

Underpinnings of Big Data Analytics and Its Applications

Anand Kumar Pandey¹, Rashmi Pandey², Ashish Tripathi³

¹CSA, ITM University, Gwalior, Madhya Pradesh, India

²CSE, ITM Group of Institute, Gwalior, Madhya Pradesh, India

³CSA, ITM University, Gwalior, Madhya Pradesh, India

ABSTRACT

Big data and Data Mining are co-related to each other and also emphasize the phenomena of extracting and analysis useful data from considerable database. The concept of Big Data analytics plays a very significant role in several fields, such as Data Mining, Education and Training, cloud computing, E-commerce, healthcare and life science, Banking and Agriculture. Big data Analytic is a technique for looking at big set of data to expose hidden patterns. A large amount of data is continuously generated every day using modern information system and technologies. As a result this paper provides a platform to investigate applications of big data at various stages. In future, it come forward to be a required for an analytical assessment of new developments in the big data technology. In addition, it also explores a new and suitable outlook for researchers to expand the solution, based on the literature survey, challenges, new ideas and open research issues.

Keywords : Data Analysis, Big Data Analytics, Big data Applications, E commerce, Descriptive Analysis.

I. INTRODUCTION

In the world of Big data, the Big data analytics has become an exceptionally fundamental element for growth of productivity, efficiency, reliability and the quality of services of big data sets. The concept of Big data analytics and its applications also plays a essential responsibility in different fields. Multifaceted and huge datasets have various types of different and important features that are closely in resemblance with “Big Data”. To administer these datasets is troublesome with the traditional information preparing frameworks [2]. Furthermore, our efforts also describe about different aspects about data valuation, data conversion, data prophecy, data stabbing, data analysis, data security, data privacy violations and allotment propose different climbing challenges that the “Big Data analytics” provides.

This review paper explained evolutionary descriptions in many useful fields with compilation of huge datasets. In general, this article refers to the gathering of huge and composite useful datasets which are complex to retrieve and compile using traditional database management tools or data processing techniques. These are available in different structured, semi-structured, and unstructured arrangement in petabytes and beyond [3]. The concept of big data analytics requires to helps operations to turn into more effective. Big data analytics technique like Hadoop helps to minimizing the cost of storage. Analyze data sets and describe conclusions about different data sets also illustrates the conceptual study that which application is much more suitable in which circumstances. Furthermore, it should be well-known that all kind of data available in any form of big data sets are not constructive for analysis or decision

making practice. There are so many challenges in Big data analytics, in recent years it has been accumulated in different domains like health care, public services, E-commerce, telecommunication, and other interdisciplinary scientific researches. Most of the data is collected through web based application, such as, social networking sites, online communities, survey, and cloud based networks. Sometimes this data needs to be analyze and sometime not.

II. AN OVERVIEW OF BIG DATA ANALYTICS

In current technological generation big data analytics become most popular research area and prominent field for IT industry and IT professionals. Big data analytics is the procedure of analytical study and mixed data sets i.e., to representation hidden patterns, unidentified correlations, advance trends, user choices and further helpful information which might assist to organizations make more-informed business decisions. Data analytics is the procedure of investigative and required data sets in order to depict conclusions regarding the information they include, increasingly with the support of specialized systems and software[1]. There are so many kinds of growing challenges in the big data world and most of the big data challenges and its related publications highlight on big data analysis and its concerned fields. Before discussing data analytics in detail, we should consider about all its relevant Contents used for its description, concepts and its objective, therefore we can visualize it with a innovative vision. As per the available research literature we try to manage a global overview of the Big Data Analytics as shown in following figure 1.

An overview of big data analytics is defined in above figure by the concept of three dimensions like reasonable, authorized and technological dimensions, but they also include the specified fields of data, processing and management, it consist the core and conceptual building blocks of data analytics [8].

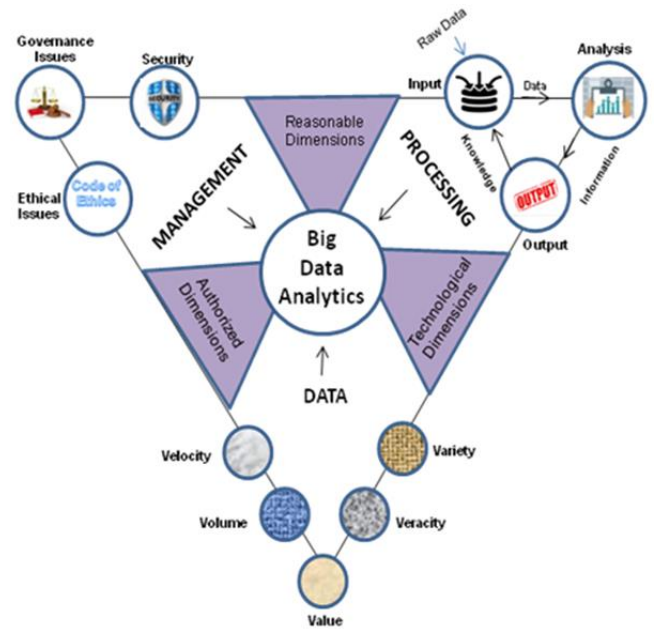


Figure 1. An Overview of Big Data Analytics

- **Data** is the fundamental part of big data analytics, it consists big quantity of data generated quickly from different sources in a variety of formats. The generated data might be non related, unpredictable, volatile and unreliable due to the probable uncertainties within them. It should be properly verified, validated, and value based.
- **Processing** is also known as Big Data stream processing, it should be always system intended. It becomes clear that partitioning a large datasets into smaller one can speed up the processing. Each organization should have its own way to model the data processing pipeline. It also consist the core functioning of data input, data analysis and data output. This model is very simple, at first incoming and only required raw data is received as an input, then this data is passed to next ready to use centre of the process section known as analysis In this analysis section the data is filter, transform, integrate, structures and discover useful information. Another and last section of processing pipeline is data output also known as post processing [1]. Here system apply system based knowledge in form of algorithm and evaluate the selected and useful information.

- **Management** is an managerial progression that includes acquiring, validating, storing, defending, and processing necessary data to make sure the convenience, consistency, and correctness of the data for its users.
- **Technological Dimensions** is concern with methods, algorithm and tools used to carry out tasks for the purpose of best technical aspects. It also emphasizes the various software and hardware architecture used for data storage and data processing.
- **Authorized Dimensions** might be implicit for the convergence of legal aspects, depends on input data and management. In this authorized viewpoint we try to focus on corresponding legal compliant regarding big data use and valuable use of the organization’s big data resources using appropriate software agent.
- **Reasonable Dimensions** also known as economic dimensions, it uses to revolve the contents of big data in to fruitful knowledge about the essential performance of the social and economical software agents. The reasonable dimensions be positioned between content delivery paradigm and new business model.

III. DIFFERENCE BETWEEN DATA ANALYSIS AND DATA ANALYTICS

In this section, we describe Data Analytics vs Data Analysis, we will look at their Meaning and Head to head Comparison. Sometimes, Analysis and analytics are treated as exchangeable term. The following variation between data analysis and data analytics is came from a eminent term of which data analysis generates a subcomponent[4].

Table 1. COMPARISON TABLE

S. N	Key Points	Data Analysis	Data Analytics
1	Concept	It is a simple process of	It is defined as “a process of

		identify, cleaning, transforming, and working out of the data and helps in decision-making.	transforming data into action through analysis
2	Cost	Less Expensive	More Expensive
3	Form	This is a precise form of data analytics.	This is simple type of analytics which is used with data pools.
4	Structure	It consist of defining a data, study, identifying, transforming the data to provide a meaningful conclusion.	It consists of data collection and scrutinize in general and it can have multiple users.
5	Sequence	It is sequence of data collection, gathering and retrieving.	It have some different sequence of steps like Data recognition, Data attainment, Data mining, Data Validation & Cleansing etc.
6	Tools	Most of the available Data analysis tools are Open source .	Only authentic and registered tools implement big Data Analytics.
7	Usage	It may be used in different domains for the basis of past experience.	It used in pre-defined patterns with some futuristic and beneficial approached.

IV. TYPES OF DATA ANALYTICS

Data analytics is a work of expertise data scientist. To effectively work with data, we need to choose any specific kind of analytic technique. There are five basic types of big data analytics, describes as follows.

A. Prescriptive Analytics

This is the most underused and important kind of analytics technique. It applied when we have to endorse an activity, so that the decision makers can use the data in this way. It needs less focus to provide required solutions. It helps to find out the finest solution in less time among multiple available choices[5]. This kind of analytics, generally applied to take decisions for customer related services.

It focused on optimal decisions and future situations. It uses simple rules to provide solutions for complex models.

B. Diagnostic Analytics

This type of analytics is used to apply when something is going to happen. When some trend or restrictions are set to take decisions then data scientist try to diagnose proper solutions. This analytics is mostly used in healthcare and life science applications.

For example during backward looking, target dependent variables, and during inferred explanatory situation of decisions.

C. Descriptive Analytics

This is the most straightforward kind of analytics. It also use some kind of investigation process to describe the proper analyze the set of data. It also suggests and provide to use historic data that use to analyze the proper solution.

This technique normally used forward looking, it focused on multiple prescribed descriptions, comparisons, and pattern detections, so that it is most time exhaustive technique.

D. Predictive Analytics

This is the most frequently used technique; predictive analytics use models to forecast what might happen in specific scenarios. It uses a variety of calculable, displaying, information mining, and machine learning strategies to study soon and verifiable information, along these lines permitting experts to make forecasts about what has to come.

This technique also used forward looking. It focused on non-discrete predictions of future states, used by most of the modern applications.

E. Outcome Analytics

This is also known as consumption analytics. This technique provides approaching into customer behaviour that drives exact outcomes. This analysis is meant to help you know your customers better and learn how they are interacting with your products and services.

This technique used backward looking, real time as well forward looking. It also Focused on utilization patterns and associated business outcomes.

V. APPLICATIONS OF BIG DATA ANALYTICS

In the world of data science the big data analytic have enormous applications. Some of the applications are real time operation based. We will also discover a variety of sources of Big Data along with Big data applications in different domains All these applications are classified in following groups shown in Fig 2.

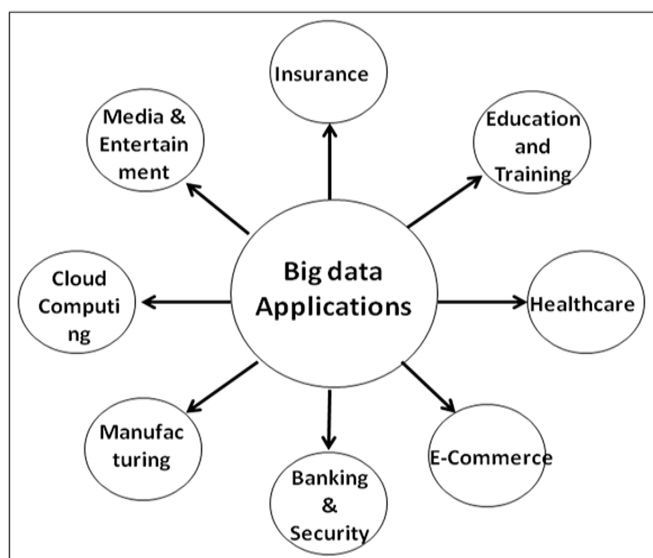


Figure 2. Applications of Big Data Analytics

A. Insurance

Big data analytics is very much useful in insurance industry. Big data is particularly promising and differentiating for insurance companies. With no physical products to manufacture, data is possibly one of their most important assets. Financial, actuarial, claims, risk, consumer, producer/wholesaler and many other types of data form the basis for practically every decision an insurer makes.

B. Education and Training

Education and training sector is enormous amount of data associated with different courses, students, training, and curriculum. These analytical study can be used to progress the operational effectiveness and working of educational institutes and in training sectors[6]. Following are some of the areas in the education and training industry that have been transformed by big data analytics.

C. Media and Entertainment

With the grow in digitization, media and entertainment companies are managing enormous amounts of data like never before. It has been expected that Facebook alone gathers and processes more than 500 TB of data. While the largest search engine in the world, Google handles 3.5 billion requests every day

wherein over 1000 computers are employed to tackle a single user query! Netflix, the most popular online media service provides almost 500 events that consume 1.3 PB/day and 8 million events that consume 24 GB/Second during peak time.

D. Healthcare

The healthcare and life science domains are connected to each other to take care the basic driving factors about minimizing the risk factor. The technique of data analytics acquires the disruptive growth and solutions about hidden and potential data[5]. This basic application involve in different domains in following kinds of expertise.

- Data Analysis for disease pattern and its occurrence.
- Patient care quality and treatment analysis.
- Data Analytics regarding drug discovery and development.
- Scientific studies based on real world data.
- Mix optimization for Market of pharmaceutical.
- Cost categorization and patient risk stratification.
- Big data Analytic driven activity master for patient index performance.

E. E-Commerce

The impact of Big data Analytics on E-commerce business is surprising. Using this analysis the organization can take better decisions. Projecting analytics is used by e-commerce to forecast what the consumer will buy.

- Enhance the shopping pattern analysis of customer.
- Best way of understanding customer's shopping behaviors.
- Analyze suitable offers at right time.
- The real time analytics can help to be aware about change of pricing.
- Forecast about market condition and easy to take decision.

F. Cloud Computing

The cloud can help us to process and analyse your big data faster, leading to insights that can improve your products and business. We can say that Big data and Cloud computing a perfect combination? Well, there are data points in support of it[4]. Besides that, there are also some real-time challenges to deal with. Merging big data with cloud computing is a powerful combination that can transform your organization.

G. Manufacturing

Big data analytics can help manufacturing companies in doing so by providing them perceptive insights from the information that is being collected from various sources. These sources can be their operational machines, databases keeping a log of the number of units produced, and employee records. Advanced big data analytics is a hot topic for the manufacturing industry. Manufacturers are generating vast amounts of data through their systems, but are they using it to optimize overall operations.

H. Banking and Security

Big Data has provided biggest chance to companies like Citi bank to see the big picture due to balancing the responsive nature of the data for delivering value to clients along with prioritizing the privacy and protection of information [7]. It has been fully adopted by many companies to drive business growth and enhance the services they provide to customers.

Big data analytics can aid banks in understanding customer behavior based on the inputs received from their investment patterns, shopping trends, inspiration to invest and personal or financial backgrounds.

VI.CONCLUSION

This paper introducing the basic thoughts of big data analytics discussed with the current applications and

its complexity encountered by huge literature survey. In this literature we have also identify the future aspects that could be burden in this field. Big data analytics cannot be complete and finish at any point in an technical field, where a considerable part of the research is yet to be completed. The data collection, analysis and testing in all relevant field is demonstrate every day. The idea of analyze data in proper form with diversity of data improvement is growing because of the expansion of sensor and mobile phones with simulated web applications. Data formed by this way, is the best resource for enterprises in significant business dealings polices. In recent years data are generated at a dramatic pace. Analyzing these data is challenging for a general man. The banking, e-commerce and Cloud computing were described to prepare and break down wonderful measure of data and it has malformed into the new Big Data model to obtain care of the on-demand administrations.

VII. REFERENCES

- [1]. Gandomi and M. Haider, Beyond the hype: Big data concepts, methods, and analytics, International Journal of Information Management, 35(2) (2015), pp.137-144.
- [2]. 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.
- [3]. S.Kuchipudi, T.S.Reddy, "Applications of Big data in Various Fields", International Journal of Computer Science and Information Technologies (IJCSIT), Vol.6, No.5, Pp.4629-4632, 2015.
- [4]. Safeeullah Soomro, " Big data Analytics and Its Applications" International Journal of Annals of Emerging Technologies (AETiC), Vol 1, No. 1, 2017.
- [5]. S.Mukherjee, R.Shaw, "Big Data–Concepts, Applications, Challenges and Future Scope" International Journal of Advanced Research in

- Computer and Communication Engineering, Vol.5, No.2, 2016.
- [6]. Rashmi Pandey and Anand Kumar Pandey ” Framework For MAS Based Big Data Analysis Using Agent Mobility”. In Proceedings of the sixth International IEEE conference on Computing for Sustainable Global Development, IndiaCom-2019.
- [7]. L.Venkata, S.Narayana, “A Survey on Challenges and Advantages in Big Data, International Journal of Computer Science and Technology Vol.6, No.2, 2015.
- [8]. T. Hansmann and P. Niemeyer, “Big data-characterizing an emerging research field using topic models,” in Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2014 IEEE/WIC/ACM International Joint Conferences on, vol. 1. IEEE, 2014.
- [9]. Anand Kumar Pandey and Rashmi Pandey “ Data Modelling and Performance Analysis Approach of Big data”, Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Amity University Rajasthan, Jaipur - India, February 26- 28, 2019

Cite this article as :

Anand Kumar Pandey, Rashmi Pandey, Ashish Tripathi, "Underpinnings of Big Data Analytics and Its Applications", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 6 Issue 2, pp. 75-81, March-April 2020. Available at doi : <https://doi.org/10.32628/CSEIT206222>
Journal URL : <http://ijsrcseit.com/CSEIT206222>