International Research Journal of Computer Science (IRJCS)
Issue 06, Volume 6 (June 2019)
SPECIAL ISSUE - 5th International Conference - "ACCE – 2019"
ISSN: 2393-9842
www.irjcs.com

A REVIEW PAPERBASED ON BIG DATA ANALYTICS

Rashmi
Department of Computer Science and Engineering,
Srinivas University, Mukka, Mangalore, India
hiremathrashmi8@gmail.com
B.R Kishore
Professor & HoD, Computer science and Engineering
Srinivas University, Mukka, Mangalore, India
cshod@srinivasgroup.com

Manuscript History
Number: **IRJCS/RS/Vol.06/Issue06/JNCS10083**
Received: 29, May 2019
Final Correction: 30, May 2019
Final Accepted: 02, June 2019
Published: June 2019
doi://10.26562/IRJCS.2019.JNCS10083
Editor: Dr.A.Arul L.S, Chief Editor, IRJCS, AM Publications, India

Abstract-In recent times the volume of data produced is in tremendous amount which is of the size varying from terabytes to zettabytes and with the data sets including structured, semi-structured and unstructured data which is called as big data. Data is generated from the different sources like social media, sensors, transactional applications, video/audio, networks etc. It is important to extract a useful data from a big data by using the processing framework and to analyse it so as to achieve benefits in business, better customer service, more effective marketing. The objective of this paper is to inspect the big data at various stages. It describes the characteristics of big data. It describes the history and evolution of big data analytics. It describes the types of technologies to extract the valuable information. It helps the researchers to find the solution by considering the challenges and the issues.

Keywords— Big data, Structured data, Unstructured data, Hadoop, Big data analytics.

I. INTRODUCTION

(1.1) Overview

The term "big data" can be defined as data that becomes so large that it cannot be processed using conventional methods. The hugeness of the data which can be viewed to be big data is a persistently changing factor and newer tools are continually changing factor and newer tools are continually being developed in order to handle this big data. Data is being generated at massive amount. In fact 90% of the data in the world today was produced in last two years.The complex process of examining large and varied big data to uncover data including hidden patterns, unknown correlations, market trends and customer preferences that can help organizations make informed business decisions hence called as big data analytics. Generally, Data warehouses have been used to store the large dataset. In this case extracting the precise intelligence from the available big data is a major concern. Most of the presented approaches in data mining are not usually able to manage the large datasets successfully. The key issuein the analysing the big data is the absence of coordination betweendatabasesystemsmoreoverwithanalysistoolslike data mining and statistical analysis. These challenges generally emerge when we wish toper form knowledge discovery and representation for its practical applications. An underlying issue is how to quantitatively illustrate the essential characteristics of big data. There is a need for epistemological implications in describing data revolution [2]. Moreover, the study on complexity theory of big data will help understand essential characteristics andformation of complex patterns in big data, simplify its representation,Gets better knowledge abstraction, and guide the design of computing models and algorithms on big data [1]. Much research was carried out by various researchers on big data and its trends [3], [4], [5]. However, it is to be noted that all data available in the form of big data is not useful for analysis or decision making process. This paper focuses on characteristics and challenges in big data and its available technologies.

(1.2) History and evolution of big data analytics

The concept of big data has been around for the years. Most of the organizations now understand that if they capture all the data that streams into their businesses, they can apply analytics and get significant value from it. But even in the 1950s, decades before anyone come out with the term "big data," businesses were using basic analytics (necessarily numbers in a spreadsheet that were manually examined) to uncover insights and trends. The new benefits that big data analytics brings to the table are however speed and efficient. Whereas a few years ago a business would have gathered information, run analytics and unearthed information that could be used for future decisions, today that business can identify the perceptions for immediate decisions.

II. TYPES OF BIG DATA

1. Structured Data

Structured Data refers to the data which is already stored in databases, in an ordered form. It interprets for about 20% of the total existing data, and is used the most in programming and computer-related activities. There are two sources of structured data such as machines and humans. All the data received from sensors, web logs and financial systems are grouped under machine-generated data. This comprise of GPS data,medical devices, data of usage statistics collected by servers and applications and the massive amount of information that regularly move through trading platforms. Human-generated structured data mainly incorporatesall the data a human inputs to a computer, example his name and other personal details. When a person clicks a link on the internet, or even makes a move in a game, data is created- this can be used by companies to figure out their customer behaviour and make the appropriate decisions and modifications.

2. Unstructured data

Structured data resides in the conventional row-column databases, unstructured data have no clear format in storage. The rest of the data generated, about 80% of the total account for unstructured big data. Most of the data a person encounters belongs to this category and till not long ago, there was not much to do to it except storing it and analysing it manually. Unstructured data is also classified based on its source, into machine-generated or human-generated. Machine-generated data relates for all the satellite images, the scientific data from various experiments and radar data collected by different facets of technology. Human-generated unstructured data is produced abundance across the internet, since it consists social media data, mobile data and website content. This means that the pictures we upload to our Facebook or Instagram handles, the videos we watch on YouTube and even the text messages we send all contribute to the nemerous heap that is unstructured data.

3. Semi-structured data

The line between unstructured data and semi-structured data has always been unclear, since most of the semi-structured data appear to be unstructured when looked briefly. Information that is not in the traditional database format as structured data, but contain some organizational values which make it easier to process, are included in semi-structured data. For example, NoSQL documents are considered to be semi-structured, since they contain keywords that can be used to process the document easily. Big Data analysis has been found to have a definite business value, as its analysis and processing can help a company achieve cost reductions and considerable growth. So it is essential that you do not wait too long to exploit the potential of this excellent business opportunity.

III. CHARACTERISTICS OF BIG DATA

In order to make sense out of this overwhelming amount of data it is often broken down using five V's: Velocity, Volume, Value, Variety, and Veracity.

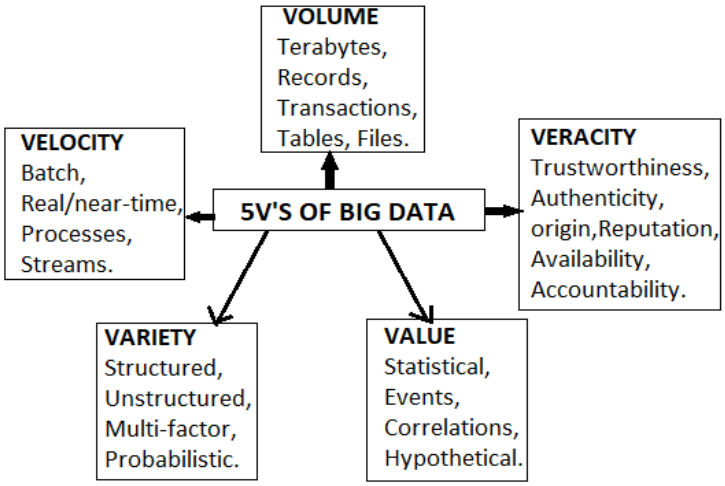

Fig. 1Characteristics of Big Data

3.1. Velocity

Velocity refers to the speed at which massive amounts of data are being produced, collected and analyzed. Every day the number of emails, twitter messages, photos, video clips, etc. increases at quick speeds around the world. Every second of every day data is increasing. Not only must it be analyzed, but the speed of transmission, and access to the data must also remain immediate to allow for real-time access to website, credit card verification and instant messaging.

3.2. Volume

Volume means the incredible amounts of data produced each second from social media, cell phones, cars, credit cards, M2M sensors, photographs, video, etc. The extensive amounts of data have become so large in fact that we are no longer able to store and analyse data using traditional database technology. We now use distributed systems, where parts of the data is stored in various locations and brought together by software. With just Facebook alone there are 10 billion messages, 4.5 billion times that the "like" button is pressed, and over 350 million new pictures are uploaded every day. Collecting and analyzing this data is clearly an engineering challenge of tremendously vast proportions.

3.3. Value

Value refers to the worth of the data being extracted. Having endless amounts of data is one thing, but unless it has value it is useless. While there is a clear link between data and insights, this does not always mean there is value in Big Data. The most important part to begin with big data initiative is to understand the costs and benefits of collecting and analyzing the data to ensure that ultimately the data that is gathered can be analysed.

3.4. Variety

Variety refers to the various types of data we can now use. Data today looks very different than data from the past. We no longer just have structured data such as name, phone number, address, financials, etc. That fits exactly and neatly into a data table. Today's data is unstructured. In fact, 80% of all the world's data fits into this category, including photos, video sequences, social media updates, etc. New and innovative big data technology is now allowing structured and unstructured data to be harvested, stored, and used simultaneously.

3.5. Veracity

Veracity refers to the quality or trustworthiness of the data. It defines the data accuracy. For example, think about all the Twitter posts with hash tags, abbreviations, typos, etc., and the reliability and accuracy of all that content. Extracting loads and loads of data is of no use if the quality or trustworthiness is not accurate. Another good example of this relates to the use of GPS data. Satellite signals are lost as they bounce off tall buildings or other structures. When this happens, location data has to be fused with another data source like road data, or data from an accelerometer to provide accurate data.

IV. CHALLENGES IN BIG DATA ANALYTICS

In recenttimes big data has been gathered in various areas such as public administration, healthcare, retail, bio-chemistry, and other interdisciplinary scientific researches. Usually web-based applications confront big data like social computing, internet text and documents, and internet search indexing. Social computing comprise of social network analysis, online communities, and prediction markets although internet search indexing includes ISI, IEEE, XPlorer, Scopus, Thomson Reuters etc. Taking into an account these advantages of big data it provides a new chance in the knowledge processing tasks for the upcoming researchers. Although opportunities always follow some challenges. In order to handle the challenges we need to know various computational complexities, info security, and computational method to analyse big data. For example, many statistical methods that perform efficiently for small data size do not scale to huge data. Likewise, many computational techniques that perform well for small data face notable challenges in analysing big data. Different challenges that the health sector face was being researched by many researchers [6]. Hence, the challenges of big data analytics are categorized into four broad categories specifically data storage and analysis, Knowledge discovery and computational complexities, scalability, and visualization of data, and info security.

A) Data Storage and Analysis

The datasize in recent years has grown exponentially by several methods like mobile devices, aerial sensory technologies, remote sensing, radio frequency identification readers etc. This data is stored by spending much cost althoughfinallythey are ignored or deleted because of not having the sufficient space to store the data. Hence, the first challenge for big data analysis is storage means and higher input/output speed. In such cases, the accessing the data must be on the top priority for the knowledge discovery and representation. The main reason is being that, it must be accessed easily and promptly for further analysis. In past decades, analyst use hard disk drives to store data but, it slower random input/output execution than sequential input/output. To conquer this constraint, the concept of solid state drive (SSD) and phrase change memory (PCM) was developed. Although the available storage technologies cannot possess the required execution for processing big data. Other challenge with Big Data analysis is assigned to diversity of data. With the ever growing of datasets, data mining tasks has significantly expanded. When dealing with large datasets additional data reduction, data selection, feature selection is a necessary task.

This presents an outstanding challenge for researchers. Because, existing algorithms may not always respond in an appropriate time when dealing with this high dimensional data. In present days the major challenge is automation of process and developing new machine learning algorithms to ensure consistency.The primary consideration is clustering of large datasets that help in analysing the big data [11]. Recent technologies likeHadoop and mapReduce make it possible to collect huge amount of semi structured and unstructured data in a reasonable amount of time. The main challenge is how to efficiently analyse this kind of data for getting betterknowledge.One of thestandardmethod istotransform the semi structured or unstructured data into structured data, and then apply data mining algorithms to extract knowledge. A framework to analyse data was explained by Das and Kumar. The detailed explanation of data analysis for public tweets was also explained by Das et al in their paper. One of the greatchallengeinthisinstanceisto concentratemoreon designing storage systems and to elevate effective data analysis tools which provide guarantees on the output when the data comes from various sources. Moreover, designing machine learning algorithms to analysethe data is necessary for improving effectiveness and scalability.

B. Knowledge Discovery and Computational Complexities is a key concern in big data isKnowledge discovery and representation. It incorporates number of sub fields likeauthentication, archiving, management, preservation, information retrieval, and representation. There are many tools for knowledge discovery and representation like fuzzy set, rough set, soft set, near set [7], formal concept analysis, principal component analysisetc. Moreover many hybrid tools are also developed to process the real life problems but these techniques are not problem independent. Later some of these techniques may not suit for large datasets. Hence the data size keeps on increasing at a faster rate, the techniques and tools which are available may not be enough to process this kind of data for getting the useful data. The data warehouses and data marts are themost popular approach for management of large dataset. In order to store the data produced from operational systems data warehouses are used. While the data mart rests on a data warehouse and ease the analysis. Much computational complexities are required for the analysis of large data sets. The main concern is to handle inconsistencies and uncertainty in the datasets. Generally, systematic modelling of the computational complexity is used. It isdifficult to create a comprehensive mathematical system that is broadly applicable to big data .By understanding the complexities domain specific data analytics can be run easily. A series of these development could reproduce big data analytics for various areas. By using machine learning approach with the less memory requirement ample of research and survey has been carried out in this area. The key aim in this area of research is to reduce computational cost for processing and complexities. Although, at present big data analysis tools have poor performance in handling computational complexities, uncertainty and inconsistencies. Hence,it is a big challenge for developing techniques and technologies that can deal with computational complexities, uncertainties and inconsistencies in a effective manner.

C. Scalability and Visualization of Data

Scalability and security of big data analysis techniques one of the major challenge. To enhance the data analysis and to increase the speed of processes by Moor's law researchers have paid attention in last decades. To enhance the data analysis it is important to build sampling, on-line and multiresolution analysis techniques. Incremental techniques have a better scalability mechanism with respect to big data analysis. There is a dramatic shift taking place in the processer technology by the number of cores being increased because of the increase in the size of the CPU speeds. Hence by increasing the processor speeds parallel computing is being developed. Parallel computing is required in the real time applications like navigation, social networks, finance, internet search, timeliness etc. The aim of data visualization is to present them more appropriately using techniques of graph theory. The relationship between data with proper interpretation is provided by the graphical visualization. Although, online sellers like Flipkart, Amazon, e-buy have millions of users and billions of goods sold every month. Hence huge amount of data is generated. Some companies uses tablean tool for the visualization of big data. It has the ability to convert the huge and complex data into spontaneous pictures. This will help employees of the company to forever search the relevance, monitor customer feedback, and the sentiment analysis. Although, present data tools for visualization mainly have improper implementation in functionalities, scalability, and response in time. We can notice that big data have build many challenges for the developments of the hardware and software which cause parallel computing, cloud computing ,distributed computing, visualization process, scalability. In order to conquer this issue we have to coordinate more mathematical models to computer science.

D. Information Security

In order to secure the huge amount of the sensitive data the organizations apply different policies. Huge amount of data is analysed, coordinated and mined to get the meaningful patterns out of it. The great amount of security risk is incorporated with big data. Hence data security is becoming a issue for the big data analytics. The techniques such as authentication, authorization, and encryption can be used to increase the security of big data. Several security measures that big data applications come across are scale of network, verity of devices, real time security monitoring, and lack of intrusion system.[9]Big data security challenge has attracted the attention for security of information. Hence, it is important to pay attention towards developing a multilevel security policy model and prevention system. While much research work is followed for securing the big data still the development is required [9]. The big challenge is in developing a multi-level security, privacy preserved data modal for the big data.

V.BIG DATA TECHNOLOGIES

Organizing and manipulating of excess volumes of structured data, semi-structured data and unstructured data is the big data management. The focus of big data management is to make sure the quality of high level data and availability of data for business intelligence and big data analytics applications. There are diverse amount of tools which can be used for big data management from data acquisition, data storage to data visualization. This part outlines those tools and related tools. Few of the tools which are used for various objectives are explained below:

A. Data Analysis

1) Hadoop: Hadoop is an open source platform for treating big data and its analytics. It is user friendly and flexible to work with various data sources, by collecting different sources of data or fetching the data from a database to run processorintensive machine learning process [12]. This tool has various types of applications such as location based data from weather, traffic sensors and social media data.

2) Map Reduce: Map Reduce is the programming environment which permits larger job implementation scalability against group of server. Map Reduce implementation has two main tasks: Map task converts input dataset into a various set of value pairs. The Reduce task combines several outputs of the Map task to form reduced tuples.

3) Hive: It is the SQL-like bridges which provides predictable business applications to run SQL queries against a Hadoop cluster. It was developed earlier by Facebook, then it has been made open source software tool, and it is a high level perception of the Hadoop which allows all to make queries against data stored in a Hadoop storage medium just as if they were manipulating a conventional data store.

4) PIG: Pig is a high level scripting language which is used with Apache Hadoop. PIG contains of a Perl like language which allows the query execution over data stored on a Hadoop instead of a SQL.

5) Platform: It is the big data discovery and analytical tool. It is a platform which takes the user's queries into Hadoop jobs automatically, therefore creating an abstraction layer which can exploit to simplify and organize datasets by anyone.

6) Rapidminer: It's software platform evolved by the company of the same name which provides an integrated platform for machine learning, data mining, text mining, predictive analytics and business analytics. It is used for business and commercial applications also for research, education, training, rapid prototyping, and application development and supports all procedures of the data mining process including dataset preparation, validation, results visualization and optimization.

B. Storage Technologies

As the size of data extends in gigantic way, there is a need for efficient and effective storage technologies to handle the big data. The main advancements in this space are associated with data compression and storage virtualization.

1) HBase: hive is a NoSQL databases which runs on top of HDFS. Apache HBase is an open source NoSQL database environment which provides real time read and write access to those large databases. An HBase scale linearly to handle very large data sets with few billions of rows and millions of columns, and it easily combines data sources that use a wide variety of various structures and schemas [11]. HBase is natively integrated with Hadoop and it works seamlessly with access engine YARN.

2) SkyTree: It is a high-performance data analytics and machine learning platform which focuses specifically on big data analytics and handling. Machine learning, is a needed part of big data, because the high data volumes make the exploration manually. Automated data exploration approaches are too expensive.

3) NoSQL (Non- Relational Databases): NoSQL (Not only SQL) database, is also called as Not Only SQL, it is an approach to data administration and database design which is useful for the big volume of data sets in distributed background. The most popular NoSQL database is built using Apache Cassandra. Other NoSQL databases implementations includesSimpleDB, Google BigTable, Map Reduce, MemcacheDB, Cassandra, MongoDB and Voldemort. Companies which uses NoSQL include social Medias Netflix, LinkedIn, and Twitter.

C. Visualization Tools

There are many open source visualization tools available in market.Few of them are listed below.

1) R Tool: R is a well-known programming language and software tool for graphic and statistical computing based data visualization. It is supported by R Foundation for Statistical Computing. The R Tool is broadly used among statistical area and data miners for developing statistical software and data analysis.

2) Tableau: This is the tool is used to visualize the result in the form of charts, maps, graphs and many other graphics form. A desktop application is available for visual analytics.

3) Infogram: There are easy three steps processes in this tool used to select among many visual templates, differentiate this tool with additional visualizations like charts, map, videos and ready to share your visualization. It supports accounts for video/audio files publishers and for the journalists of research script publisher, branded policies for businesses and classroom accounts for educational projects.

4) ChartBlocks: Chart block is an easy-to-use free online tool it doesn't requires any complex coding, and creates visualizations from databases and spreadsheets.

5) Ember Charts: Ember chart tool is based on the framework called Ember.js framework and it uses D3.js under the hood. Ember Charts features scatter charts, bar, pie, time series. It is very well-designed and easy to use tool.

6) Tangle: Tangle is a data visualization tool beyond the visualization, allows the designers and program developers to generate reactive programs which gives a better understanding of data relationships.

D. Big Data and Other Technologies

This section describes few of the important techniques which are closely related to big data. They are listed as follows:

1) Association with Cloud Computing: Cloud computing is the technology which can be used to store large volume of data in web. The main aim of cloud computing is to use high level computing and huge volume resources under firm management, so as to arrange for big data applications with well-defined computing capability. The development of cloud computing provides solutions for the storage and processing of big data. The development of big data accelerates the enlargement of cloud computing.

2) Association with Internet of Things (IoT): There are large number of networking sensors are inserted into different IoT devices and machines around the world. Those implanted sensors may gain different types of data, like network communication data, geographical data, environmental data, astronomical data, and logistic data [10]. Since the sources of data gathered from various environments, IoT produced big data has various types of characteristics when it is compared with normal big data. This data has some special characteristics such as heterogeneity, variety, noise, and redundancy. An authenticated report from Intel Corporation says that big data in IoT has three classic characteristics. They are as follows: (i) Plentiful terminals producing massive data. (ii) Data produced by IoT is commonly semi structured or unstructured (iii) IoT data will be useful only if it is analyzed.

3) Association with Data Center: The data center is not only a standard for storage of data, also supports more responsibilities including gathering of data, processing of data, organizing data, and optimizing the data values and operations in the big data paradigm. Date center has huge volume of data that organizes and accomplishes data according to its objective. The development of big data provides better opportunities and challenges to data centers.

VI. BIG DATA APPLICATIONS

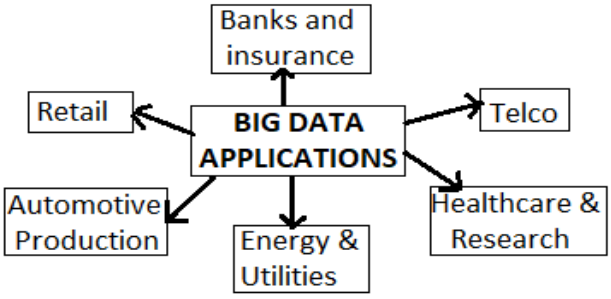

Fig. 2 Big Data Application Areas

1) Fraud Recognition and Control: Business operations face many types of fraudulent claims or transaction processing. Hence fraud recognition and control is most resounding big data application. In most cases, fraud is discovered long after the fact, at which point the loss has been done and all that's left is to minimize the harm and adjust policies to prevent it from happening again. Big data platforms that can verify, analyze, claims and transactions in real time, identifying large scale patterns across so many transactions or detecting inconsistent behaviour from an individual user, can change the fraud detection game.

2) Call Centre Analytics: Now we turn to the customer related big data application examples, in which call center data analytics are specifically powerful application. The current way of process in a customer's call center is often a great barometer and influencer of market sentiment, but without a big data solution, much of the awareness that a call center can provide will be ignored or revealed too late. Big data solutions can help ascertain recurring problems or customer and staff behaviour patterns on the fly not only by making intellect of time or quality resolution metrics, but also by capturing and processing call content itself.

3) Log Analytic in IT: IT departments and consultancies are generates a huge amount of logs and trace data. Without a big data solution, huge volume of the data may go unexamined. All organizations naturally do not have the source or manpower to agitate through all that information by hand, let alone in real time. With the help of big data solution, however both logs and trace data may be put to better use. Within this list of big data application examples, IT log analytics is the most largely applicable.

4) Social Media Analysis: Of the customer-facing Big Data application examples could discuss, analysis of social media activity is one of the most important. Everyone and their mothers are on social media these days, whether they like company pages on Facebook or tweeting complaints about products on Twitter. A big data solution built to produce and investigates social media activity, like IBM's Cognos Consumer Insights, a fact solution running on IBM's big Insights big data platform, may make the sense of the chatter. Social media data can provide real time insights into how the market is responding to products and campaigns. With those insights, companies can adjust their pricing, promotion, and campaign placement on the fly for optimal results.

5) Finance Analysis: Big Data analytics can be used to analyze the financial status and prediction in enterprises. For Example, the tool is analyzing the critical stock market moves and supports in making global financial prediction and decisions. Even though this is not a fool-proof process, it is definitely advancement in the field.

6) Agriculture: In agriculture, biotechnology centers use sensor data to enhance crop efficiency. It does test the crops and simulates to measure the plants reaction to various conditions. Its environment continuously adjusts to changes in the characteristics of various data including water level, temperature, growth, output, and gene sequencing of each and every plant in the testing environment called test bed.

VI. CONCLUSION

In recent years the data is generated at extremely higher rate in the volumes of zettabytes from various sources such as sensors/meters and activity records from electronic devices, social interactions, business transactions, electronic files, broadcastings. Analysing such a tremendous data is a big challenge. The paper describes the history and evolution of big data analytics, it describes the various technologies used for processing the big data. Big data analytics provide the ability to help enterprises understanding their business environments, their customer behaviour and needs and their competitor's activities. The paper also illustrates the challenges in the big data analytics and the big data analytics applications. Hence, analysing and storing such a huge data is a challenging area where the new technologies have to be developed.

REFRENCES

1. X. Jin, B. W.Wah, X. Cheng and Y. Wang, Significance and challenges of big data research, Big Data Research, 2(2) (2015), pp.59-64.
2. R. Kitchin, Big Data, new epistemologies and paradigm shifts, Big Data Society, 1(1) (2014), pp.1-12.
3. C. L. Philip, Q. Chen and C. Y. Zhang, Data-intensive applications, challenges, techniques and technologies: A survey on big data, Information Sciences, 275 (2014), pp.314-347.
4. 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.
5. S. Del. Rio, V. Lopez, J. M. Bentez and F. Herrera, On the use of mapreduce for imbalanced big data using random forest, Information Sciences, 285 (2014), pp.112-137.
6. MH. Kuo, T. Sahama, A. W. Kushniruk, E. M. Borycki and D. K. Grunwell, Health big data analytics: current perspectives, challenges and potential solutions, International Journal of Big Data Intelligence, 1 (2014), pp.114-126.
7. J. F.Peters, Near sets. General theory about nearness of objects, Applied Mathematical Sciences, 1(53) (2007), pp.2609-2629.
8. O. Y. Al-Jarrah, P. D. Yoo, S. Muhaidat, G. K. Karagiannidis and K. Taha, Efficient machine learning for big data: A review, Big Data Research, 2(3) (2015), pp.87-93.
9. Z. Hongjun, H. Wenning, H. Dengchao and M. Yuxing, Survey of research on information security in big data, Congresso da sociedadaBrasileira de Computacao, 2014, pp.1-6.
10. CheikhKacfahEmani, Nadine Cullot, Christophe Nicolle, Understandable Big Data: A survey, Mobile New Applications 2014, 171-209
11. Mike Barlow, Real-Time Big Data Analytics: Emerging Architecture, ISBN: 978-1-449-36421-2, 2013
12. Cheng-Long Ma , Xu-Feng Shang , Yu-Bo Yuan, International conference on machine learning and cybernetics, 2012, vol:4

