

Potential impact of Big Data in Real World Sectors

Srijay Soni^{1*}, Riya Shah², Sagar Vyas³

¹Department of Information and Technology, Gandhinagar Institute of Technology,, Gandhinagar, Gujarat, India;
srijaysoni2198@gmail.com

²Department of Information and Technology,, Gandhinagar Institute of Technology,, Gandhinagar, Gujarat, India;
riya.shah5555@gmail.com

³Department of Information and Technology, Gandhinagar Institute of Technology,, Gandhinagar, Gujarat, IndiaIndia;
vyassagar62@gmail.com

*Correspondence: *srijaysoni2198@gmail.com*

Abstract: In this information and digital era vast amount of data have become available on hand to decision makers. Big data represents not only the huge amount of data but also the data which are high in variety and velocity. Due to the galloping growth of the data, solutions are needed to handle, analyse, and extract the information from the datasets. Big data has a potential impact on the digital world. After examining the impact of big data on various fields and sectors, this paper presents a brief review of what is big data and the use of big data in our real world. In this paper, we have addressed the various type of fields where big data is used such as Digital Space, Healthcare, Finance, Education and Government.

I. Introduction

In this digital era with the growth of technologies and services, a large amount of data is produced that can be structured, semi-structured and unstructured from different sources. In simple terms, data is facts and statistics collected together for reference or analysis. Data consumes various details of organization, transaction, and documents. If there is no data storage detail about a person, transaction or organization will be lost after the use of data. Big Data is a massive volume of both structured and unstructured data. Big data is so huge and complex that it is difficult to process and work on them using traditional database and software techniques. For any huge organization or company, the data storing, analyzing, predicting will be in huge amount, so it is necessary to work with big data. Due to the availability of Big data it made the traditional marketing tools to work more powerfully and effectively. Many wide varieties of scalable database tools have evolved like Hadoop, NoSQL, MongoDB, etc. (Ishwarappa and Anuradha, 2015). While Big Data is more of a marketing term than anything, the implication is usually that you have so much data that you can't analyze all the data at once because the amount of memory (RAM) it would take to hold the data in memory to process and analyze it is greater than the amount of available memory. This means that analyses usually have to be done on random segments of data, which allows models to be built to compare against other parts of the data. There are four characteristics of Big data. They are volume, velocity, variety and veracity. Volume refers to the size or amount of data. Nowadays data volume is increasing from gigabytes to petabytes (Smitha and Kumar, 2013). 40 Zettabytes of data will be created by 2020

which is 300 times from 2005 [IBM]. Velocity is the speed at which data is created, stored, analyzed and visualized. Variety refers to many sources and types of data. Data may be in different forms such as Text, numerical, images, audio, video, social media data (Smitha and Kumar, 2013). On Twitter, 400 million tweets are sent per day, and there are 200 million active users on it (IBM report). Veracity means uncertainty or accuracy of the data.

Advantages of Big Data:

- Big Data analysis concludes innovative solutions. It helps in understanding and targeting customers and helps in optimizing business processes.
- Big data helps in improving science and research.
- Big data improves healthcare and public health with the availability of record of the patients.
- Big data helps in financial trading, digital marketing, sports, polling, security enforcement, education, etc.
- Every second addition is made through Big data.
- One platform can carry vast information.

Disadvantages of Big data:

- Storage can cost a lot of money to store Big data.
- Many data are unstructured in Big data.
- It can be used for exploiting the customer data/records.
- It is not useful for a short period. It needs to be analyzed for long durations to credit its benefits.
- Due to quick and speedy updates in big data, it might mismatch the real figures.

This study is about the capability and impact of Big data in various sectors where it can be more helpful and innovative than the present days.

II. Impact of Big Data in Digital Space:

Big data is like a secret ingredient, raw material. However, it is not that data is not important. However, rather, Big data is used for making intelligent marketing decisions. The decisions made and action taken afterwards makes much difference.

By combining the integrated strategy with big data, the organizations make a large impact on this key areas:

- Customer Engagement: Big data just not tells where your customers are, but who they are, what they want, and how and when they should be contacted.
- Customer loyalty and holding: Big data helps you to discover that what influences customer loyalty and what keeps the customer coming back.
- Marketing Achievement: With Big data, one can find the optimal marketing spends across the world and it can also optimize the marketing programs through the testing, analysis and measurement.



Figure 1: The Frameworks of Digital Marketing in Big Data

Grishikashvili et al., 2014 has explained the above framework (Figure 1).

- Commune:
What: Customize interaction and value delivery
How: Personalized/Customized communications and Product services, Real-time interaction, Linkage to core Business.
- Attract:
What: Attract consumer to the applications.

How: Audience creation, mnemonic branding, “Piggyback” advertising.

- Involve:
What: Generate Interest and Participation.

How: Intuitive interface/navigation, interactive content, User-generated content.

- Reserve:
What: Make sure customers come back.

How: Dynamic content, Transaction capabilities, Digital communities.

- Master:
What: Learn about customer preferences.

How: Information capture, continuous preference learning

In Future the big data will impact on this factor:

- Amount of Data:

The amount of data marketers in the past few years has tremendous access to data—especially due to smartphones.

IBM has stated that 90% of the world’s data was created in the past two years, and this will continue as we have more and more connected devices used in our everyday lives which produce a vast amount of data.

- Social Changeover:

According to the current scenario, approximately 5.5 million “Internet of Things” or “IOT” are connected every day in appliance-like Television, Fridges, Air conditioners, etc. The data stored by the customers or person interactions are stored and analysed by marketers to judge the customers. Now that connected homes becoming more popular, and smart speaker devices like Amazon Echo and Google Home with their own assistants are user friendly. Consumers will use the devices for making purchases or just for surfing the web. Big Data will have a big role to play in judging how to find the customers and how to appeal or interact with them.

- Customers are always connected:

Customers can choose when to receive or read the content whenever or wherever they are. Big Data can show us which device to root on and how the users interact with online stores and brands depending on the device. The marketers need to figure out how to connect with the customers.

- Location Marketing:

To market to existing customers based on their location is an excellent use of Big Data. There is a balance for having privacy, but the customers who choose to share their location

history provide a better deal of insight into their behaviour. Big data will also help the marketers analyse the timing of location-based offers or other techniques. The best example of that is Uber that has business off the back of location data.

III. Healthcare

The healthcare industry can generate a huge collection of data. By using these data, we can find a proper diagnosis for diseases which needlessly consume healthcare money. Moreover, also this can avoid curable death. The way of life has been changed by the Big Data. The way of life has been changed by big data, analyze and leverage data in any industry. One of the most promising areas where it can be applied to make a change in healthcare. Healthcare analytics, which has the potential to reduce costs of treatment, predict outbreaks of epidemics avoid preventable diseases and improve the quality of life in general. World population increases along with the average human lifespan, which poses new challenges to today's healthcare industry and treatment delivery methods. Looking for the best strategies to use these numbers, health professionals are capable of collecting a massive amount of data.

There are many areas which can be taken as an instance for big data analytics, but few specific areas which reflect a broader view of big data analytics in the health sector (Belle et al., 2015).

- Image Processing

Therapeutic images are a vital source of data frequently used for diagnosis, therapy assessment and curing (Ritter et al., 2011). Medical image data ranges from one megabyte to hundreds of megabyte per single case study when a massive data is concerned large storage capacity is required if data has to be stored for a long period. CT (Computed Tomography), MRI (Magnetic Resonance Imaging), X-ray, molecular imaging, ultrasound are some of the examples of imaging techniques that are well established and much used nowadays.

- Genomics

The budget to sequence the human genome (encompassing 30,000 to 35,000 genes) is rapidly decreasing with the development of high-throughput sequencing technology (Lander et al., 2001; Darmanac et al., 2010). With implications for current public health policies and delivery of care, analyzing genome-scale data for developing actionable recommendations promptly is a significant challenge to the field of computational biology.

For disease prediction, some computer-aided methods have been developed. For example, for predicting inpatient mortality, APACHE III is a prognostic scoring system. There are also other specific models for specific disease and specific conditions, such as heart conditions (Jenkins et al., 2002), hepatitis (Piscaglia et al., 2006), Alzheimer's disease (Liu et al., 2007), and cancer (Mould, R.F., 2003). In one general predictive system which consists of massive data containing symptoms of various diseases are used to detect a person's disease from which he is suffering. Using Big Data, specifically collaborative filtering, based on data from similar patients we generate predictions focused on other diseases. These predictions can lead to improve management and prevention strategies, and potentially empower the patient to have a dialogue that leads to improved wellbeing. They may also guide some rare diseases and complications that could confound a physician but are elucidated by the data-driven integration of experiences of many physicians and patients.

Figure 2 illustrates the theoretic platform for CARE. When someone arrives in an office with his or her medical history, this medical history is compared with all the other patients' medical histories that one may have access to based on defined similarity constraints. The similarity could be defined by symptoms, family histories, lab results, urban/rural residencies, occupation, demographics, etc. Based on the similarity computation, a pool of patients most similar to the patient under consideration is selected. Once the similar universe of patients is selected, we apply collaborative filtering using inverse frequency and vector similarity (Breese et al., 1998). The functioning of collaborative filtering can be specified mathematically.

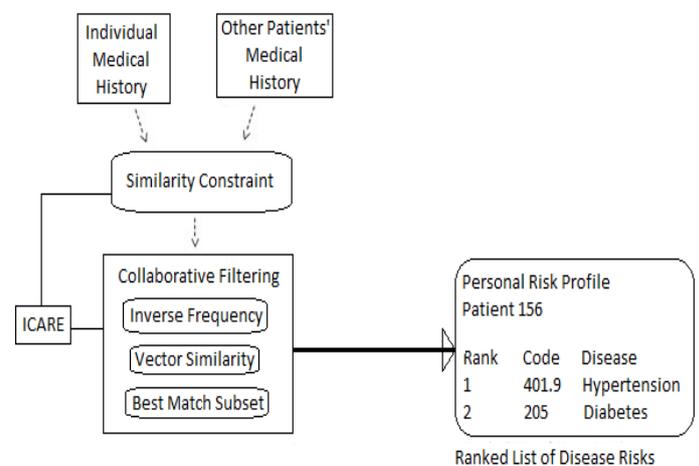


Figure 2. CARE process diagram

IV. How Finance sector involves in Big data

The ability to analyze, access and manage vast volumes of data while rapidly evolving the Information Architecture has long been critical to financial services companies as they improve business efficiency and their performance (Stackowiak et al., 2015). Recently, bank profitability has been on the move from a lower position to a higher one, particularly in locales of the reality where financial conditions are great. Budgetary administrations associations will proceed on the fundamental and focal spotlight on income development and higher edges through operational effectiveness, better hazard administration, and enhanced client closeness. Banks will likewise grow new income streams by entering new administration zone and Showcasing.

Financial business forecasts from various sources such as:

- » News
- » Industry information

Modern huge information alludes to an immense measure of differentiated time arrangement produced at a rapid by mechanical hardware, and it is known as the Internet of Things(IOT). This huge sum information alludes to Huge information.

- » Exchanging information

At the point when the Hashtag expression and word has been utilized different occasions it at that point turns into a Slanting Theme.

- » Administrative information

Managed and secret information is information controlled by Government laws in the region of Medicinal services, fund, and Training.

- » Investigator reports (interior and contending banks)
- » Cautions about occasions (blogs, News, Twitter and other informing channels)

The financial distress modelling stream use data mining techniques to detect of interest to auditors to assist with their going concern evaluations and forecast the financial distress (or financial failure) of companies (Gepp et al., 2018). This financial distress might not be limited to a company or an organization, but may also extend to corporate stakeholders.

Khandani et al., (2010) use machine learning techniques to construct nonlinear, nonparametric models of consumer credit risk at the level of the individual and the customer, rather than the corporation. They combine customer transactions and credit bureau data and can use machine

learning to improve classification rates on credit card default and delinquencies significantly. This model throughout the recent financial calamity suggests that aggregated consumer-credit risk analytics may have major applications in forecasting systemic risk.

Singh et al., (2015) were inspired by animal ecology studies to analyse the transactions of thousands of people; they found that individual financial outcomes are associated with Traditional financial decision systems (e.g. credit), and this economy had to rely on explicit individual traits like age, gender, job type, and marital status, while being oblivious to spatiotemporal mobility or the habits of the individual involved with spatiotemporal traits (e.g., exploration and exploitation) and that these traits are over 30% better at predicting future financial difficulties than comparable demographic models.

V. Education:

In the education sector, Big Data offers unaccustomed opportunities for educators to reach out and instruct students in new ways. It will give them a deeper understanding of educator's education experience, and help them to understand the education system. Here are some of the important ways in which big data can impact the education offers.

The Value of Big Data in Education

Big Data can change research, as well as education. The Cloud Computing is the cooperation Technologies on which the Big Data depends. These advancements can enhance instructive administrations, giving understudies or students alike access to ease content, online educators, and networks of individual students. Big Data can bolster the great instructive framework helping instructors to examine what understudies know and which systems are best for every understudy. Along these lines, educators can likewise ready to learn new strategies and techniques for their training work (West, 2012). Technologies can give a quick criticism of understudies and educators about their scholarly execution, for example, Data mining and Data analytics. These techniques can give a profound examination of some training examples and concentrate important information from them. In this way, aggregate and huge scale information can predict whom student needs more assistance from the education system, maintaining a strategic distance from the peril of disappointment or drop out. This has accordingly to discover educational methodologies that appear to be most dynamic with specific understudies and exceptional needs. Then again, as Siemens and Gasevic say "Big Data can easily find apply to online education." As we can see, online education has a very big development in recent years and has a very increasing impact on the education sector.

Moreover, computerized learning is a gathering of data and analytics which can add to educating and learning. Along these lines, numerous understudies take an interest in on the web or portable realizing, where are made new data. This new data, additionally with the assistance of interpersonal organizations, are helping the understudies with the distinctive foundation to correspond among them and help them to comprehend centre course ideas.

Advantages of Big Data and Open Data in Education

- Improved instruction:

With the use of analytics, the courses can be adjusted from teachers to enhance understudies' execution and learning capacities making the exercises more close to home.

- Matching students to programs:

By using Open Data, students and parents can find the best school or education program.

- Matching students to employment:

Different Companies and candidate employees can make more effective tools with better performance to use open data to enable their skills with the required skills. What's more, understudies can likewise discover and find programming for occupations which can coordinate with their capability, more proficient than previously.

- Transparent education financing:

TEF opens the students to have additionally choosing power in advanced education and furthermore influence them to take an interest in training exercises, which already they will not be able to do as such.

- Efficient system administration:

Efficient system administration helps the School education systems to develop a skilful school supply chain which will provide a hand to administrators to allow more effective educational resources. In the other route around this will build up an anchor, superior, and bear the cost of an adaptable and keen arrangement for future education interests.

VI. The consequence of big data in Government:

The government additionally produces Terabytes and petabytes of information consistently. It requires an innovation that plays out a constant examination on the huge measure of the informational collection. This will assist the government in providing esteem added services to the citizens. Big data analytics helps in discovering valuable decisions by understanding the data patterns and the relationship between them with the help of machine learning algorithms(Patel and Shah,2017).

- The requirement for Big Data in Government:

Big data analytics helps the government in various

things. Big data use building smart cities by providing faster and reliable services to its citizens.

1. Addressing Basic Requirement Quickly:

2. Providing quality education:

3. To provide pension to senior citizens without any delay.

4. To ensure that benefits provided by government reach all the people.

5. To control traffic in peak times based on the live streaming data about vehicles.

6. To monitor the need for mobile ambulance facilities

7. To reduce the unemployment rate

Here is the best example of implementing big data in the government sector by India.

Indian population is 1.3 billion, so it is difficult to find a person data. Indian government follows many person identification documents like PAN card, voter card, driving license but there is no any unique identification then The Unique Identification Authority of India (UIDAI) is a statutory authority established under the provisions of the Aadhaar Act 2016 on 12 July 2016 by the Government of India, under the Ministry of Electronics and Information Technology.(Rajuet al., 2017). Under the Aadhaar Act 2016, UIDAI is responsible for Aadhaar authentication and enrolment including operation and management of all stages of Aadhaar life cycle, developing the procedure, policy and system for issuing Aadhaar to perform authentication and numbers to individuals and the security of identity information and authentication records of individuals. Aadhaar card contains the statistic highlights, for example, the name of the resident, Father/Mother's name, Date of Birth, Sex, the address of the citizen, phone number and biometric highlights, for example, photo, fingerprints and iris (eye) subtle elements. Aadhaar is not meant to replace existing identification documents like passport, driving license PAN Cards etc. However, it can be used as a single identification document because every person has unique biometric data. All the biometric information and statistic are put away into one incorporated database, and this task has been accounted for as a world's biggest database administration and Biometric ID framework individually by The Seasons of India and Forbes. Aadhaar is a 12-digit remarkable character number issued to every single Indian occupant in light of their biometric and statistic information. It was initially a document for identification purpose. However, it has now been made mandatory for some of the government-run programmers. This project also provides online support to change data of Aadhaar card using SSUP

(self-service update portal) from UIDAI official website(uidai.gov.in). (Raju et al., 2017).For an occurrence to change the name, one needs to present the Newspaper Notice of India specifying that 'required individual's name has been changed from old name to new name.To update DOB (Date of Birth), the required reports are Birth Testament issued by the Area Metropolitan Organization, and for the general population who were conceived 1989 and didn't have a birth endorsement, they can give an affirmation to change their DOB. Further, if one does not have the required document to change the DOB, then Passport or SSLC (Secondary School Leaving Certificate) can also be taken into consideration.For changes in address, landline charge, Mastercard bill and power charge less than three months old, bank passbook,a rental agreement Voter ID, or a Passport and the examined duplicate of verification of personality is adequate. Changes can likewise be made to the Portable number and Sexual orientation too, and evidence of character is required for these reasons. For all the statistic changes the confirmation is being checked through an OTP (One Time Secret key) sent to the versatile enlisted number. However, the biometric information cannot be changed. Nowadays the government of India is linking the Aadhaar card with many government functionalities such as Railway Reservation System,ATM Security, Cloud-BasedE-Voting, Aadhaar e-KYC Services,E-health Care,Municipal Corporations and Aadhaar Pay.

VII. Conclusion:

In this review paper, an overview is provided on applications of big data in various sectors.Big data is used for making intelligent marketing decisions so it can be used in digital space/marketing and it can also swing to shape a modern and dynamic education system which can be helpful for the education sector.Big data digs on insights related to the information given, thereby giving a progressive future to the healthcare and government sector.Big data also offers the ability to provide a global vision of different factors and areas related to the financial risk that can have a huge impact on business.

This paper discussed the potential positive impact of big data in fields like digital space,healthcare,education,finance and government.This paper aims to identify theoretical dilemmas and the importance of the digital revolution.If properly applied,Big data has the potential to provide a basis for advancements in technical and scientific levels.

Declaration

Availability of data and material

All relevant data and material are presented in the main paper.

Funding

Not Applicable

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

Authors acknowledge the support provided by Department of Information Technology, Gandhinagar Institute of Technology,Gandhinagar, Gujarat, India. Authors are also thankful to the college for giving permission to publish this research.

References

- [1]. Belle, A., Thiagarajan, R., Soroushmehr, S.M.R., Navidi, F., Beard, D.A., Najarian, K., 2015. Big Data Analytics in Healthcare. *BioMed Research International*. 1-16.
- [2]. Breese, J.S., Heckerman, D., Kadie, C. 1998. Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*. 1998: 43–52. Morgan Kaufmann Publishers Inc.
- [3]. Drmanac, R, Sparks, A.B., Callow, M.J., Halpern, A.L., Burns, N.L., Kermani, B.G., Carnevali, P., Nazarenko, I., Nilsen, G.B., Yeung, G., Dahl, F., Fernandez, A., Staker, B., Pant, K.P., Baccash, J., Borcherding, A.P., Brownley, A., Cedeno, R., Chen, L., Chernikoff, D., Cheung, A., Chirita, R., Curson, B., Ebert, J.C., Hacker, C.R., Hartlage, R., Hauser, B., Huang, S., Jiang, Y., Karpinchyk, V., Koenig, M., Kong, C., Landers, T., Le, C., Liu, J., McBride, C.E., Morenzoni, M., Morey, R.E., Mutch, K., Perazich, H., Perry, K., Peters, B.A., Peterson, J., Pethiyagoda, C.L., Pothuraju, K., Richter, C., Rosenbaum, A.M., Roy, S., Shafto, J., Sharanovich, U., Shannon, K.W., Sheppy, C.G., Sun, M., Thakuria, J.V., Tran, A., Vu, D., Zaranek, A.W., Wu, X., Drmanac, S., Oliphant, A.R., Banyai, W.C., Martin, B., Ballinger, D.G., Church, G.M., Reid, C.A. 2010. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*, 327(5961) 78–81.
- [4]. Gepp, A., Linnenluecke, M.K., Smith, T. 2018. Big Data Techniques in Auditing Research and Practice: Current Trends and Future Opportunities. *Journal of Accounting Literature*, 1-44.
- [5]. Grishikashvilli, K., Dibb, S., Meadows, M., 2014. The investigation into Big Data Impact on Digital Marketing. *International Conference on Communication, Media, Technology and Design*. 146-150.
- [6]. Ishwarappa., Anuradha, J., 2015. A Brief Introduction on Big Data 5Vs Characteristics and Hadoop Technology. *Procedia Computer Science*. 48, 319-324.
- [7]. Jenkins, K.J., Gauvreau, K., Newburger, K., Spray, T.L., Moller, J.H., Lezzoni, L.I. 2002. Consensus-based method for risk adjustment for surgery for congenital heart disease. *J Thorac Cardiovasc Surg*. 123(1):110–118.
- [8]. Khandani, A.E., Adlar J. K, Andrew W.L. 2010 . Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance* 34, 2767-2787.
- [9]. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dear, K., Doyle, M., Fitzhugh, W. 2001. Initial sequencing and analysis of the human genome. *Nature*, 409, 860-921.

-
- [10].Liu, Y.H., Chen, S.M., Lin, C.Y., Huang, C.I., Sun, Y.N., 2007. Motion tracking on elbow tissue from ultrasonic image sequence for patients with lateral epicondylitis. In 29th Annual International Conference of the IEEE. 95–98.
- [11].Mould, R.F. 2003. Prediction of long-term survival rates of cancer patients. *Lancet*. 361(9353):262.
- [12].Patel, R., Shah, N. 2017. Big Data Analytics Applications in Government Sector of India. *International Journal of Scientific Research in Engineering*. 1(1), 50-54.
- [13].Piscaglia, F., Cucchetti, A., Benlloch, S., Vivarelli, M., Berenguer, J., Bolondi, L., Pinna, A.D., Berenguer, M. 2006. Prediction of significant fibrosis in hepatitis C virus-infected liver transplant recipients by artificial neural network analysis of clinical factors. *Eur J Gastroenterol Hepatol*. 18(12):1255–1261.
- [14].Raju, R.S., Singh, S., Khatter, K. 2017. Aadhaar Card: Challenges and Impact on Digital Transformation. 1-21. (Unpublished Report).
- [15].Ritter, A., Cherry, C., Dolan, W.B., 2011. Data-Driven Response Generation in Social Media. *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*. 583-593.
- [16].Singh, V.K., BBeing.ozkaya, B., Pentland, A. 2015. Money Walks: Implicit Mobility Behavior and Financial Well-being. *PLoS ONE*, 10(8), 1-17.
- [17].Smitha, T., Kumar, V.S., 2013. Application of Big Data in Data Mining. *International Journal of Emerging Technology and Advanced Engineering*. 3(7), 390-393.
- [18].Stackowiak, R., Licht, A., Mantha, V., Nagode, L. 2015. Big Data and The Internet of Things. *Apress*. 1-220.
- [19].West, D.M., 2012. Data Mining, Data Analytics, and Web Dashboards, *Gov. Stud. Brook. US Reuters*.