IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Vol.-6, Issue-12, Dec 2018 E-ISSN: 2347-2693

A Novel Framework for Big Data Analytics in Business Intelligence

Prashant Bhat[1], Prajna Hegde[2*], Pradnya Malaganve [3]

[1]School of Computational Sciences and IT, Garden City University, Bengaluru, India
[2,3]Research Scholar, School of Computational Sciences and IT, Garden City University, Bengaluru, India

[*]*Corresponding Author: prajna_rh@yahoo.co.in , Tel.: +91-9483807764*

Available online at: www.ijcseonline.org

Accepted: 26/Dec/2018, Published: 31/Dec/2018

***Abstract*— :** In recent years, due to new technologies databases are growing rapidly. This has resulted in evolution of term "Big Data". Big Data is nothing but huge data sets that can be processed and analysed to get useful information. Big Data Analytics is the process of inspecting large datasets, extracting information from it. These meaningful information obtained from large data sets can be utilized in business. Big Data Analytics helps business to take innovative, better decisions so as to improve business output. Big Data is large in size, grows very quickly hence impossible for traditional systems to process it. In this paper we review the importance of Big Data, Big Data Analytics and we propose an approach for using uncovered patterns, information of Big Data Analytics in Business Intelligence. Also an attempt is made in listing challenges in Big Data Analytics.

Keywords— Big Data, Business Intelligence, Big data Analytics, Data Clustering, Data Optimization, Classification

I. INTRODUCTION

The concept "Big Data" is getting important in recent years. Before 1950's business enterprises were using "analytics" to get insights and trends. They used to gather data, run analytics, get useful values that can be used to take business decisions. But today most of the firms capture stream of data that flows into their business. After that apply Big Data Analytics [14] so as to get values from it. The data sets that are extremely large in size, that can be analyzed to get useful patterns and trends is called Big Data. Because of its characteristics like volume, velocity and variety of data big data has got its importance in industry. Big Data is a large volume, complex in structure, streaming in very fast. Any organizations can utilize these data coming from different sources. Big data can add values in faster, better decision making in business. Techniques and tools of big data helps improvement in cost reduction, new products and services in business sector.

Big Data is large amount of data that comes from different sources. These enormous amount of data cannot be used directly anywhere as it's a raw data. In order to get enough use from it needs to be processed, analysed. There are basically three types of Big Data Analysis. 1.Descriptive Big Data Analytics. It is nothing but getting insight of what happened previously. It explains what happened in the past. 2. Diagnostic Big Data Analytics. It deals with why a particular thing happened in the past. It gives reason to a particular event. 3. Prescriptive Big Data Analytics. It is getting insight of what will happen in a particular situation. It involves predicting the future outcome using some mathematical and statistical proceedings.

The organization of the paper is as follows. Section I contain the introduction of Big Data and Big Data Analytics, Section II contain the review of Big Data Analytics in Business Intelligence, Section III contain the challenges of Big Data in Business Intelligence, Section IV contain the proposed framework for Big Data Analytics in Business Intelligence, Section V concludes the research work with future directions.

II. REVIEW OF BIG DATA ANALYTICS IN BUSINESS INTELLIGENCE

A. Big Data Analytics in Business Intelligence

The authors Zhaohao Sun, Lizhe Sun, Kenneth David Strang [1] proposed a paper which puts insights on the use of Big Data Analytics services in improving Business Intelligence. The authors presented nature of Big Data Analytics, Big Data Analytics Service Oriented Architecture (BASOA), and implementation of BASOA in Business Intelligence. Surveyed data shows BASOA is feasible for improving Business Intelligence and company information system. Authors also found that temporality, expect ability, relativity are the characters of intelligence in Business Intelligence. Both decision makers and customers expect these characteristics in BI for system, product, service of firm.

The authors Jafar Raza Alam, Asma Sajid, Ramzan Talib, Muneeb Niaz [2] presented a paper which states Big Data is a game changer for enterprises in achieving business advantages. Big data is receiving much attention in recent years. But there are some challenges which are causing decrease in growth of organizations. The reason for not implementing big data strategy is lack of knowledge about big data and its advantages.

The authors Pekka Pääkkönen, Daniel Pakkala [3] proposed a paper which says in recent years many business cases like LinkedIn, Facebook in social networking domain are making use of big d. Other Big Data use cases have concentrated on capturing the information from streaming of movies, network traffic monitoring, process improvement in industry. This paper presents technology independent reference architecture for Big Data systems. The authors have contributed in classification of relevant implementation technologies, product or services that are based on published use cases analysis and survey of associated work.

B. Big Data Analytics in Marketing Sector

Alexandre Borba, Ana Akemilkeda [4] stated Big Data can create greater impact on research and development as there is rapid increase in data generation, storage capacity, processing and analysing capacity. Big Data helps creating good understanding with customers hence it can be used in marketing field. This paper presents insights of Big Data in marketing information system. It also presents how Big Data helps decision making, some major concepts and new possibilities of using Big Data.

C. Big Data Analytics in Banking Sector

The authors Pingale Murali Manish, Sheetal Kasale, Anit Dani Simon [5] presented a paper which says after demonetization and digitalization India Banking System is undergoing huge change. Professionals have pressure of handling large amount of data every day. Many people who handle data have failed in using information which are in their own database. Big Data has characteristics like volume, velocity and variety. It can help ion improving risk management, quality, HR optimization by providing new possibilities for customer segmentation, monitoring the behaviour of clients, fraud detection. With the new possibilities and applications of new technologies Indian Banking Sector cover a way for efficient use of Big Data Analytics.

The authors Hossin Hassani, Xu Huang, Emmanual Silva [6] stated for banking sector finding the finding advanced Big Data Analytics tools like Data Mining techniques is a key that aims to reveal information from large volume of data to achieve improved strategic management and customer satisfaction. Up to date review on research status at the time about Data Mining and comprehensive review of Data Mining in banking is useful in providing future direction for research and development. This paper presents progression and most recent implementation of Data Mining post 2013. Also it helps in bringing insights related to future development of Data Mining and banking sector by collecting and analysing trends in research, data resources, technological aids and data analysis tools. Besides some obstacles are identified in this paper, summery for all those who are facing challenges of Data Mining are also presented. Srivastava, Utkarsh & Gopalkrishnan, Santosh [7] say Big Data is revolution in 21st century in banking firms as they have valuable data from many decades. It helped to know about money movements, prevent major disasters and thefts, and also helped to understand customer's behaviour. Big Data is useful to banks as they can extract information very quickly and easily from their data. Banks have started developing power of data to get utility across various spheres of their functioning, ranging from sentiment analysis, product cross selling, regulatory compliances management, reputational risk management, financial crime management etc. This paper intends to note how Big Data analytics can be used successfully in banking sector in following aspects: 1. Spending pattern of customers 2. Channel usages 3. Customer Segmentation and Profiling 4. Product Cross Selling based on the profiling to increase hit rate 5. Sentiment and feedback analysis 6. Security and fraud management. Here secondary data is used from banks in analysing primary nature. This paper reveals some of the best practices being adopted by banks globally, and can be implemented by Indian banks to improve their financial service offerings to customers.

III. CHALLENGES OF BIG DATA ANALYTICS IN BUSINESS INTELLIGENCE

Now a days, Big Data is a challenging concept in the field of technology. Data growing at the rapid speed is Big Data. Data generated from various sources in a huge volume creates Big Data. Data may be generated from social activity, government records, mobile phones, sensors, any scientific experiments etc. In this paper an attempt is made to list some of the major challenges that come on the way of adopting Big Data

A. Big Data Storage

Volume is one of the characteristic of Big Data [8], [13]. It consists of data in terms of exabytes. Storing exabytes of data in any device or in one system is not easy. It needs large storage as well as because of high volume extraction of data becomes a tedious task. Loading such a large amount of data into the storage , synchronizing data source and the storage system is a challenging task.

B. Acceptance of Data

Sometimes generated data are unnamed. Often, these unnamed data is confused to be unwanted data. Rather than amount of data, what companies do with the data matters more. Many organizations are lagging behind to adopt techniques of Big Data because of these complex structure of Big Data. Data continuously streams into the system. All these data associated with organization can be utilized to get some pattern. But it becomes difficult to select a part of data that are needed.

C. Processing of Data

Any business organizations are always interested in accurate, clean and sophisticated data rather than enormous amount of data with different formats. When data are in huge size it needs more time, advanced tools to process data. Therefore reducing, transforming, classifying [12], clustering of data becomes a challenging task specially when data are in huge size. Because of the complex structure it becomes difficult for traditional systems to process Big Data.

D. Management Complexity

Processed data must be visualized [9] to extract knowledge so that it can be utilized in decision making. This process needs skill and because of lack of experts it becomes difficult to get insight into the data and get advantage of data in Business Intelligence. As transactional data grows with high speed, it becomes difficult to keep track of data. Since structure of Big Data is complex, it creates lots of management difficulty.

E. Data Privacy and Security

Because of large volume of data companies cannot keep track of data in repository. Most of the data comes from internet and public data sets it is necessary to ensure that data collected is not violating privacy of customers [9]. Also Big Data analytics needs customers and other data. Organizations need to provide security [14] from network and external risks. It must also be noted that while gathering data from different resources, privacy of the customers must be maintained.

F. Integrating Different Data Resources

Data are gathered from different resources like censors, social media, web camera etc. Data must be integrated into single platform [13]. Synchronization of resources is a challenging task.

G. Lack of Experts

It needs skilled workers who have scientific vision along with business, customer understanding. The employees need to be updated with new emerging technologies [11] that can efficiently handle large amount of data. Since Big Data is in large volume lot of skill is required in handling it. It becomes challenging task for the companies to search for experts to handle Big Data.

It is highly difficult to adopt Big Data Analytics due to various reasons like handling growth of data, generating data in time, integrating and relating data, validating and securing data etc. Big Data with systematic approach and thought to thought view of structure can act as solution to many of the business problems. There are many Big Data Analytics tools like Apache Hadoop, NoSQL, MicrosoftHDInsight etc. which can be used to analyse large volume of data. Big Data with advanced tools and techniques can be used to get useful patterns that can be used in business to take proper decisions. With new emerging techniques [11] and advanced tools like Hadoop, Hive, PolyBase etc. Big Data can be implemented easily in future.

IV. PROPOSED FRAMEWORK

The proposed framework for Big Data Analytics is presented in Figure 1 and it consists of the following components.

A. Data Resources

Data comes from media which is the famous source of data. Now a days organizations are moving from traditional database system to cloud. Big Data may also be generated from IoT like censors, mobile phones etc.

B. *Data Storage*

Big Data generated from different sources are stored in a repository. It contains enormous amount of data. It may contain Data may be structured, semi structured and unstructured.

C. Data Preprocessing

Data collected from different sources cannot be processed directly. Preprocessing involves cleaning, editing, reduction, wrangling etc.

D. Big Data Mining Process

Data stored in repository is next processes in order to get some information. There are many technologies that can be applied on Big Data such as association which relates different data to get patterns, classification of data [12], clustering of data, decision trees, sequential patterns etc.

E. *Result Evaluation and Analysis*

Next step in the framework is evaluation of result and analysis of data that have been processed. Analysis may be descriptive, diagnostic or predictive [15] depending on the requirement.

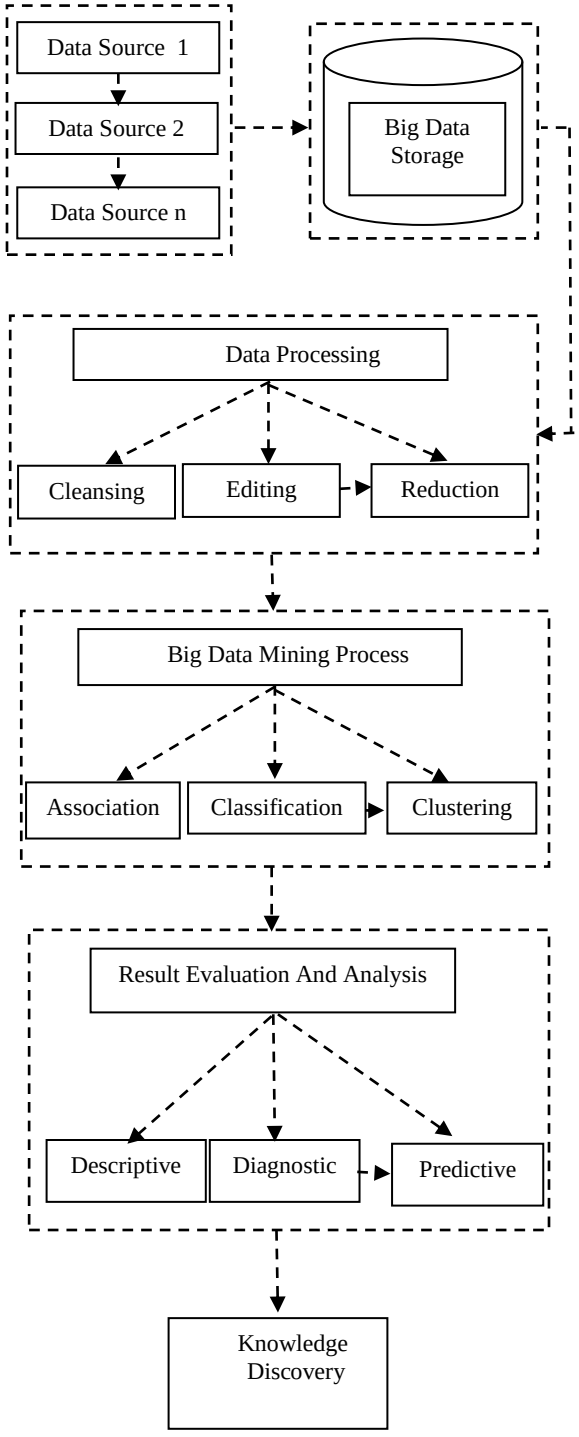

Figure 1: Proposed Framework for Big Data Analytics

F. Knowledge Discovery

Analyzed data are evaluated. These evaluated results gives answers to many of the business problems. The information discovered helps decision makers to take best decision for the scenario.

V. CONCLUSION AND FUTURE WORK

In this paper we reviewed Big Data concept, how it can be implemented in different sectors of business [10], how it helps to take business decisions. This paper describes a new framework for Big Data Analytics that can be implemented in business intelligence. Even though there are many Big Data techniques and tools [11] that can process and analyse Big Data, it is a challenging task to deal with data growth, system capacity, cost, validating data, securing data.

REFERENCES

[1] Sun, Zhaohao & Zou, Huasheng & Strang, Kenneth David. (2015), "*Big Data Analytics as a Service for Business Intelligence*", 58. 162-169. 10.1080/08874417.2016.1220239

[2] Jafar Raza Alam, Asma Sajid, Ramzan Talib, Muneeb Niaz, "*A Review on the Role of Big Data in Business* ", Joint Chiefs of Staff Message Center, Vol. 3, Issue. 4, April 2014, ISSN 2320–088X

[3] Pekka Pääkkönen, Daniel Pakkala, "*Reference Architecture and Classification of Technologies, Products and Services for Big Data Systems, Big Data Research*", Volume 2, Issue 4, 2015, Pages 166-186, ISSN 2214-5796

[4] Alexandre Borba, Ana Akemilkeda, "*Big Data Usage in the Marketing Information System*", Journal of data analysis and information processing 2014, 2, 77-85

[5] Pingale Murali Manish, Sheetal Kasale, Anit Dani Simon, "*Banking & Big Data Analytics*", OSR Journal of Business and Management (IOSR-JBM), e-ISSN: 2278-487X, p-ISSN: 2319-7668 PP 55-58

[6] Hossin Hassani, Xu Huang, Emmanual Silva, "*Digitalization and Big Data Mining in Banking*", MDPI, Received: 27 June 2018; Accepted: 17 July 2018; Published: 20 July 2018

[7] Srivastava, Utkarsh & Gopalkrishnan, Santosh, "*Impact of Big Data Analytics on Banking Sector: Learning for Indian Banks*", Procedia Computer Science. 50. 643-652. 10.1016/j.procs.2015.04.098

[8] Wani, Mudasir & Jabin, Suraiya , "*Big Data: Issues, Challenges and Techniques in Business Intelligence*", Conference: Conference: Proceedings of CSI - 50th Golden Jubilee Annual Convention, Springer (AISC Series), December 2015

[9] Althaf Rahaman.Sk, Sai Rajesh.K, .Girija Rani K, "*Challenging tools on Research Issues in Big Data Analytics*", 2018 IJEDR | Volume 6, Issue 1 | ISSN: 2321-9939

[10] Hsinchun Chen, Roger H. L. Chiang, Veda C. Storey, "*Business intelligence and analytics: from big data to big impact*", MIS Quarterly Vol. 36 No. 4/December 2012

[11] P. Soumya Sree Laxmi, P. Sree Pranathi, "*Impact of Big Data Analytics on Business Intelligence-Scope of Predictive Analytics*", E-ISSN 2277 –4106, P-ISSN 2347 5161, Accepted 22 March 2015, Available online 29 March 2015, Vol.5, No.2 (April 2015)

[12] Siddu P. Algur, Prashant Bhat, "*Web Video Mining: Metadata Predictive Analysis using Classification Techniques*", I.J Information Technology and Computer Science 2016 in MECS, DOI:10.5815/ijitcs.2016.02.09

[13] Taylor-Sakyi, Kevin. (2016). "*Big Data: Understanding Big Data*", Cornell University Library, arXiv

[14] Hugh J. Watson, "*Tutorial: Big Data Analytics: Concepts,Technologies, and Applications*", Communications of the Association for Information System, Volume 34, Article 65

[15] Zakir, Jasmine, (2015), "*Big Data Analytics*", Conference: International Association for Computer Information Systems, Volume 16

Authors Profile

Dr. Prashant Bhat is currently working as Assistant Professor and Research Supervisor in the School of Computational Sciences and IT, Garden City University, Bengaluru. He received Ph.D. in Computer Science from Rani Channamma University Belagavi, M.Sc. (Computer Science) from Karnataka University Dharwad and B.Sc. (Computer Science/Electronics/Mathematics) from Karnataka University Dharwad in the year 2017, 2012 and 2010 respectively. He has more than 5 years of experience in teaching and research. He is specialized in the area of Algorithms, Data Science, Web Mining Finite Automata and Theory of Computation, Big Data and Computer Networks. His research area is Data Mining, Web Mining, Multimedia Mining and Predictive Analytics. He published more than 33 research papers in various peer reviewed UGC recognized International Journals and conferences. He is a member of professional bodies like IAENG (International Association of Engineers in Computer Science and Data Mining), Member of Editorial/Review Board of International Journal of Application or Innovation in Engineering and Management (IJAIEM) and Universal Journal of Mathematics (UJM).

Ms. Prajna Hegde is pursuing Ph.D. in Computer Science at Garden City University, Bengaluru, Karnataka, India. She received BSc(Computer Science, Mathematics, Electronics) and MSc degrees from Karnatak University, Dharwad, Karnataka, India in 2012 & 2014 respectively. Her research interest includes Data Mining, Big Data, Web Data Mining, Web Multimedia Mining etc.

Ms Pradnya Malaganve is pursuing Ph.D. in Computer Science at Garden City University, Bengaluru, Karnataka, India. She received BCA degree from Karnatak University Dharwad, Karnataka, India in 2012. She received MSc degree from Rani Chennamma University, Belagavi, Karnataka, India in 2014. Her research interest includes Data Mining, Web Mining, Web Multimedia Mining etc.

