Proceedings of International Conference on Sustainable Development (ICSD 2019)
In Association with Novateur Publications IJIERT-ISSN No: 2394-3696
ISBN. No. 978-93-87901-05-6
February, 14th and 15th, 2019

ICSD/34

BIG DATA ANALYTICS: CHALLENGES, CAPABILITIES AND POTENTIAL BENEFITS

Ms. Asmita Namjoshi
Research Scholar, Department Of Computer Science, Tilak Maharashtra Vidyapeeth, Pune
asmitasrisri@yahoo.co.in

Abstract
Big Data sphere is one of the most endowed Information Technology spectrums with sizeable potential both on the surface of maturing economy and design swing in the domain of data storage administration and analytics. Nevertheless, nowadays, the altitude of intricacy achieved and the need of standardization of Big Data administration architectures stand for a mammoth barricade towards the espousal and implementation of analytics especially for those corporate and organizations not counting an adequate amount of competences and acquaintance. The occupied prospects of Big Data Analytics can be achieved through the definition of technologies that bring about Big Data users' outlook and requirements, also when these requirements are fuzzy and unclear. Under these grounds, authors discuss issues and challenges from the Data Analytics design and development perspective and explore possible opportunities to overcome these challenges.

Keywords: Big Data, Big Data analytics, Issue and challenges, Big data analytics, Big data architecture, Health care

1. Introduction

It is an undoubted fact that Big Data has haggard enormous interest from scientists in information sciences, policy and decision makers in governments and IT corporate. As the rate of information enlargement crosses Moore's Law at the commencement of this new century, unwarranted data is making huge dilemma to human beings. However, there are so much prospective and extremely helpful standards concealed in the gigantic amount of data. A novel methodical prototype is instinctive as data-intensive scientific discovery (DISD), also known as Big Data problems. A great number of application domains, ranging from economic and business actions to civic government, from nationwide defense to methodical researches in many domains, involve with Big Data problems. On the either side, Big Data is tremendously expensive to produce productivity in enterprises and revolutionary leap in scientific disciplines, which give us a bunch of opportunities to make immense advancements in numerous domains. There is no distrust that the prospect struggle in corporate yield and technologies will certainly join into the Big Data expedition. On the other hand, Big Data also emerges with many interrogations, such as constraints in data imprison, data storage, data investigation and data visualization. This paper is expected to exhibit a detailed sight about Big Data, including Big Data applications, Big Data opportunities and challenges, as well as the state-of-the-art techniques and technologies authors formerly take up to contract with the Big Data problems. Authors also talk about quite a few fundamental methodologies to hold the data cascade, for example, granular computing, cloud computing, bio-inspired computing, and quantum computing.

1.1 Challenges with reference to Deep Learning Techniques

Big Data Analytics and Deep Learning are two elevated-focal points of data science. Big Data has turn out to be significant as numerous organizations both communal and confidential have been gathering gigantic amounts of domain-specific information, which can hold constructive information about critical issues such as national intelligence, cyber security, fraud detection, marketing, and medical informatics. Organizations such as Google and Microsoft are analyzing great volumes of data for business analysis and judgments, affecting existing and future technology. Deep Learning algorithms take out high-level, multifaceted abstractions as data presentations through a hierarchical learning procedure. Multifaceted abstractions are learnt at a given level based on comparatively straightforward abstractions formulated in the preceding level in the ladder. A significant advantage of Deep Learning is the analysis and learning of colossal quantity of unsupervised data, making it a precious instrument for Big Data Analytics where raw data is largely unlabeled and un-categorized. In the study hand, authors discover how Deep Learning can be deployed for addressing some vital problems in Big Data Analytics, including retrieval of composite patterns from gigantic volumes of data, semantic indexing, data tagging, rapid information retrieval, and simplifying discerning tasks. Authors also examine some aspects of Deep Learning research that necessitate further looking at to integrate specific

challenges introduced by Big Data Analytics, including streaming data, high-dimensional data, scalability of models, and distributed computing [1].

1.2 Understanding challenges and benefits to Healthcare organizations

In continuation with this discussion, authors would like to attract reader's attention to benefits of data analytics in some promising applications including healthcare, especially to identify challenges. So far, health care industry has not fully utilized the impending reimbursement to be gained from big data analytics. While the persistently increasing corpse of academic study on big data analytics is frequently technology oriented, an enhanced understanding of the premeditated implications of big data is immediately needed. To deal with this need, this study scrutinizes the past expansion, architectural design and component functionalities of big data analytics. The analytics challenges address analytical potential for patterns of care, unstructured data analytical capability, decision support capability, predictive capability, and traceability. Researchers are required to map the reimbursement determined by big data analytics in terms of information technology (IT) infrastructure, operational, organizational, managerial and strategic areas. Authors, further advocate some strategies for organizations that are considering adopting big data analytics technologies. Author's findings will help healthcare organizations understand the big data analytics capabilities and potential benefits and support them to formulate more effective data-driven analytics strategies [2].

1.3 Understanding challenges in Data Management

Furthermore, to study challenges in developing Big Data applications has happened to ever more imperative in the last few years. In effect, quite a lot of organizations from diverse sectors depend ever more on information extracted from gigantic volumes of data. However, in Big Data framework, conventional data techniques and platforms are less well-organized. They demonstrate a sluggish receptiveness and lack of scalability, performance and accuracy. To visage the composite Big Data challenges, a large amount of efforts has been carried out. As a outcome, a variety of types of distributions and technologies have been urbanized. This paper is a evaluation that investigates up to date technologies developed for Big Data. It aims to assist to pick and take on the accurate amalgamation of different Big Data technologies according to their technological requirements and specific applications' rations. It presents not only a worldwide outlook of main Big Data technologies but also evaluates them according to different system layers such as Data Storage Layer, Data Processing Layer, Data Querying Layer, Data Access Layer and Management Layer. It sorts outs and discusses foremost technologies features, advantages, restrictions and usages [3].

1.4 Understanding challenges of systematic analysis and synthesis of data

To extend this discussion, authors now discuss potential valued insights of big data applications for enhanced decision-making method; have lately fascinated extensive interest from both academics and practitioners. Big Data Analytics is more and more flattering a trending practice that many organizations are adopting. The analytics procedure, together with the deployment and use of analytics tools, is observed by organizations as a instrument to enhance functional efficiency though it has strategic capability, compel new revenue streams and grow competitive recompense over business competitors. However, there are diverse types of analytic applications to believe. Consequently, preceding to rapid use and buying expensive Big Data tools, there is a call for for organizations to first understand the Big Data Analytics background. Known the noteworthy nature of the Big Data and Big Data Analytics, this paper presents a state-of-the-art assessment that presents a holistic view of the Big Data challenges and Big Data Analytics methods adopted by organizations to aid others appreciate this landscape with the goal of making vigorous decisions. In doing so, methodically analyzing and synthesizing the existing research published on Big Data and Big Data Analytics, domain. Exclusively, the authors look for answering and following two major questions: such as what are the different types of Big Data challenges speculated/suggested/affronted by corporate? and What are the diverse types of Big Data Analytics methods speculated/suggested/affronted to conquer Big Data challenges? [4].

This organized literature review is conceded out from end to end observing and understanding the historical trends and extant patterns/themes in the Big Data Analytics research domain, evaluating offerings, shortening knowledge, thereby identifying boundaries, implications and potential further research avenues to hold up the academic group of people in exploring research themes/patterns.

Thus, to outline the accomplishment of Big Data strategies extracted from the Scopus database. The investigation presented in this paper has acknowledged relevant Big Data research studies that have contributed both theoretically and experientially to the extension and accumulation of intellectual prosperity to the Big Data Analytics in technology and organizational resource management discipline.

1.5 Understanding challenges theorized / proposed / employed by organizations

Data Analytics using Deep learning is at present an exceptionally full of life research domain in machine learning and pattern recognition fraternity [5]. It has acquired enormous successes in an extensive area of applications such as speech recognition, computer vision,

and natural language processing. With the absolute dimension of data accessible today, brings huge opportunities and transformative potential for various sectors; on the other hand, it also presents extraordinary challenges to harnessing data and information. As the data keeps getting superior, deep learning is coming to occupy a key role in providing big data predictive analytics solutions. In this paper, authors provide a succinct impression of deep learning, and emphasize existing research endeavor and the challenges to big data, as well as the prospective trends.

1.6 Challenges in Big Data and Internet of Everything (IoE)

Furthermore as far as the market trends are concerned, Big Data has been playing an essential role in almost all fields such as healthcare, education, business organizations and scientific research. There is a sturdy association in Big Data and Internet of Everything (IoE). Universally, IoE applications are used to incarcerate or observe some specific standards to discover the unseen standards and take enhanced decisions. When the gadget is connected to the Internet, it always senses the specific parameters and stores those parameters into a associated data supplies. This would amplify the amount of the data stored in a data store. Hence, high end devices and scalable storage systems are desirable to store such enormous size of data. The amount of data to be stored and processed becomes an imperative problem in day to day life. Relational data base management system (RDBMS) is generally used to store the conventional data but day by day the volume, velocity and variety of sensor data is on the rise towards the Exabyte. This necessitates sophisticated tools and techniques to store, process and display such large amount of sensor data to the end users. Hence, Big Data tools are frequently used to practice such gigantic amounts of data. This would enlarge the market of the Big Data analytics [6].

1.7 Challenges in supply chain management

Authors now consider the fast rising attention from both academics and practitioners in the application of big data analytics in supply chain management (SCM) in order to explore new challenges [7]. This review responds to the need by proposing a narrative categorization structure that provides a full picture of modern literature on where and how Big Data Analytics has been functional within the Supply Chain Management context. The categorization structure is architecturally based on the content analysis method that addresses four research questions: (1) in what aspects of Supply Chain Management is Big Data Analytics being planned for deployment? (2) At what level of analytics is Big Data Analytics used in these aspects? (3) What types of Data Analytics models are used in Supply Chain Management? (4) What Data Analytics techniques are working to build up these models? The argument tackling these four questions reveals a number of research gaps, which guides in future research actions [7].

1.8 Challenges in Data Analytics for analytical and mathematical methods

To proceed with our discussion with reference to challenges of Data Analytics in Supply Chain Management, few issues have developed more dialogue lately. Given their acquaintance of analytical and mathematical methods, operations research (OR) scholars would appear well suspended to take a escort role in this discussion [8]. Regrettably, some have suggested there is a mis-match between the work of OR scholars and the requirements of enthusiastic managers, particularly those in the pasture of operations and supply chain management where data-centric decision-making is a main connstituent of most job descriptions. In this paper, authors endeavor to address this misalignment. Authors inspect both functional and intellectual applications of OR-centric big data analytical tools and techniques within an operations and supply chain management domain to emphasize their future potential in this domain. This paper contributes by providing recommendations for scholars, educators, and practitioners that aid to illustrate how OR can be instrumental in solving big data analytics problems in hold of operations and supply chain management [8].

2. Exploration of opportunities to overcome Data Analytics Challenges

Data is a treasured asset in the world today. The economics of data are constructed on the notion that data value can be extracted through the practice of analytics. Though Big data and analytics algorithms are yet in their preliminary evolution stage, their reputation cannot be underestimated. As big data begins to enlarge and grow, Prominence of big data analytics will endure to grow in everyday lives, both personal and professional. Moreover, the size and volume of data is increasing every single day, making it significant to address the way in which big data is addressed every day. According to investigations being accomplished many enterprises are initiating deployment of big data analytics in their day-to-day working. With intensify in admirations of Big data analytics, it is noticeable that capitalizing in this mediocre is what is going to shelter the future evolution of corporations and trademarks.

The crucial way to data assessment is Big Data Analytics and that is why it is foremost task to emphasize on that characteristic of analytics. Many corporations use diverse approaches to employ Big Data analytics and there is no mystic solution to magnificently instigating this. While data is vital, even more imperative is the development through which companies can gain insights with their help. Gaining insights from data is the goal of big data analytics and that is why investing in a system that can deliver those insights is extremely crucial and important. Successful implementation of big data analytics therefore requires a combination of skills, people and processes that can work in perfect synchronization with each other.

Today, establishments are emerging at a hasty pace and therefore encroachments in giant technologies. This means that brands must be prepared to model and acclimatize big data in such a manner that they become a fundamental feature of the information management and analytics substructure.

With astonishing potential, big data is today an emergent troublesome strength that is composed to become the next giant gadget in the arena of assimilated analytics, thereby transmuting the way in which brands and companies perform their duties across stages and economies.

With prodigious impending and prospects however come great challenges and obstacles. This means that corporations must be capable to crack all the apprehensive obstacles so that they can reveal the full potential of big data analytics and its concerned arenas. When big data analytics challenges are discussed in a suitable fashion, the accomplishment rate of executing big data clarifications automatically escalates. As big data styles its way into corporations and brands around the world, addressing these challenges are tremendously vital.

This review further consolidates the results of challenges explored in Big Data Analytics implementations that summarize some common issues. To achieve promising results, one needs to address these issues. With respect to some emerging applications, these issues are identified and presented collectively in figure 1.

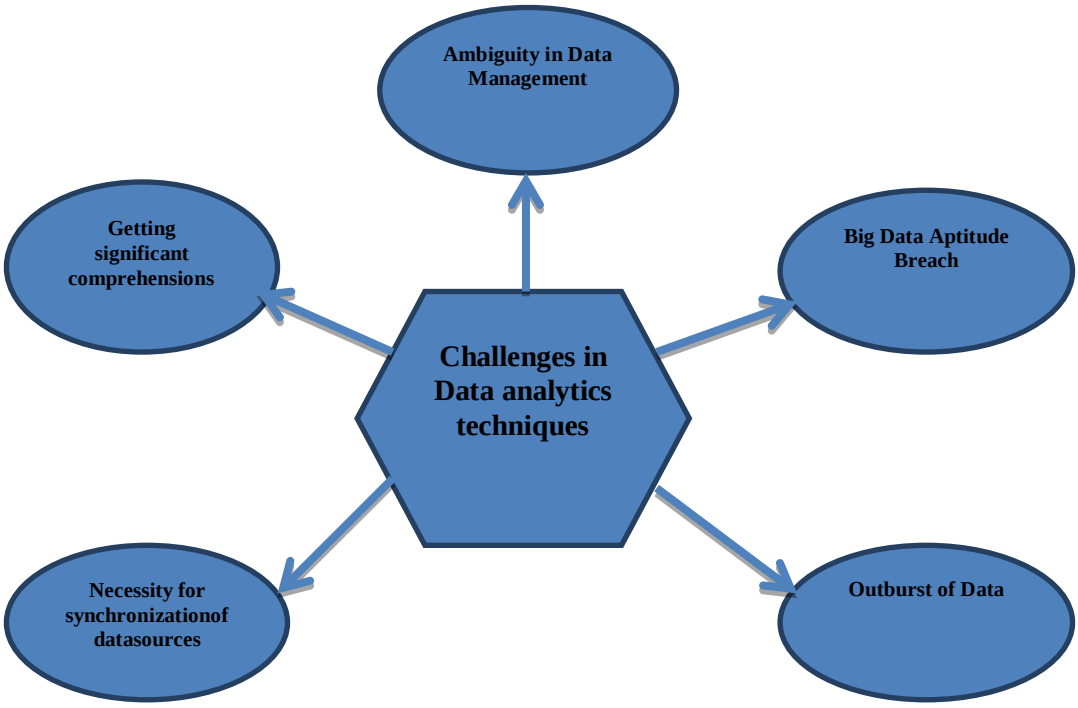

Figure1: Major challenges in Data Analytics

3. Summarization of discussion on challenges

Some of the key challenges that big data analytics drive is fronting today include the following:

3.1 Ambiguity of Data Management Infrastructure

Because big data is uninterruptedly escalating, there are new establishments and skills that are being industrialize devery day. A big challenge for corporations is to discover which technology works paramount for them without the outline of new risks and problems.

3.2 The Big Data Aptitude Breach

While Big Data is a increasing field, there are very scarce specialists existing in this field. This is because Big data is a multifaceted field and people who comprehend the intricacy and sophisticated nature of this field are very

rare and between. Another foremost challenge in the field is the talent gap that prevails in the industry.

3.3Outburst of Data

Corporations have to confront unbounded amount of data on a consistent basis. The scale and diversity of data that is accessible today can devastate any data experts and that is why it is imperative to make data approach ability modest and convenient for brand managers and owners.

3.4 Necessity for synchronization of data sources

As data sets become more assorted, there is a prerequisite to integrate them into an analytical podium. If this is overlooked, it can generate holes and lead to incorrect perceptions and communications.

3.5 Getting significant comprehensions through the usage of Big data analytics

It is vital that establishments achieve appropriate insights from big data analytics and it is essential that the correct A major challenge in the big data analytics is bridging this gap in an effective fashion.

4. Potential Opportunities

Now authors further explore some opportunities to face the challenges being discussed in various implementations. Consider for example, the Health care which is a data-rich industry. Administrative databases embrace an incredible number of transactions for each patient under treatment. The growth of the implementation of electronic health records with the help of Health Information Technology is mounting the quantity of data obtainable exponentially. At rest, the health care industry has been sluggish to influence the immeasurable data to advance care and health care operations [10]. The espousal of value-based procurement programs with augmented claim for organizations aggravated by the reasonably priced Care Act are moving both suppliers and consumers to use data to enhance operations. Health care's big data has the capability to revamp the process of health care with the most efficient and effective treatment pathways. Price-based procurement programs are incenting both health care providers and insurers to investigate new ways to leverage health care data to measure the quality and efficiency of care. The use of analytics in health care data presents of a number of daunting challenges, but also rich opportunities.

Predictive modeling can be used by the health care organizations for fraud deterrence [11]. Predictive modeling utilizes statistical techniques and past data to approximate the likelihood of potential results. Many systems are using these techniques to determine which health insurance claims are likely to be deceptive preceding the payment for the service.

Since many years, the largest challenge Big Data practitioners are facing in this domain with massive data is rendition of business to models. Since it is just an embryonic area, it is thorny even to define a problem painstakingly for example non technical staff can't converse the problem definition to the technical data scientists. A huge quantity of data is produced from structured, semi structured and unstructured sources which make it very hard to administer information about customers practice. Principally it's the necessitate of gigantic Shopping websites like Amazon and Flipkart to process enormous quantity of data so as to practice substantial datasets to find concealed patterns. These data have to be recovered, altered and encumbered to revise patterns related to customer conduct and their interactions. Companies can perk up the overall performance by using the customer feedbacks .investigation results from big data demonstrate them where they stand today in the spirited souk. Crowd sourcing is a proposal which is different from Big Data, crowd sourcing projects have exposed the prospective of using a wide group of genuine people to pull together useful and accurate data. There cannot be trustworthy data management without some human participation to ensure errorless environment and the administrators can without difficulty make their own mistakes and place their personal subjective incline on the filing of information. Big data corporate need crowd sourcing in their functioning to guarantee objectivity and assortment, thwart in opposition to errors more successfully and let social trends amuse a part in data examination. This amalgamation of big data and crowd sourcing provides a mode to contemporary data collection for customer experiences to make information more valuable, authentic and manageable. Big reimbursement can be reaped by coupling up crowd sourcing with big data.

Big Data encompasses all types of data which prevails today. From Electronic Health records and e-data to the overpowering quantity of government red tape information, which is digitized, there is necessity of efficient analytics techniques to be in place.

We can't classify Big Data with a single definition or portrayal, because we are yet operational on it. The grand obsession about information technology is that it has forever been on hand for technology enterprises, corporate and all types of organizations.

Surfacing of cloud computing has made it easier to endow with the best of technology in the most lucrative packages. Cloud computing not only condensed costs, but also prepared a broad assortment of applications accessible to the smaller enterprises. While the cloud is rising gradually we are also witnessing an detonation of information across the web. Social media is a entirely diverse world, where both advertisers and users engender heaps of data every day. This tremendous pouring in data eventually becomes difficult to manage. It appears that cloud computing together with big data could be an ideal permutation.

Jointly, they will grant a solution which is mutually scalable and compliant for big data and business analytics. The analytics benefit is going to be a massive benefit in today's world. Envisage all the information possessions which will become easily reachable.

Conclusion

A colossal warehouse of terabytes of data is generated each day from contemporary information systems and digital technologies such as Internet of Things and cloud computing. Investigation of these huge data requires a lot of labors at manifold levels to take out acquaintance for decision making. Therefore, big data analysis is a modern area of research and development. The indispensable purpose of this paper is to discover the potential impact of big data challenges, open research issues, and various tools associated with it. As a result, this paper provides a platform to explore big data at plentiful stages. in addition, it opens a new-fangled perspective for researchers to expand the solution, based on the challenges and open research issues. Authors wrap up by exploring insights into pertinent future mechanisms by posturing some questions, including defining data sampling criteria, domain adaptation modeling, defining criteria for obtaining useful data abstractions, enhancing semantic indexing, semi-supervised learning, and active learning. In addition, Big Data presents many unexplored also including large scale, heterogeneity, noisy labels, and non-stationary distribution. To comprehend the complete potential of Big Data, we therefore have to to tackle these technical challenges with innovative ways of philosophy and transformative solutions. We consider that these research challenges posed by Big Data are not only well-timed, but will also fetch plenty opportunities.

References

[1]. Maryam M Najafabadi, Flavio Villanustre, Taghi M Khoshgoftaar, Naeem Seliya, Randall Wald and Edin Muharemagic, Deep learning applications and challenges in big data analytics, Journal of Big Data (2015) 2:1 DOI 10.1186/s40537-014-0007-7.

[2]. Yichuan Wang a, LeeAnn Kung b , Terry Anthony Byrd, Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations, 2018, TFS-18447; No of Pages 11.

[3]. Ahmed Oussous , Fatima-Zahra Benjelloun, Ayoub Ait Lahcen, Samir Belfkih , Big Data technologies: A survey, Journal of King Saud University – Computer and Information Sciences 30 (2018) 431–448.

[4]. Uthayasankar Sivarajah , Muhammad Mustafa Kamal, Zahir Irani, Vishanth Weerakkody, Critical analysis of Big Data challenges and analytical methods, Journal of Business Research 70 (2017) 263–286.

[5]. XUE-WEN CHEN and XIAOTONG LIN, Big Data Deep Learning: Challenges and Perspectives, IEEE Acess, 2014, Digital Object Identifier 10.1109/ACCESS.2014.2325029.

[6]. Gunasekaran Manogaran , Chandu Thota and Daphne Lopez , Human-Computer Interaction With Big Data Analytics, HCI Challenges and Privacy Preservation in Big Data Security, 2018, Pages: 22, DOI: 10.4018/978-1-5225-2863-0.ch001.

[7]. Truong Nguyena , Li ZHOUa, Virginia Spieglerb , Petros Ieromonachoua , Yong Lina, Big data analytics in supply chain management: A state-of-the-art literature review, 2018.oone and Raymond R. Hill, Back in Business: Operations Research in support of Big Data analytics for Operations and supply Chain Management, November 2018, Volume 270, Issue 1-2, pp 201-211,

[8]. Martin Hilbert, Big Data for Development: A Review of Promises and Challenges, 13 December 2015, https://doi.org/10.1111/dpr.12142, Willey Online Library.

[9]. Susan White, A review of big data in health care's: challenges and opportunities, Open Access Bioinformatics 2014:6, 13–18, https://www.researchgate.net/publication/267865661.

[10]. Debi Prasanna Acharjya, A Survey on Big Data Analytics: Challenges, Open Research Issues and Tools,*in* International Journal of Advanced Computer Science and Applications · February 2016, DOI: 10.14569/IJACSA.2016.070267.

[11]. Laika Satish and, Norazah Yusof, A Review: Big Data Analytics for enhanced Customer Experiences with Crowd Sourcing, at 2nd International Conference on Computer Science and Computational Intelligence 2017, ICCSCI 2017, 13-14 October 2017, Bali, Indonesia.