A Review on the Study and Analysis of Big Data using Data Mining Techniques

Anupama Jha
Assistant Professor, VIPS, GGSIPU
New Delhi, India

Dr. Meenu Dave
Professor, Jagannath University
Jaipur, India

Dr. Supriya Madan
Professor, VIPS, GGSIPU
New Delhi, India

Abstract-- Big Data is an emerging concept that describes innovative techniques and technologies to analyze large volume of complex datasets that are exponentially generated from various sources and with various rates. Data mining techniques are providing great aid in the area of Big Data analytics, since dealing with Big Data are big challenges for the applications. Big Data analytics is the ability of extracting useful information from such huge datasets. This paper presents a literature review that include the importance, challenges and applications of Big Data in various fields and the different approaches used for Big Data Analysis using Data Mining techniques. The findings of this review give relevant information to the researchers about the main trends in research and analysis of Big Data using different analysis domains.

Keywords-- Big Data, Big Data Analytics, Big Data Application, Data Mining,

I. INTRODUCTION

In this digital era, analysts have enormous amounts of data available on hand. Big Data is the term for a collection of unstructured, semi-structured and structured datasets whose volume, complexity and rate of growth make them difficult to be captured, managed, processed or analyzed by using the typical database software tools and technologies. Different varieties are in the form of text, video, image, audio, webpage log files, blogs, tweets, location information, sensor data etc. . Discovering useful insight from such huge datasets requires smart and scalable analytics services, programming tools and applications [1].

Data mining is also known as Knowledge Discovery in Database (KDD) is an analytical process used in different disciplines to search for significant relationships among variables in large data sets. Analyzing fast and massive stream data may lead to new valuable knowledge and theoretical concepts. Big data has potential to help organizations to improve operations and make faster & more intelligent decisions.

II. BIG DATA

Big Data means not only an enormous volume of data but also other features that differentiate it from the concepts of "very large data" and "massive data". In fact several definitions for Big Data are found in the literature.

International Data Corporation (IDC) defines Big Data as: "Big Data technologies describe new generation of technologies and architectures designed to economically extract value from very large volumes of a wide variety of data, by enabling high-velocity capture, discovery and/or analysis"[2].

McKinsey Report defines Big Data as "data sets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze"[3].

Big Data is also defined in terms of 3 Vs i. e Volume, Variety, and Velocity. Volume represents the size of the data. Velocity refers to the speed of both data generation and data delivery of real-time data. Variety makes the data too big as data comes from the various sources shown in Table 1.

TABLE I: DIFFERENT TYPES OF BIG DATA & ITS SOURCES

Data Types	Sources	Formats
Structured	Business Applications such as retail, finance, bioinformatics etc.	RDBMS, OLAP, Data warehousing
Semi-structured	Web Applications such as web logs, email, webpages	XML, CSV, HTML, RDF
Unstructured	Images, Audio, Video, Sensor data, Blogs, Tweets etc.	User generated text content

The above definitions for Big Data provide a set of tools to compare the emerging Big Data with traditional data analytics. This comparison is summarized in Table 2 under the dimension of 3V's Volume, Velocity and Variety.

TABLE II: COMPARISON OF BIG DATA AND TRADITIONAL DATA

Characteristics	Big Data	Traditional Data
Volume	Terabyte, Petabyte, Exabyte	Gigabyte
Velocity	More rapidly	Per hour, day
Variety	Structured, Semi-structured or Unstructured	Structured
Data integration	Difficult	Simple
Data access	Real time or batch	Interactive
Source of data	Fully distributed	Centralized

Big Data and its analysis is the centre of modern science and business areas. Large amount of data is generated from the various heterogeneous sources, hence it become difficult to store, extract, transform and load [4]. A recent study estimated that every minute Google receives over 4 million searching queries, e-mail users send over 200 million emails, YouTube users upload 72 hours of video, Facebook users share over 2 million pieces of content, 350 GB of data is processing on facebook , more than 570 websites are created every minute and Twitter users generate 277,000 tweets [5]. The estimation about the generated data is that till 2003 it was represented about 5 exabytes, then until 2012 was 2.7 Zettabytes and till 2016 it is expected to increase 4 times[6].

Big datasets are available in astronomy, social media/entertainment and networking sites, life sciences, healthcare, video surveillance, government data, sensor technology and networks, mobile devices etc.

The Fig. 1 shows the data in Terabytes (TB) for the year 2001-12 [7]. From Year 2005-12, it would appear from this graph that the amount of data was exponentially increased within this period due to the significant contribution of Big Data Analytics.

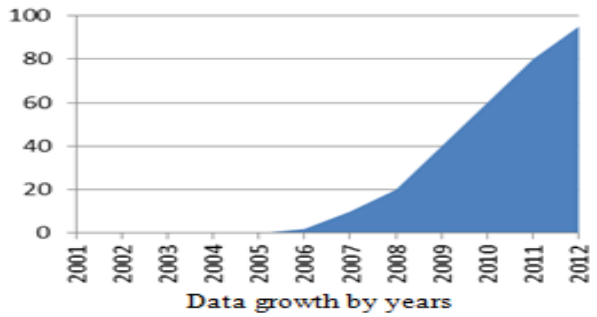


Fig. 1 2005-12 Exponential Data Growth (Data in TB)

Capturing of Big Data, its duration, storage, sharing, analysis, visualization of the massive data and of course the most important technology are several challenges that need to be faced by the enterprises or media when handling Big Data. For that the technology needs new architecture, algorithms and techniques for its implementation. It also requires technical skills. So experts are needed for this new technology to deal with Big Data. The analysis of Big Data also needs certain challenges such as Technical Challenges, Data management and sharing, Privacy, security and trust & Misuse of Big Data.

III. Data Mining

Data mining started to be an interest target for information industry because of the existence of huge data containing large amounts of hidden knowledge. In fact we could say that data mining is a relatively new scientific research area in the field of statistics, machine learning, database management science and visualization to discover and present knowledge in a form which is easily understandable to us [8]. The function of data mining is to discover hidden knowledge from the large volumes of data sets. This
extracted knowledge can help and support organizations to make better and more intelligence decisions.

Data mining systems have potential for generating millions of patterns and rules. An interesting pattern represents knowledge. Measures of pattern interestingness either objective or subjective can be used to guide the discovery process.

There are many popular models that can be efficiently used in different data mining problems. Naïve Bayes, Decision Trees, Neural Networks, Support Vector Machines, K-means are few among them.

IV. Big Data Analytics & Process Model

Big Data Analytics is the application that enables organizations to analyze large sets of data to discover patters and other useful information. Due to the significant contribution of Big Data Analytics, the amount of data was exponentially increased within the past decade i. e 2005-15.
The technological advances in storage, processing, and analysis of Big Data include:

- The rapidly decreasing cost of storage and CPU power in recent years.
- The flexibility and cost-effectiveness of datacenters and cloud computing for elastic computation and storage and
- The development of new frameworks such as Hadoop, which allow users to take advantage of these, distributed computing systems storing large quantities of data through flexible parallel processing.

In this section, we focus on the Process model for Big Data analytics. The overall model is divided into 2 Sub-Processes: Data Management and Analytics, which further broken down into 5 stages as shown in Table 3[9].

Table III: Processes for extracting information from Big Data.

Big Data Processes	
Data Management	**Analytics**
1. Acquisition & Recording 2. Extraction, Cleaning & Annotation 3. Integration, Aggregation & Representation	4. Modeling & Analysis 5. Interpretation

Data analysis is the most important stage of the Big Data Process model. The goal of this model is to extract useful insights, conclusions that will support in decision-making. The objective of data analytics is to retrieve as much information as possible that is pertinent to the subject under consideration. In Big Data Analytics, it was found by the researcher that the generated data is divided into various domains of Big Data application. Some of

them are Structured Data Analytics, Text Analytics, Web Analytics, Multimedia Analytics and Mobile Analytics.

Table 4 below summarizes all the 5 types of Big Data application, organized by data type covering data characteristics, their different approaches and techniques and their sources of data[10]. Data analytics addresses information obtained through observation, measurement, or experiments about a phenomenon of interest.

V. LITERATURE REVIEW

This section presents a comprehensive literature review
from different journals, academicians and other internet sources. It is divided into two parts. The first part presents a review based on the importance, challenges and applications of Big Data in various fields. The second part summarizes the different approaches & their outcomes for Big Data Analysis with different Data Mining techniques.

A. Literature Review: Big Data

Wei Fan and Albert Bifet in 2012 presented an overview of the topic big data mining, its current status, controversy and forecast to the future[11].

In 2013, S.Vikram Phaneendra and E. Madhusudhan Reddy illustrated that how big data differs from other data in 5 dimensions such as volume, velocity, variety, value, veracity and complexity. They explained the hadoop architecture to handle big data systems. The authors also focused on the challenges such as data privacy, search analysis etc that need to be faced by enterprises while handling Big Data [12]

The authors Shilpa and Manjit Kaur described various issues regarding Big Data. They also explained how Big Data analysis solves some problems regarding operational, financial and commercial in aviation that were previously unsolvable within economic and human capital constraints [13].

Kishor, D discussed a change to the basic definition V^3 (3V) of Big Data to C^3 (3C) so that the Big Data analytics may be better explained with mathematical and statistical techniques [14].

In 2013, Dheeraj Agarwal provides a comprehensive study for data mining, models, issue, and focuses its application[15].

Sagiroglu, S. & Sinanc, D. described the Big Data content, its scope, methods, privacy, security, samples, advantages and challenges. They found that the challenges were not only to collect and manage the data but also how to extract the useful information from that collected data [16].

Richa Gupta, Sunny Gupta and Anuradha Singhal in 2014 provide an overview on big data, its importance, technologies to handle big data and how Big Data can be applied to self-organizing websites which can be extended to the field of advertising in companies [17].

Xindong Wu, Gong-Quing Wu and Wei Ding presented a HACE theorem in 2014 that characterizes the features of
Big Data revolution and proposes a Big Data Processing model from the Data Mining Perspective [18].

In 2014, Bharti Thakur and Manish Mann overviewed types of big data and important challenges in big data management and analytics that arise from the nature of data i.e large, diverse, and evolving [19].

The authors Sabia and Sheetal Kalra presented various real time applications of big data that include healthcare, networking security, market & business, education system, telecommunication etc. [20].

The author Vatsal Shah in July 2015 analyzed large scale unstructured data in the form of video format & proposes solution to the same using Hadoop platform [21].

B. Literature Review: Big Data Analytics approach

Table 5 below summarizes the review of some of the research paper studied based on different approaches used for Big Data Analytics in the recent years.

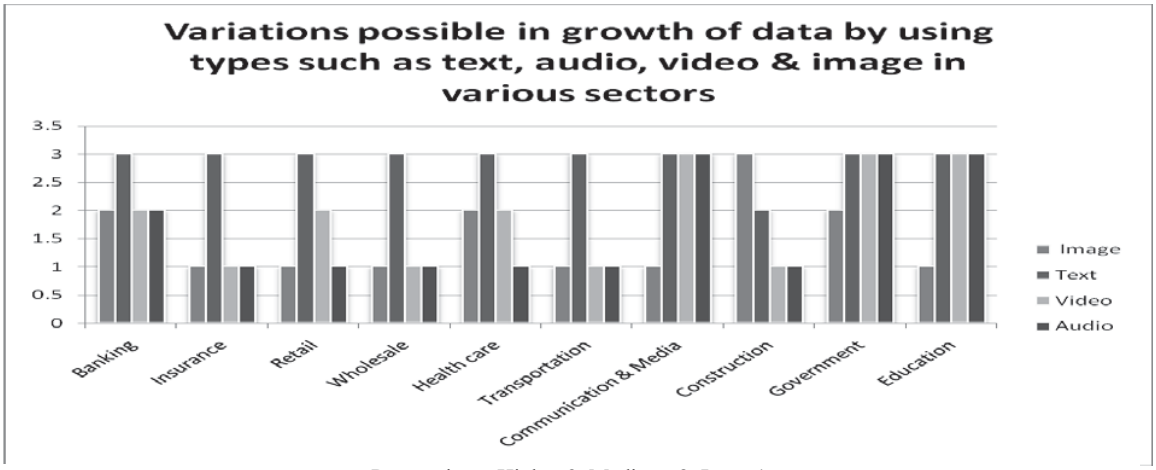
The data is from different formats such as audio, video, images and text and differ from industry to industry [22]. Banking, Insurance and Health care sectors are responsible for text data. Communication and Media are highly responsible for audio and video type of data. The Figure 2 below illustrates that there are variations shows in the amount of data stored in different sectors by using the types of data generated and stored.

IV. Conclusion and Future work

The exponential growth in terms of capacity and complexity of data in last decade has led to substantial research in the field of big data technology. In this paper, we have made an attempt to summarize the recent literature review year wise in the area of Big Data & its analysis using different analytics approaches. Text analytics which is considered to be the next generation of Big Data, now much more commonly recognized as mainstream analysis to gain useful insight from millions of opinion shared on

social media. The video, audio and image analytics technique has scaled with advances in machine vision, multi-lingual speech recognition and rules-based decision engines due to the intense interest existence of real time data of rich image and video content. They are the potential solutions to economical, political and social issues.

Our future work would primarily focuses on the Big Data analytics approach discussed above using various data mining techniques.



Penetration: High = 3, Medium=2, Low=1

Fig. 2 The types of data generated and stored varies by different sector

TABLE IV: TAXONOMY OF BIG DATA ANALYTICS

Analysis Domain	Characteristics	Approaches	Techniques	Source of Data
1. Structured Data Analytics	Structured records, less volume and real time.	Statistical Analysis	Statistical Machine Learning. RDBMS, OLAP, Data Warehouse	Business & Scientific Data.
2. Text Analytics	Language dependent, Unstructured, Semantic, Rich Textual.	Text Mining, Opinion mining or Sentiment Analysis, Summarization, Question Answering System, Natural Language Processing (NLP).	Clustering, Neural Network, SVM, Decision Tree, Naïve Bayes	Emails, logs, Blogs, Twitter, Facebook, Webpage, Corporate documents, Government Rules and Regulations etc.
3. Web Analytics	Integration of Text and Hyperlink, Symbolic, Metadata	Web Content Mining, Web Usage Mining, Web Structure Mining	Classification and Clustering.	Various Web Pages

4. Multimedia Analytics	Collection of Audio, video & image	Indexing & Retrieval, Summarization, Error detection, annotation	Multimedia annotation, Video Summarization, audio Summarization	User generated multimedia, surveillance, healthcare media, corporation produced multimedia.
5. Mobile Analytics	Fragmented data, person specific	Monitoring	Location based mining	Mobile phone applications, sensors

TABLE V: LIST OF EXISTING APPROACHES AND THEIR OUTCOMES USED IN BIG DATA ANALYTICS

Approach	Year	Title of Research paper	Authors	Outcomes
1. Structured Analytics	2013	Social Business Intelligence Using Big Data [23].	Gautam Shroff, Lipika Dey and Puneet Agrawal	Authors described how the fusion of social and business intelligence is defining the next generation of business analytics applications using a new AI driven information management architecture that is based on big data technologies.
	2015	A Review of Applications of Data Mining in the Field of Education [24].	Hardeep Kaur	Described data mining applications in the field of education and also the basics of educational data mining system.
		Data Mining for the Internet of Things: Literature Review and Challenges [25].	Feng Chen, Pan Deng, Jiafu Wan, Daqiang Zhang, Athanasios V. Vasilakos and Xiaohui Rong	Reviewed data mining in knowledge view, technique view, and application view, including classification, clustering, association analysis, time series analysis and outlier analysis.
2. Text Analytics	2011	Sentiment analysis of social media content using N-Gram graphs [26].	Fotis Aisopos, George Papadakis & Theodora Varvariqos	Significant improvements over other methods used in this context not only with respect to effectiveness but also to efficiency.
	2013	Scalable Sentiment Classification for Big Data Analysis Using Naıve Bayes Classifier [27].	Bingwei Liu, Erik Blasch, Yu Chen, Dan Shen and Genshe Chen	Evaluated the scalability of Naïve bayes classifier in large datasets. It was found that accuracy improved 82% when the dataset increases.
		Automatic Sentiment Analysis for Unstructured Data [28].	Jalaj S. Modha, Gayatri S. Pandi , Sandip J. Modha	Very useful to identify and predict current and future trends, product reviews, people opinion for social issues etc. Also useful for big organizations like SAP, SAS and TCS
	2014	Big Data Analytics: A Text Mining Based Literature Analysis [29].	Ahmed Elragal, Moutaz Haddara	Reviewed literature based on text mining and clustering techniques
		Real Time Sentiment Analysis of Twitter Data Using Hadoop [30].	Sunil B. Mane, Yashwant Sawant, Saif Kazi, Vaibhav Shinde	Access time is reduced and accuracy is found to be 72.27 %.
		Text Mining Approach for Big Data Analysis Using Clustering and Classification Methodologies [31].	Somesh S Chavadi & Dr. Asha T	It works for all kind of real time input. The model successfully works for Big data by speeding up the computation for large data sets.
		Data Mining and Data Pre-processing for Big Data [32].	Ashish R. Jagdale, Kavita V. Sonawane, Shamsuddin S. Khan	Performance analysis of the proposed system is evaluated on the basis of different factors such as the execution time of the algorithm, data scalability, flexibility and data heterogeneity

		A Big Data Methodology for Sentiment Analysis of Twitter Data [33].	Supraja.G.S, Dr Jharna Majumdar, Shilpa Ankalaki	Concluded that Birch algorithm is more efficient to perform clustering and give better result for sentiment analysis.
	2015	Opinion Mining for Reputation Evaluation on Unstructured Big Data [34].	Uma Gurav and Nandini Sidnal	Focused on combination of different classifiers techniques to overcome the challenges and enhance the granularity of opinion capturing. Identified many research challenges in sentiment analysis that are yet to be addressed.
		Parallel Implementation of Big Data Pre-Processing Algorithms for Sentiment Analysis of Social Networking Data [35].	V.Jude Nirmal and D.I. George Amalarethinam	Authors described the 3 approaches machine learning, lexicon based methods and linguistic analysis for sentimental analysis in terms of efficiency, processing power and memory bandwidth.
3. Web Analytics	**2004**	Using Web Analytics to Measure the Activity in a Research-Oriented Online Community [36].	Catherine Dwyer, Xuzhong Zhou and Starr Roxanne Hiltz.	Applied Web analytics to a research oriented virtual community to measure member's usage characteristics and interaction with the site and in e-commerce to predict and/or influence shopper's choices. They used data mining techniques to uncover browsing patterns by examining the content of Web server logs and also highlight Web design problems.
	2010	Using Web Analytics Data to support Social Software Users [37].	Alexander W. Schneider and Boltzmannstr	Discussed Web Analytics Technologies Challenges in a Social Software Context. Also reviewed how social software encompasses all web-based applications which support human communication and interrelation. Several scenarios have been developed for implementation.
	2014	Mining Web Log Files for Web Analytics and Usage Patterns to Improve Web Organization [38].	Sana Siddiqui and Imran Qadri	Reviewed the method of discovering helpful insights from the online server log file of an educational institute. Implemented on applications like net traffic analysis, economical web site administration, website modifications, system improvement and personalization and business intelligence etc.
4. Multimedia Analytics	**2008**	Visual Analytics: Scope and Challenges [39].	Daniel A. Keim, Florian Mansmann, Jorn Schneidewind, Jim Thomas, and Hartmut Ziegler	Authors gave an overview of the current state of visual analytics field and its challenges in various applications such as business, environmental monitoring, security, disaster management, healthcare etc.
	2015	An Intelligent Video Surveillance Framework with Big Data Management for Indian Road Traffic System [40].	R. Balaji Ganesh and S. Appavu	Focused on refining a framework for large scale video analytics while incorporating the simple, light-weight aspects of a video surveillance algorithm, and makes an insight by adopting blob tracking based video surveillance algorithm for large scale video analytics.
		Big Video Data Analytics using Hadoop [41].	Vatsal Shah	Analyzed video format of large scale unstructured data and also discussed the current issues and proposes solution to the same using Hadoop platform.
5. Mobile	**2014**	Use of Big Data Technology in	Punam Bedi, & Vinita Jindal	Evaluated the performance in terms of processing time. It is greatly reduced while

Analytics		Vehicular Ad-hoc Networks [42].		increasing the number of nodes in Hadoop framework,.

REFERENCES

[1] Puneet Singh Duggal and Sanchita Paul, Big Data Analysis : Challenges and Solutions
[2] Gantz, J., & Reinsel, D. (2011). The 2011 Digital Universe Study: Extracting Value from Chaos.
[3] Internet source, aisel.aisnet.org
[4] Sagiroglu, S and Sinanc D., Big Data: A Review, International Conference on Collaboration Technologies and Systems (CTS), pp. 42-47, 20-24 May 2013
[5] http://www.domo.com/blog/2014/04/data-never- sleeps-2-0/
[6] Garlasu, D.Sandulescu, V. Halcu, I. and Neculoiu, G., A Big Data implementation based on Grid Computing, 17-19 Jan. 2013
[7] Neil Raden, Big Data Analytics Architecture, Hired Brains Inc, 2012
[8] Han J. and Kamber M, Data Mining: Concepts and Techniques, Morgan Kaufmann Publishers,San Francisco, 2000.
[9] Ankita S. Tiwarkhede and Vinit Kakde, A Review Paper on Big Data Analytics, IJSR, Volume 4 Issue 4, April 2015
[10] Han Hu, Yongyang Nen, Tat Seng Chua, Xuelong Li, Towards Scalable System for Big Data Analytics: A Technology Tutorial, IEEE Access, Volume 2, Page No 653, June 2014.
[11] Wei Fan and Albert Bifet, Mining Big Data: Current Status, and Forecast to the Future, SIGKDD Explorations, Volume 14, Issue 2, 2012.
[12] S.Vikram Phaneendra and E.Madhusudhan Reddy, Big Data- solutions for RDBMS problems- A survey, IEEE/IFIP Network Operations & Management Symposium (NOMS 2010),Osaka Japan, Apr 19-23 2013.
[13] Shilpa and Manjit Kaur, Big Data and Methodology-A review, IJARCSSE, Volume 3, Issue 10, October 2013.
[14] Kishor, D., Big Data: The New Challenges in Data Mining, IJIRCST, 1(2), pp. 39-42, 2013.
[15] Dheeraj Agarwal, A comprehensive study of data mining and applications, IJARCET, Vol , issue 1, January 2013.
[16] Sagiroglu, S. and Sinanc, D., Big Data: A Review, International Conference on Collaboration Technologies and Systems (CTS), pp. 42-47, 20-24, May 2013.
[17] Richa Gupta, Sunny Gupta and Anuradha Singhal, Big Data : Overview, IJCTT, Vol 9, Number 5, March 2014.
[18] Xindong Wu, Gong-Quing Wu and Wei Ding, Data Mining with Big Data, Jan 2014.
[19] Bharti Thakur and Manish Mann, Data Mining for Big Data: A Review, IJARCSSE, Volume 4, Issue 5, May 2014.
[20] Sabia and Sheetal Kalra, Application Of Big Data: Current Status and Future Scope, IJACTE, Vol 3, Issue 5, 2014.
[21] Vatsal Shah, Big Video Data Analytics using Hadoop, IJARCSSE, Vol 5, issue 7, July 2015.
[22] James Manyika, Michael Chui, Brad Brown, Jacques Bhuhin, Richard Dobbs, Charles Roxburgh and Angela Hungh Byers, Big Data: The next frontier for innovation, competition and productivity, June 2011.
[23] Gautam Shroff, Lipika Dey and Puneet Agrawal, Social Business Intelligence Using Big Dat , 2013.
[24] Hardeep Kaur, A Review of Applications of Data Mining in the Field of Education, IJARCCE, Vol. 4, Issue 4, April 2015.
[25] Feng Chen, Pan Deng, Jiafu Wan, Daqiang Zhang, Athanasios V. Vasilakos and Xiaohui Rong, Data Mining for the Internet of Things: Literature Review and Challenges", Hindawi Publishing Corporation, International Journal of Distributed Sensor Networks, Article ID 431047, March 2015.
[26] Fotis Aisopos, George Papadakis and Theodora Varvariqos, Sentiment analysis of social media content using N-Gram graphs , WSM'11 Proceeding of the 3rd ACM SIGMM international workshop on Social Media, Nov'28-Dec'01, 2011.
[27] Bingwei Liu, Erik Blasch, Yu Chen, Dan Shen and Genshe Chen, Scalable Sentiment Classification for Big Data Analysis Using Naıve Bayes Classifier, 2013
[28] Jalaj S. Modha, Gayatri S. Pandi and Sandip J. Modha, Automatic Sentiment Analysis for Unstructured Data, IJARCSSE, Volume 3, Issue 12, December 2013.
[29] Ahmed Elragal and Moutaz Haddara, Big Data Analytics: A Text Mining Based Literature Analysis , NOKOBIT, Vol 22, Issue 1, ISSN 1892-0748, 2014.
[30] Sunil B. Mane, Yashwant Sawant and Saif Kazi, Vaibhav Shinde, Real Time Sentiment Analysis of Twitter Data Using Hadoop, IJCSIT, Vol. 5 (3), 2014, 3098 – 3100.
[31] Somesh S Chavadi and Dr. Asha T, Text Mining Approach for Big Data Analysis Using Clustering and Classification Methodologies, IJETAE, Volume 4, Issue 8, August 2014.
[32] Ashish R. Jagdale, Kavita V. Sonawane & Shamsuddin S. Khan, Data Mining and Data Pre-processing for Big Data", Vol 5, Issue 7, July 2014.
[33] Supraja.G.S, Dr Jharna Majumdar and Shilpa Ankalaki, A Big Data Methodology for Sentiment Analysis of Twitter Data, IJIRCCE, Vol. 3, Issue 7, July 2015.
[34] Uma Gurav and Nandini sidnal, Opinion mining for reputation evaluation on unstructured Big Data, IJARCET, Vol 4 Issue 4, April 2015.
[35] V.Jude Nirmal and D.I. George Amalarethinam, Parallel Implementation of Big Data Pre-Processing Algorithms for Sentiment Analysis of Social Networking Data", IJFMA Vol. 6, No. 2, 2015, 149-159, ISSN: 2320 –3242 (P), 2320 –3250 (online)Published on 22 January 2015.
[36] Catherine Dwyer, Yi Zhang and Starr Roxanne Hiltz, Using Web Analytics to Measure the Activity in a Research-Oriented Online Community, Paper Submit to the Tenth Americas Conference on Information Systems, New York, , August 2004, pp 2668-78.
[37] Alexander W. Schneider and Boltzmannstr, Using Web Analytics Data to support Social Software Users, 2010, srvmatthes5.in.tum.de
[38] Sana Siddiqui and Imran Qadri , Mining Web Log Files for Web Analytics and Usage Patterns to Improve Web Organization, Siddiqui et al., IJARCSSE 4(6), June - 2014, pp. 794-802
[39] Daniel A. Keim, Florian Mansmann, Jorn Schneidewind, Jim Thomas, and Hartmut Ziegler, "Visual Analytics: Scope and Challenges", http://nbn-resolving.de/urn:nbn:de:bsz:352-opus- 68426, Lecture notes in computer science, No. 4404(2008) pp. 76- 90.
[40] R. Balaji Ganesh & S. Appavu, "An Intelligent Video Surveillance Framework with Big Data Management for Indian Road Traffic System", IJCA (0975 – 8887) Volume 123 – No.10, August 2015
[41] Vatsal Shah, Big Video Data Analytics using Hadoop IJARCSSE, Vol 5, issue 7, July 2015.
[42] Punam Bedi & Vinita Jindal, Use of Big Data Technology in Vehicular Ad-hoc Networks, http://www.researchgate.net/publication/266148960, 2014.