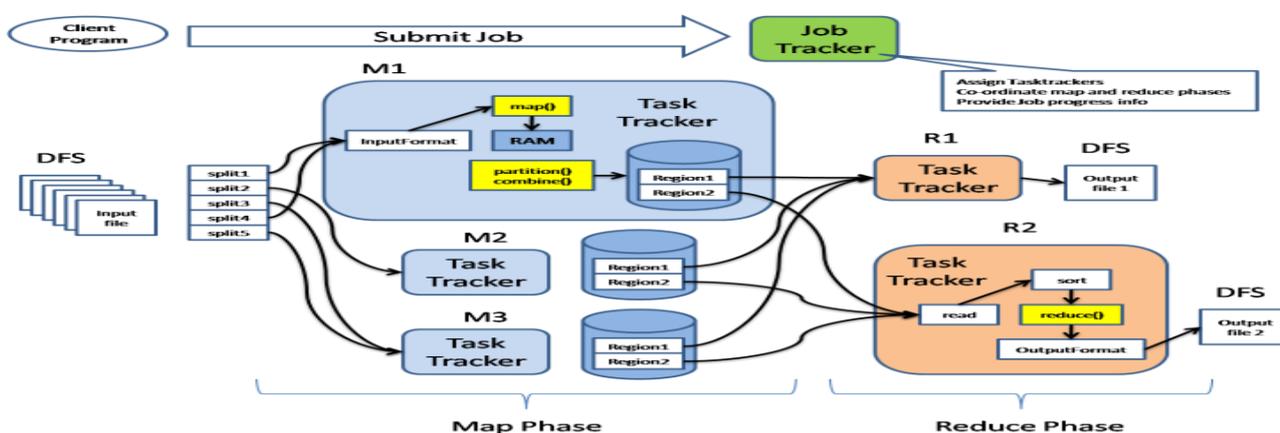


Fig 5. Architecture of Hadoop [18]

Apache Spark

Apache Spark is also an additionally open supply tool that's developed by in AMP science laboratory at UC Berkley [7]. It is framework with stream processing abilities. Engineered victimization several of a similar principles of Hadoop's MapReduce engine, Spark focuses totally on dashing up process and instruction execution workloads by providing full in-memory computation and processing improvement. Through this we are able to do In-memory analytics .Speed of this is often (100 times) quicker than Hadoop. it's extremely compatible with Hadoop's storage. For

implementation of associate degree in-memory batch computation, Apache Spark uses a model known as RDDs (Resilient Distributed Datasets), to Fig with information [3]. Stream process: Model Stream process capabilities area unit equipped by Spark Streaming. Spark itself is meant with batch-oriented workloads in mind. To manage the inequality between the engine styles and also the characteristics of streaming workloads, Spark implements an inspiration known as micro-batches.

This strategy is meant to treat streams of information as a series of terribly little batches which will be handled victimization the native linguistics of the batch engine. The clear reason to use Spark over Hadoop MapReduce is speed and its advanced DAG programming. Another of Spark's major benefits is its skillfulness. It is deployed as a standalone cluster or integrated with associate degree existing Hadoop cluster.

MongoDB

MongoDB is a database that is based on JSON documents. It is written in C++ and launched in 2009, and is still expanding. MongoDB may be information that may be used each in little comes that have many hundreds of users. MongoDB database basically holds the set of data that has no defined schema. There is no predefined format like tables, and can stores data in the form of BSON documents. BSON are binary encoded JSON like objects. User can go for MongoDB rather than MySQL if the requirement is knowledge intensive because it stores information and queries [23, 24]. MongoDB belongs to the NoSQL family

and based on C++. MongoDB is especially engineered for storage of information and retrieval of stored information. It can do processing and measurability. It supported C++ and belongs to the NoSQL family. It doesn't trust that aren't time-sensitive, making relative tables. It uses document format for storage of its records. Detailed Comparison among Apache Hadoop, Apache Spark and MongoDB is shown in Table II.

Table II: Comparison among Hadoop, Spark and MongoDB

Criteria	Hadoop	Spark	MongoDB
Processing Model	Batches	Mini batches	Single Document
Latency	High latency	Very Low latency	Low
Fault tolerance	Uses replications	Uses RDD	Replica Set
Supported languages	Java	Scala, Java ,Python	C++
Batch Framework	HDFS	Core Spark API	BSON
Security	High Secure	Less secure	Secure
Advantages	Storage and processing speed, low cost, secure	Scalable, high speed, powerful caching.	Expandable
Weakness	'Name Node', is the single point of failure	Expensive, less no. of algorithm, small files problems.	Fault tolerance issue.
Hardware Cost	Cost is more	Cost is more.	Cost effective because it is a single product.

C. Data filtering and extraction tools:

Some tools are used for data filtering and extraction. These tools are very helpful to collect useful information from Internet. These are:

Mozenda

Mozenda provides technology, delivered as either code (SaaS and on-premise options) or as a managed service, which enables folks to capture unstructured internet information, convert it into a structured format. Internet scraping is that method of mechanically mining information or collection info from the Wide internet.

"Mozenda uses a point-and-click code tool to show websites into structured information. Acknowledged by numerous names, like internet scraping, internet information gathering, internet information extraction, information extraction and

information scraping, the technology "enables users to draw upon web-scraped information for business intelligence and large information applications and makes information extracted from websites usable as CRM, ERP, or different transactional information." With Mozenda, users will capture information lists, produce helpful combine lists, collect information from tables and extract information from common formats.

Octoparse

It is the last word tool for information extraction (web locomotion, information locomotion and information scraping). You'll be able to flip the complete net into a structured format with Octoparse internet scraping tool. So as to attain automatic internet scraping during a real sense, the.

Octoparse team has never stalled its pace in creating information additional accessible and prepared to everyone. It's unmovable in our belief that within the era of huge information, anyone ought to be blessed the potential to gather information therefore on harness the ability of huge information. With precise information at hand, you'd be able to conduct information analysis, selling strategy, sentiment analysis, crusade, lead generation and additional.

Webhose.io

Webhose.io gives direct access to period of time. It provides the structured information from hundreds of on-line records. The online hand tool helps in extracting web information. It supports more than 240 languages and storing the output information in numerous formats.

Spinn3r

Spinn3r permits to fetch the complete information from various social media application. It is distributed with the API of fire station and that manages ninety fifth of the assortment work. It JSON files to saves the extracted information. The net hand tool perpetually scans the net and find out the updates from various multiple sources to induce you period of time. Its admin is console that allows to management of crawls. The full-text search permits creating advanced queries on information.

Scraper

Scraper could be a Chrome extension with restricted information extraction options however it's useful for creating on-line analysis, and commerce information to Google Spreadsheets. This tool is meant for beginners yet as consultants UN agency will simply copy information to the writing board or store to the spreadsheets exploitation OAuth. Scraper is also a free of cost tool that works in the browser and auto-generates smaller XPath's for outlining URLs. It doesn't offers the benefit of larva locomotion like Import, however it's additionally a profit for novices as you don't have to be compelled to handle untidy configuration.

D. Data cleaning and validation tools:

Data cleansing, and validation is an important stage. Various validation rules are used to confirm the necessity and relevance of data extracted for analysis. Sometimes it may be difficult to apply validation constraints due to complexity of data. Data cleaning tools are very helpful because they help in minimizing the processing time. They also reduce the computational speed of data analytics tools. A significant overview of latest data cleaning tools is provided as below.

DataCleaner

DataCleaner is a tool that is integrated with Hadoop. Data transformation, validation and reporting are its main features. It is a tool which is an application for data quality analysis. There is a profiling engine in its core to profile the data. This can be extensible by adding data cleansing, transformations, reduplication, matching merging and enrichment. It profiles and analyses the database within minutes, discovers patterns with the Pattern Finder, finds frequency of data using Value Distribution profile, filters the

contact details, detects duplicates by using fuzzy logic, Merge the duplicates values etc. These are some features of DataCleaner [25]. It can access almost any data store like MySQL, Oracle, MongoDB, dbase, and many more.

RapidMiner

With RapidMiner Studio accessing, loading and analyzing of any type of data is possible. Data can be both structured and unstructured like text, still images, and media. It extracts useful information from traditional structured data and can transformations from unstructured format to structured format. Using this access of more than 35 file types including ARFF, SAS and via URL is possible. It has wizards for various database connections like Microsoft Excel & Access, CSV etc. It is capable to access to Cassandra, NoSQL databases MongoDB, Dropbox, web pages, PDF and many more applications.

OpenRefine

OpenRefine is an open source power tool for messy data. It working with messy and unstructured data and improving it. Openrefine is a manipulation tool that cleans, reshapes and intelligently edits the data batches [26]. It is free and its code can be reused in other applications too. It offers many features like clustering, editing cells, faceting, reconciling, which helps to clean the data effectively. Openrefine is very easy as excel and a powerful tool like access database.

IV. CONCLUSION

In present scenario, a large volume of data is generated in different forms i.e. text, audio, video, image etc. It is very difficult to handle this large amount of data i.e. big data and to search out associations, patterns and to investigate the massive information sets from this. This paper gives the idea about the analysis of big data. A number of tools and techniques are discussed to analyze big data. For batch-only workloads, Hadoop could be a nice selection that's probably less costly to implement than another solutions. Batch-only workloads aren't time-sensitive. Hadoop provide a properly-tested instruction execution model that's best fitted to handling terribly massive datasets wherever time isn't a big issue. A spark could be a nice possibility for those with various process workloads. Spark instruction execution offers unimaginable speed benefits. Spark Streaming could be a smart stream process answer for workloads that value throughput over latency. It has integrated libraries, wide storage, and flexible integrations. MongoDB is excellent in data consistency and data partitioning. The purpose of this paper is to present various big data handling techniques and tools used to handle a large volume of data which is collected from various sources and can improve the overall performance of systems.

REFERENCES

1. B.Thillaieswari., "Comparative Study on Tools and Techniques of Big Data Analysis" International Journal of Advanced Networking & Applications (IJANA) Volume: 08, Issue: 05 Pages: 61-66 (2017) Special Issue.

2. Elgendy, N. "Big Data Analytics in Support of the Decision Making Process", MSc Thesis, German University in Cairo, p. 164 (2013).
3. G. George and D. Lavie, "Big data and data science methods for management research", Academy of Management Journal, vol 59, issue 5, pp. 1493 – 1507, 2016.
4. Kaisler S, Armour F, Espinosa JA, Money W. "Big data: issues and challenges moving forward" In: System sciences (HICSS), 2013 46th Hawaii international conference on, IEEE. 2013. pp. 995–1004.
5. Kubick, W.R. "Big Data, Information and Meaning", In: Clinical Trial Insights, pp. 26–28 (2012)
6. Khan N, Yaqoob I, Hashem IAT, et al. "Big data: survey, technologies, opportunities, and challenges", Sci World J. 2014;2014:712826.
7. TechAmerica: "Demystifying Big Data: A Practical Guide to Transforming the Business of Government", In: TechAmerica Reports, pp. 1–40 (2012)
8. <https://spark.apache.org/>
9. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Byers, A.H., "Big Data: The Next Frontier for Innovation, Competition, and Productivity", In: McKinsey Global Institute Reports, pp. 1–156 (2011)
10. Mouthami, K., Devi, K.N., Bhaskaran, V.M., "Sentiment Analysis and Classification Based on Textual Reviews", In: International Conference on Information Communication and Embedded Systems (ICICES), pp. 271–276 (2013).
11. Plattner, H., Zeier, A, "In-Memory Data Management: An Inflection Point for Enterprise Applications", Springer, Heidelberg (2011).
12. P. Perner (Ed.): ICDM 2014, LNAI 8557, pp. 214–227, 2014.
13. Russom, P, "Big Data Analytics. In: TDWI Best Practices Report", pp. 1–40 (2011).
14. Sanchez, D., Martin-Bautista, M.J., Blanco, I., Torre, C, "Text Knowledge Mining: An Alternative to Text Data Mining", In: IEEE International Conference on Data Mining Workshops, pp. 664–672 (2008).
15. errat, O, "Social Network Analysis. Knowledge Network Solutions", 28, 1–4 (2009).
16. Song, Z., Kusiak, A, "Optimizing Product Configurations with a Data Mining Approach", International Journal of Production Research 47(7), 1733–1751 (2009).
17. Thuan L. Nguyen, "A Framework for Five Big V's of Big Data and Organizational Culture in Firms", IEEE International Conference on big Data Mining Workshops ,pp 5411-5413(2018).
18. <https://hadoop.apache.org/docs/>
19. Ms. Komal, "A Review Paper on Big Data Analytics Tools" (IJTIMES), e-ISSN: 2455-2585 Volume 4, Issue 5, May-2018, pp 1012-1017.
20. <https://tdwi.org/articles/2017/02/08/10-vs-of-big-data.aspx>
21. Nirmal Kaur, Gurbinder Singh, "A Review Paper On Data Mining And Big Data", International Journal of Advanced Research in Computer Science, Volume 8, No. 4, May 2017, ISSN No 076-567, pp 407-409.
22. J.Nageswara Rao, M.Ramesh, "A Review on Data Mining & Big Data, Machine Learning Techniques", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7 Issue-6S2, April 2019, pp 914-916.
23. <https://www.google.com/imgres>.
24. MongoDB, Inc. (2015, Aprilie), MongoDB Ops Manager Manual Release 1.6,[Online].Available:<https://docs.opsmanager.mongodb.com/current/opsmanager-manual.pdf>
25. R. P Padhy, M. R. Patra, S. C. Satapathy, "RDBMS to NoSQL: Reviewing Some Next-Generation Non-Relational Database's", International Journal of Advance Engineering Sciences and Technologies, Vol. 11, Issue No. 1, 015-030, 2011.
26. <http://datacleaner.github.io>
27. <https://github.com/OpenRefine/OpenRefine>

AUTHORS PROFILE



Ms Ritu Ratra, Research Scholar, Department of Computer Science & Applications, Maharshi Dayanand University, Rohtak, Haryana



Dr. Preeti Gulia, Assistant Professor, Department of Computer Science & Applications, Maharshi Dayanand University, Rohtak, Haryana