Different Types of Data Analytics using Big Data

Anuj Rapaka

Assistant Professor, Department of CSE, Shri Vishnu Engineering College for Women (A), Vishnupur, West Godavari District, Bhimavaram, Andhra Pradesh, Idnia

ABSTRACT

Various types of machine automated systems are generating huge amount of data in various forms like dynamic, Content Type format, audio type, video type, sensor type, and bio-metric format that emerges the term Big Data. In this article, we are discussing issues, challenges, and application of these types of Big Data with the consideration of big data volumes. Here we are analyzing multi-media data analytics, content-based analytics, Content-type analytics, audio-type, and video-type analytics their issues and various application areas. It will influence researchers to address these issues of storage management, and accessing of data known as Big Data. As well as the utilization of Big Data in India is also highlighted.

Keywords: Big Data, Big Data Analytics, Social Media Analytics, Content Based Analytics, Content Type Analytics, Audio Type Analytics and Video Type Analytics.

I. INTRODUCTION

The term big data is used to define the growth and the availability of huge amount of structured and unstructured data. Big data which are beyond the ability of commonly used software tools to create, manage, and process data within a suitable time. Big data[3] is important because the more data we collect the more accurate result we get and able to optimize business processes. The Big data is very important for business and society purpose. The data came from everywhere like sensors that used to gather climate information, available post or share data on the social media sites, Video Type movie Audio Type etc. This collection of information is called —[1] Big Data. Now a day, this big data is used in multiple ways to grow business and to know the world. In most enterprise scenarios the data is huge or it moves quickly or it exceeds current processing capacity. Big data has the potential to help companies improve operations and make quicker, more intelligent decisions. Big data generally includes data sets with sizes beyond the ability of commonly used software

tools to capture, curate, manage, and process data within a tolerable elapsed time. Big data is a set of techniques and technologies that require latest forms of integration to uncover large invisible values from data sets that are diverse, complex, and of a massive scale. WalMart[4] handles more than 2 million user transaction every hour. Facebook holds 50 billion information from its user base. Big data require some technology to efficiently process large quantities of data. It use some technology like, data fusion and integration, genetic algorithms, machine learning, and signal processing, simulation, natural language processing, time series Analytics and visualization.

II. BIG DATA ANALYTICS

Big Data Analytics refers to the process of collecting, organizing, analyzing large data sets to discover various patterns and other useful information. Big data analytics is a set of technologies and techniques that require latest forms of integration to disclose large invisible values from large datasets[14] that are different from the usual ones, more complex, and of a

large enormous scale. It mainly focuses on solving latest issues or old issues in better and effective ways. The main goal of the big data analytic is to help organization to make better business decision, upcoming prediction, analysis large numbers of transactions that done in organization and update the form of data that organization is used. Examples of big data Analytics are big online business website like flip kart, snap deal uses Facebook or Gmail data to view the user information or behavior.

Analyzing big data allows analysts, researchers, and business users to make better and quicker decisions using data that was previously inaccessible or unusable. Using state-of-the-art analytics techniques such as Content Type analytics, machine learning, predictive analytics, data mining, statistics, and natural language processing, businesses can study previously untapped data sources independent or together with their existing enterprise data to gain latest insights resulting in significantly better and quicker decisions. It helps us to uncover hidden patterns, unknown correlations, market trends, user preferences etc. It leads us to more effective marketing, revenue opportunities, better user service etc. Big Data can be analyzed through predictive analytics, Content Type analytics, statistical analytics and data mining. Types of big data analytics are:

a) **Prescriptive:** - This type of analytics help to determine what actions should be taken. It very valuable but not used largely. It focuses on answer specific question like, hospital management, diagnosis of cancer patients, diabetes patients that determine where to focus treatment.

b) **Predictive:** - This type of analytics helps to predict upcoming or what might be happen. For example some companies use predictive analytics to take decision for sales, marketing, production, etc.

c) **Diagnostic:** - In this type look at past and study the situation what happen in past and why it happen. And how we can conquer this situation. For example weather prediction, user behavioral analysis etc.

d) **Descriptive:**-It describes what is happening currently and prediction near upcoming. For

example market analysis, compactions behavioral analysis etc. By using appropriate analytics organization can increase sales, increase user service, and can improve operations. Predictive Analytics allow organizations to make better and quicker decisions.

III. PREDICTIVE ANALYTICS

Predictive Analytics is a method through which we can extract information from existing data sets to predict upcoming outcomes and trends and also determine patterns. It does not tell us what will happen in upcoming. It forecasts what might happen in upcoming with acceptable level of reliability. It also includes what if-then-else scenarios and threat assessment. Applications areas of

a. **Clinical Decision Support:** Predictive Analytics helps us to determine that which patients are at threat of developing certain conditions like diabetes, asthma, lifetime illness etc. Collection Analytics: Predictive Analytics helps financial institutions for the allocation for collecting resources by finding most effective collection agencies, contact strategies etc. to each user.

b. **Cross Sell:** An Organization that offers multiple products, Predictive Analytics can help to study user's spending, their behavior etc. This can help to lead cross sales that mean selling additional products to current users.

c. **User Retention:** As the number of competing services is increasing, businesses should continuously focus on maintaining user satisfaction, rewarding loyal users and minimize user reduction. If Predictive Analytics is properly applied, it can lead to active retention strategy by frequently examining user's usage, spending and behavior patterns.

d. **Direct marketing:** When marketing consumer products and services, there is the challenge of keeping up with competing products and consumer behavior. Apart from finding prospects, predictive analytics can also help to find the most effective combination of product versions, marketing material,

communication channels and timing that should be used to target a given consumer.

e. **Fraud detection:** Fraud is a big problem for many businesses and can be of various types: inaccurate credit applications, fraudulent transactions (both offline and online), identity thefts and false insurance. These issues plague firms of different sizes in many industries. Some examples of likely victims are credit card issuers, insurance companies, retail merchants, manufacturers, business-to-business suppliers and even services providers. Predictive analysis can help to find high-risk fraud candidates in business or the public sector.

f. **Portfolio, product or economy-level prediction:** These types of issues can be addressed by predictive analytics using time series techniques. They can also be addressed via machine learning approaches which transform the original time series into a feature vector space, where the learning algorithm finds patterns that have predictive power.

g. **Threat management:** When employing threat management techniques, the results are always to predict and benefit from a upcoming scenario. Predictive analysis helps organizations or business enterprises to find upcoming threat, Natural Disaster and its effect. Threat management helps them to take correct decision on correct time.

h. **Underwriting:** Many businesses have to account for threat exposure due to their various services and determine the cost needed to cover the threat. For example, auto insurance providers need to accurately determine the amount of premium to charge to cover each automobile and driver. For a health insurance provider, predictive analytics can study a few years of past medical claims data, as well as lab, pharmacy and other records where available, to predict how costly an enrollee is likely to be in the upcoming. Predictive analytics can help underwrite these quantities by predicting the opportunist of illness, default, bankruptcy, etc. Predictive analytics can streamline the process of user acquisition by predicting the upcoming threat behavior of a user using application level data.

IV. CONTENT TYPE ANALYTICS

Most of all information or data is available in textual form in databases. From these contexts, manual Analytics or effective extraction of important information are not possible. For that it is relevant to provide some automatic tools for analyzing large textual data. Content Type analytics or Content Type mining refers process of deriving important information from Content Type data. It will use to extract meaningful data from the Content. It use many ways like associations among entities, predictive rules, patterns, concepts, events etc. based on rules. Content Type analytics widely use in government, research, and business needs. Content simply tells you what people did but Content Type analytics tell you why. From unstructured or semi structured Content Type data all information will retrieve. From all type data it will extract important information. After extracting information it will be categorized. And from these categorized information we can take decision for business.

Steps for Content Type Analytics system:

a) **Content Data:** In initial stage data is unstructured.

b) **Content Type Data processing:** All information will transfer in Semantic Syntactic Content Type.

c) **Content Type Data transformation:** In it important Content Type will extract for future use.

d) **Feature selection:** In it data is counted and display in Statistics format.

e) **Data mining:** All data is classified and clustered.

Content Type Analytics applications areas:

a) **Security application:** It will we monitoring and analyzing internet blogs, news, Social Media sites etc. for national security purpose. It will use full detect unethical thing on internet.

b) **Marketing application:** By analyzing Content Type data we can identify which type of product customer most like.

c) **Analyzing open – ended survey responses:** In survey research one company ask to user some question like, pros and cons about some products or

asking for suggestion. For analyzing these Type of data, Content Type analytics is requiring.

d) **Automatic process on emails and messages:** By using big data analytics we can filter huge amount of emails based on some terms or words. It is also useful when you want to automatically divert messages or mails to appropriate department or section.

Various Type of Content in Social Media:

a) **Time Sensitivity:** An important feature of the social media services is their real-time nature. With the rapid growth of the content and communication styles, [12]Content Type is also changing. As the time sensitivity of the textual data the people's thoughts also changes from time to time.

b) **Short Length:** Successful processing of the short Content Type is essential for the Content Type analytics method. As the messages are short, it makes people more efficient with their participation in Social Media networking websites. Short messages are used in social media which consists of few phrases or sentences.

c) **Unstructured Phrases:** An important difference between the Content Type in social media and traditional media is the difference in the quality of content. Different people posts different things according to their knowledge, ideas, and thoughts. When composing a message also many new abbreviations and acronyms are used for e.g. How r u? —Gr8‖ are actually not words but they are popular in social media.

Applying Content Type Analytics to Social Media:

a) **Event Detection:** It aims to monitor a data source and detect the occurrence of an event that is to be captured within that source. These data sources include images, Video Types, Audio Types, Content Type documents.

b) **Collaborative Question Answering:** As Social Media networking websites has emerged, the collaborative question answering services have also emerged. It includes several expert people to answer the questions posted by the people. A large number of questions and answers are posted on the Social Media networking websites.

c) **Social Media Tagging:** Tagging of the data has also increased to a great extent. For example when any particular user is looking or searching for a recent event like —Election‖ then the system will return the results that are tagged as —‖Election‖. Textual data in social media provides lots of information and also the user-generated content provides diverse and unique information in forms of comments, posts and tags.

V. AUDIO TYPE ANALYTICS

Audio Type analytics is the process of compressing data and packaging the data in to single format called Audio Type. Audio Type Analytics[9] refers to the extraction of meaning and information from Audio Type signals for Analysis. There are two way to represent the Audio Type Analytics is 1) Sound Representation 2) Raw Sound Files.

Audio Type file format is a format for store digital Audio Type data on a system. There are three main Audio Type format: Uncompressed Audio Type format, Lossless compressed Audio Type format, Lossy compressed Audio Type format.

Application Area of Audio Type Analytics:

The Audio Type is the file format that used to transfer the data to one place to another. Audio Type analytics is used to check whether given Audio Type data is available in proper format or in similar format that sender sends. The Application of Audio Type Analytics is many:

a) **Surveillance application:** Surveillance application is based on approach for systematic choice of Audio Type classes for detection of crimes done in society. A surveillance application is based on Audio Type Analytics framework is the only way to detect suspicious kind of activity. The application is also used to send some important information to surveillance at some crisis situation urgently.

b) Detection of Threats: The Audio Type mechanism is used to identify the thread that take place between sender and receiver.

c) Tele-monitoring System: New technology have camera with the facilities to record the Audio Type also. Audio Type Analytics may provide effective detection of screams, breaking glass, gun sound, explosions, calling for help sound etc. Combination of Audio Type Analytics and Video Type Analytics in single monitoring system result as a good threat detection efficiency.

d) Mobile Networking System: The Mobile networking system is used to talk or transfer information to one place to another place. Sometimes due to some network problem the Audio Type sound is not work properly at that time Audio Type Analytics is used to find the information that not send properly due to some problems.

VI. VIDEO TYPE ANALYTICS

Video Type is a major issue when considering big data. Video Type and images contribute to 80 % of unstructured data. Now a day, CCTV[12] cameras are the one form of digital information and surveillance. All these information is stored and processed for further use, but Video Type contains lots of information and is generally large in size. For example YouTube has innumerable Video Type being uploaded every minute containing massive information. Not all Video Type are important and viewed largely. This creates a situation where Video Type create a junk and hard-core contribution to big data problems. Apart from Video Type, surveillance cameras generate a lot of information in seconds. Even a small Digital[11] camera capturing an image stores millions of pixel information in mille seconds.

Video Type Data Analytics dimensions –

Volume: Size of Video Type being more, takes the network as well as the server, time for processing. Low bandwidth connections create traffic on network as these Video Type deliver slowly. When stored on mass storage on secondary storage requires huge amount of space and takes more time retrieving as well as processing.

Variety: Video Type consisting of various format and variety such as HD Video Type, Blu-ray copies etc.

Velocity: It is speed of data. Now a day, Digital cameras process and capture Video Type at a very high quality and high speed. Video Type editing makes it to grow in size as it contains other extra information about the Video Type. Video Type grow in size faster as they are simply nothing but collection of images.

Application of Video Type analytics:

a) Useful in accident cases: With the use of CCTV cameras we can identify what happened at the time of accident it's also used for security reason and parking vehicles etc.

b) Useful in schools, traffic police, business, security etc.

c) Video Type Analytics for investigation (Video Type Search): Video Type analytics algorithms is implemented to analyze Video Type, a task that is challenging and its very time consuming for human operator especially when there is large amount of data are available using Video Type analytics we can search particular Video Type when we required.

d) Video Type analytics for Business Intelligence: It uses to extracts statistical and operational data. Rather than having operator that review all the Video Type and tally all the people or cars moving in certain area, or checking which traffic routes are most commonly taken, Video Type analytics can do it automatically.

e) Target and Scene Analytics: Video Type Analytics for business Intelligence involves target and scene Analytics. Target Analytics provides details information about the target movement, patterns, appearance and other characteristics which can be used for identification of target.

f) Direction Analytics: Direction Analytics is the ability to distinguish behavior by assigning specific values (low to high) to areas within a camera's field of view.

g) Remove the human equation through the automation: It removes the tedium involved in giving one or more set of eyes on a monitor for an extended period of time. The automation of Video Type analytics allows the insertion of human judgment at the most critical time in the surveillance process.

VII. CONCLUSION

Now, computer industry accepts Big Data as a latest challenge for different types of machine automated systems. There are many issues in storage management and accessing of data known as Big Data. The major issue is how we can use this data for increasing business and improvement in living standard of people. In this paper, we are discussing the issues, challenges, application as well as proposing some actionable insight for Big Data. It will influence researchers for finding knowledge from the huge amount of data available in various forms in various regions.

VIII. REFERENCES

[1]. Web content accessible on the connection: -http://www.sas.com/en_us/experiences/huge information/what-is-bigdata.html‖ on the dated: 16-08-2015

[2]. Web content accessible on the connection: -http://www-01.ibm.com/programming/information/bigdata/what-is-bigdata.html‖ on the dated: 16-08-2015

[3]. Web content accessible on the connection: -http://www.dqindia.com/8-creative cases of-enormous datausage-in-india/‖ on the dated: 16-08-2015

[4]. Web content accessible on the connection: -http://searchbusinessanalytics.techtarget.com/definition/enormous dataanalytics‖ on the dated: 16-08-2015

[5]. Web content accessible on the connection: -https://www.statsoft.com/course reading/content mining‖ on the dated: 16-08-2015

[6]. Web content accessible on the connection: -https://www.predictiveanalyticstoday.com/content analytics‖ on the dated: 16-08-2015

[7]. Web content accessible on the connection: -https://gigaom.com/2014/01/24/why-Video Type-is-the-following enormous thingin-huge information/‖ on the dated: 16-08-2015

[8]. Web content accessible on the connection: -http://searchbusinessanalytics.techtarget.com/definition/Social Mediamedia-analytics‖ on the dated: 16-08-2015

[9]. Web content accessible on the connection: http://en.wikipedia.org/wiki/E-commerce_in_India#cite_noteOnline_shopping_touched_new_heights_in_India_in_2012-1‖ on the dated: 16-08-2015.

[10]. Amir Gondomar, Mustafa Hider, -Beyond the buildup: Big information ideas, techniques, and analytics‖, International Journal of Information Management 35 (2015) 137– 144, diary landing page: www.elsevier.com/find/ijinfomgt

[11]. Radhakrishnan, Ajay Divakaran and Paris Smaragdis, -AUDIO TYPE ANALYSIS FOR SURVEILLANCE APPLICATIONS‖, 2005 IEEE Workshop on Applications of Signal Processing to Audio Type and Acoustics, October 16-19, 2005, New Plats, NY.

[12]. C.L. Philip Chen, Chun-Yang Zhang, 2014, -Data-escalated applications, difficulties, procedures and innovations: A study on Big Data‖, Contents records accessible at Science Direct Information Sciences, 275 (2014) 314– 347

[13]. Stephen Kaiser, Frank Armor, J. Alberto Espinosa, William Money, 2013, -Big Data: Issues and Challenges Moving Forward‖, 2013 46th Hawaii International Conference on System Scion 1530-1605/12, 2012 IEEE,

[14]. Web content accessible on the connection: -http://analyticstraining.com/2014/the-energy of-Social Media-mediaanalytics/‖ on the dated: 17-10-2015

[15]. Demon Bengali, James Horsey, 2012, -Design Principles for Effective Knowledge Discovery from Big Data‖, 2012 Joint Working Conference on Software Architecture and sixth European Conference on Software Architecture, 978-0-7695-4827-2/12, IEEE