

Review of Recent Technologies in Big Data Analysis

Manvi Sharma
BTech. Student,
Department of CS,
MACERC, Jaipur, India
manvi.95sharma@gmail.com

Manish Mathuria
Assistant Professor,
Department of CS,
MACERC, Jaipur, India
manish.4598@gmail.com

Ashish Sharma
Associate Professor,
Department of CS,
MAIIT, Kota India
ashishsharma.pce@gmail.com

Abstract—The advancement in image processing techniques made better computer vision. This computer vision is used for authentication and authorization of different entities at different places. Digital authentication done by UID verification is now used at different level for customer identification. Face recognition is totally based on image processing technique. This processing completes in two steps: Verification and Identification. The pre-processing is done to optimize image quality while it removes noise. The post processing is done for feature extension and pattern matching. The availability of efficient hardware and advanced software's is possible to uniquely identify human being. So this research proves the availability of finger print technology is successful in India.

Keywords—Digital Image Processing, Face Recognition, Face Matching;

I. INTRODUCTION

Data has always been a product of whatever work humans do, be it in Healthcare, police reports, sports and anything and everything with human involvement produces data. Historically people use to create data manually, they used to appoint workers and write down information in files. With evolution of computers, this task became simpler but with this the workload For increased and so did the amount of data being produced, now it's like a progression with the human evolution and advancement, the data being generated is showing an exponential growth. If we look at the current scenarios, there is humungous development in the every field. Technology has made it possible for humans to disregard the statement that "sky's the limit", as there are footprints on moon and we are reaching out to mars and places beyond. All because of technology. Now thinking about it, we have Internet of Things, everything is connected to internet and producing data at an unimaginable pace and amount. All kinds of data available either structured or unstructured, in form of images, signals, videos, text and a lot more things. This whole lot of data can be summed in two small words: BIG DATA. [1]

Giving a proper definition, it can be stated "Big data usually 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 refers to technologies and initiatives that involve data that is too diverse, fast-changing or massive for conventional technologies, skills and infrastructure to address efficiently. Said differently, the volume, velocity or variety of data is too great.

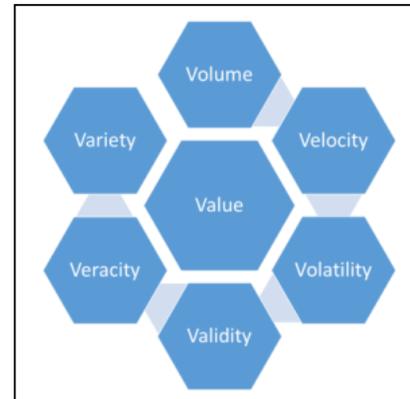


Fig. 1. : Characteristics of Big Data [1] [2]

- A. **Technical Challenges with Big-Data:** The techniques used in processing of Bid-data are capturing, storing, and presentation. The data available in depository may have different formats and status. Data digitalization is also require in some special condition. [2]

The Big Data value creation model the model is divided into four parts i.e. Big Data Assets, Big Data Capabilities, Big Data Analytics, and Big Data Value. The Big Data Assets are processed to next stage where data is identified by the capabilities in terms of organization, process, people, and system. The Big Data Analytics stage is very important because Big Data value is the next stage . [3]

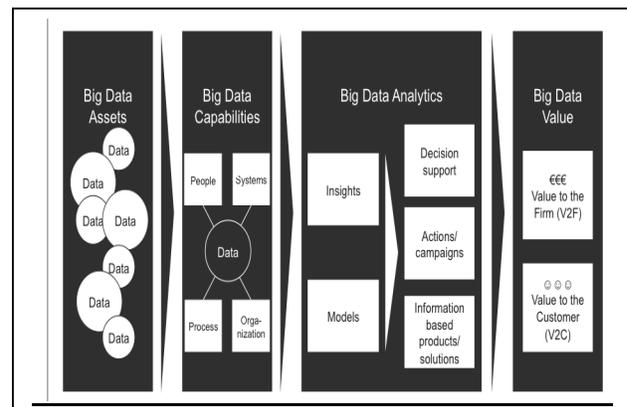


Fig. 2. Big Data Value Creation Model [3]

II. APPLICATION OF BIG DATA

With everything already being said it is somewhat predictable that Big Data finds its use in lots of applications. These applications are in almost every field, be it banking and security, health care, Government, financial, manufacturing, education, be it in the field of media and entertainment. Big data analytics finds its use everywhere. Many of the biggest examples of applications of Big Data are Google, Facebook, Amazon, Netflix and so many other things. Big Data finds its applications from satellites revolving around the earth to temperature recording devices in the weather forecast office. The favourite movies or shows on youtube and ads on your social media pages, everything is big data application. [4]

III. SOFTWARE TOOLS AVAILABLE

There are a large number of Big Data Analysis and Management tools. Each of them is proficient in providing real business value by bringing cost efficiency, better time management, and analyzing data to discover valuable business insights.

Table 1: Big Data Analysis Tools [5]

Technology	What is it? Why it matters?
Hadoop	Perhaps the most-cited technology in connection with Big Data, Apache Hadoop is an open source framework that allows for the distributed processing of the data sets across clusters of inexpensive computers. Can be used for management, data collection, storage, machine learning and security, to give few examples.
MapReduce	Usually used with Hadoop. MapReduce is a programming technique used for processing parallelizable problems across very large data sets
NoSQL databases	This is used when the underlying data relationships are not usefully understood through the traditional Relational Database model. These databases are primarily used for storing and retrieving very large data sets for real-time systems and statistical analysis.
Columnar databases	A type of database in which storage access is around columns rather than rows, which is more efficient for certain type of communication particularly those based on aggregation of columnar data sets.
In-memory databases	These databases fit large datasets into main RAM memory as opposed to disk file storage. This approach can speed query time by orders of magnitude- make or break a difference. They are useful when response time is critical, in times of emergency can be an example.

IV. RECENT TRANDS IN BIG DATA ANALYSIS

- Big Data becomes fast and approachable.
- Big Data is no longer just hadoop.
- Organisations leverage data lakes from the get-go to drive value.
- Architectures have matured to reject “one-size-fits-all” frameworks.
- Variety, not Volume or Velocity, derives Big Data investments.
- Spark and machine learning light up Big Data.
- Convergence of IoT, cloud, and big data creates new opportunities for self-service analytics.
- Self-service data prep has become mainstream as end users begin to shape big data.
- Big data grows up: Hadoop adds to enterprise standards.
- Rise of metadata catalogs helps people find analysis-worthy big data.[6] [7]

V. PRIVACY PRESERVING ISSUES IN BIG-DATA

- Privacy Issues in Big Mobile Data: Mobile data is exchange with service provider very freely, so it is very easy to identify user through his data. There are some Android Applications that can access or share user’s personal data. Application at Google Play Store found most of the application asked location-related data.
- Privacy Issues in Health Care Data: Reserches in Health and Medical Science require real time access of patient records which acctually helps doctors to take improtant decisions. Electronic Health Records contain personal information of patient which is important to encrypted before data analysis phase.
- Privacy Issues in Social Media Data: People share their life event and personal information on the social media, and sometimes people also share data about others. These activity may lead to privacy violation. [8]

VI. CONCLUSION

The review of recent technology in big data analysis are done successfully which gave meaningful result in the term of analysis of report. It is a basic paper which generally lookout for the latest application and software tools available of big data analysis. The recent trends in big data analysis also help to understand to industry requirement and future possibilities. The overall conclusion of this research work concluded that big data analysis is very important for business, medical, government and weather forecasting which actually helps to make future predictions. The analysis of big data can help organization from the big loss and help to make better management.

References

- [1] Kari Venkatram, Geetha Mary A., “Review on Big Data & Analytics- Concepts, Philosophy, Process and Applications”, Published in Bulgarian Academy of Science, Cybernetics and Information Technologies Vol. 17, No. 2.
- [2] Chunning Wang, Lichao Chen, “Exposing Library Data with Big Data Technology”, ICIS June (IEEE), 2016
- [3] Verhoef, Kooge, Walk, “Creating Value with Big Data Analytics”, Book published by Routledge (London and New York) (ACM), in 2016.
- [4] J. Archenaa and E. A. Mary Anita, “A Survey of Big Data Analytics in Healthcare and Government”, 2nd International Symposium on Big Data and Cloud Computing (ISBCC-15) 2004.
- [5] Steve Van Till, “How will Big Data Change Security?”, Online Article published at ‘securitymagazine.com’ April, 2013.
- [6] Time Winchcomb, Sam Massey, Paul Beastall, “Review of latest developments in the Internet of Things”, Cambridge Consultants, Doc Ref: P2952-R-001 v4.8, Date: 07 March 2017.
- [7] Alexei Balaganski, Dr. Sebastian Derwisch, “Big Data and Information Security”, Published in KuppingerCole Report (survey), November 2016.
- [8] Brijesh B. Mehta, Udai Pratap Rao, “Privacy Preserving Unstructured Big Data Analysis: Issues and Challenges”, ICISP 2015, Science Direct, Elsevier, 2016.