

Big Data Application in Banking Sector

Mrs. Flavia Gonsalves (Prof. MET, Mumbai)¹ Miss Samiksha Jadhav (3rd Year MCA, MET, Mumbai)²

Abstract -Banking and finance institutes generate as well as have transaction of enormous amount of money every minutes over the world. So by its virtue of its very nature, this industry is data intensive. The data present in this sector is complex information and falls under the ambit of Big Data, which has already defined as “ Set of information with large, diverse and complex which is growing at increasing rate”. These data hold more potential for bank to have better understanding of customer base, product performance and industry trend in this competitive market sector.

With the technological advancement in past decade, nearly half of adult populations have begun using digital banking. This exponentially growing numbers of tablets, mobiles and other electronic devices have made it easy for customer to perform different activities, communicates with different organizations, research products and services, purchasing new item, feedback and performing banking task. These activities and data from it can be used to create a customer profile which bank can analysis to monitor trend and foresee customer behavior and offer customize services.

This paper discuss some of the existing application of Big Data in banking as well some future opportunities related to big data in banking sector and some underlying challenges which have to be tackled.

Key words- Big Data, Banking, Risk Management, Fraud Detection, Data Usability, Customer segmentation and targeting.

1. INTRODUCTION

With the rise of the internet, smart phones and other apps, digital data has escalated. The tremendous potential of using this knowledge, also known as Big Data, is recognized by private businesses and governments alike, to generate real value for consumers and increase productivity over time. Big data might make over businesses and economies, but data science is the real game changer.

Big Data Analytics is the study of broad and varied data sets to uncover hidden patterns, consumer demands and trends, unknown associations, customer desires and other valuable knowledge that helps to achieve major marketing objectives. The astounding rise comes from both the number of data generation devices and the number of sensors in each device; about 11 billion devices

Demonetization unfolded a new era that accelerated rapid growth in India's e-Banking and e-Commerce. Following demonetization, India's government began encouraging digitization for a cashless society making way for increased e-Banking and e-Commerce transactions. Actually, there are very few mobile payment applications like; MobiKwik, PayTM, Freecharge, Google Pay, BHIM UPI, State Bank Buddy, ICICI Pockets, HDFC Chillr, Citi MasterPass, Vodafone M-pesa, etc. The potential of the business can be clearly estimated by looking at PayTM's growth – an e-Commerce website started in 2010, offering only mobile recharging succeeded in attracting personal investment from industrialist Ratan Tata by March 2015. The same month, the company received a \$575

million investment from Alibaba Group of China. As of November 2016, PayTM became India's largest mobile payment service platform with over 150 million wallets & 75 million android based app downloads.

2. LITERATURE REVIEW

The financial sector is critical for the economy and is thus a subject of concern to researchers in a wide variety of areas, such as management science, marketing, finance and IT.

Berger (2003) found evidence of a connection between technical advances and banking productivity. Constant author jointly emphasizes that banks are using applied mathematics models that have supported their money expertise for various purposes, such as credit assessment and risk analysis.

In the study, "Penetration in the Fog: Learning and Training Analytics. EDUCAUSE Review, 46(5), 30–32 "(2011), Siemens, G., & Long. P describes big data as datasets whose size goes beyond the capability to collect, store, handle and evaluate traditional database software resources.

The conclusion in the research paper "Digitalization and Big Data Mining in Banking" (2018) was that the advent of cloud computing will dramatically boost the computational efficiency of most existing frameworks while the popularization of the Internet of Things enriches big data tools and can also have a positive impact on embedded analytics and the growth of dynamic,

3. WHAT IS BIG DATA?

Big Data is such a broad and complex array of data sets that the use of one-hand information management tools or conventional data processing software is impossible to handle. Data were traditionally processed in servers, as the volume of data was very low and there was less time to process those data as well. But now the data is a so quickly in the new technological environment and people are dependent on data a lot of times. It is becoming impossible to store the data in any server at the speed at which the data grows. Traditionally, data warehousing is used to store large quantities of data and perform complex analytics.. The data center does, however, work on a small set of data that contributes to inaccuracy. When you change the sample you might get the same outcome. To get meaningful results it took a high turnaround time. Data center does not store and process vast volumes of data in a very short time. These contribute to Big Data and Big Data Analytics developing. The challenge of Big data includes capturing, storing, searching, sharing, transferring, analyzing and visualization of these data.

Following are type of Big Data

- Structured Data
- Unstructured Data
- Semi Structured Data

Structured Data

Any data that can be stored, accessed, and processed in fixed format is considered 'structured' data. Over time , computer science talent has become more active in developing techniques for working with and deriving value from these data (where the format is well understood in advance). Nowadays, though, we are foreseeing problems when a scale of these data grows exponentially, typical sizes are in the rage of multiple zettabytes.

Unstructured Data

Any data with a form or structure unknown is classified as unstructured data. As well as the size being enormous, unstructured data poses multiple challenges in terms of its processing to derive value from it. A typical example of unstructured data is a heterogeneous source of data which contains a combination of simple text files , images, videos etc. Nowadays, companies have a wealth of data with them, but sadly they do not know how to extract meaning from it, because this data is in its raw form or unstructured format.

Semi Structured Data

Semi-structured data may contain data in both types. We can see semi-structured data as structured in form but it is not

specifically defined in relational DBMS with e.g. a table description. Semi-structured data is an example of data contained in an XML file

Characteristics of Big Data

The characteristics of Big Data are 4 Vs, following are the Characteristics,

- Volume
- Variety
- Velocity
- Veracity

Volume

The name Big Data itself has to do with a scale that's massive. Sample size plays a very important role in assessing value from the results. It also depends on the volume of data whether or not a particular data can actually be considered as a Big Data. 'Size' is also one attribute that needs to be taken into account when dealing with Big Data.

Variety

Variety refers to both structured and unstructured, heterogeneous sources and the nature of data. In the past few days, spreadsheets and databases were the only data sources that most applications considered. In analytics applications, data in the form of emails, photos , videos, monitoring devices, PDFs, audio, etc. are also being considered. The abundance of unstructured data raises some storage, processing and data analysis problems.

Velocity

The word 'velocity' refers to the rate at which data is produced. How quickly the data is generated and processed to meet the demands, determines the data's real potential.

Big Data Velocity discusses how quickly data flows from sources such as business processes, program logs, networks, social media platforms, sensors, mobile devices, etc. Information flow is enormous and continuous.

Veracity

This refers to the incoherence that the data can often show, thereby hindering the process of being able to handle and manage the data effectively

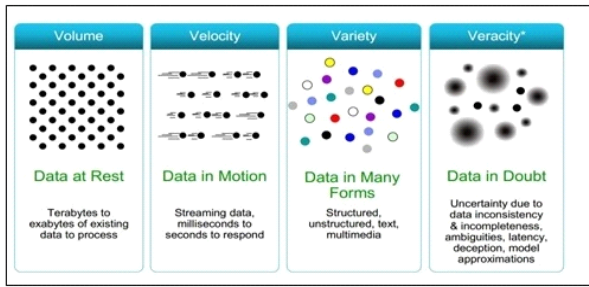


Figure 1 Characteristic of Big Data

Banking sector in India

After independence the banking industry in India has also expanded extraordinarily. A series of financial mismanagement and scandals in the 1960s and 70s had forced the government

to nationalize most banks. After 1991, however, following the footsteps of liberalization, globalization, and privatization, the private bank was reinvigorated the big way.

According to information provided by Reserve Bank of India (RBI), there are currently 26 public-sector banks, 20 private-sector banks and 43 foreign banks, approved to conduct banking operations in India. Another two entities, IDFC (Infrastructure Development and Financing Corporation) and previously a micro-finance firm, Bandhan, have got banking licenses. There are 61 national rural banks and over 90,000 cooperative banks, as well. India's banking sector has a net worth of 81 trillion Rs (\$1.31 trillion). According to research by KPMG and CII, India's banking industry is all set to become the fifth largest banking sector in the world by 2020 and the third largest by 2025.

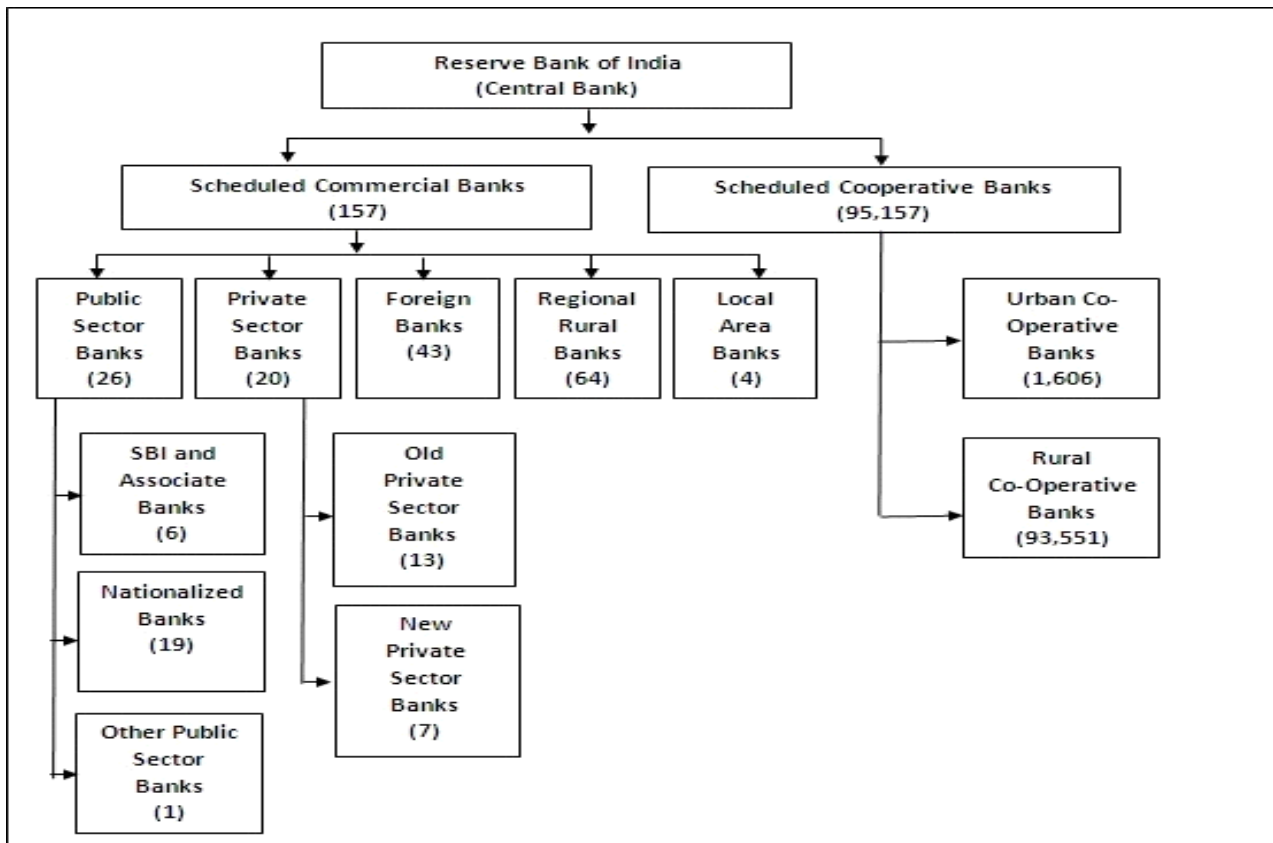


Figure 2 Banking structure in India

Innovative approaches to banking have also seen a increase in the new millennium. Banks have also leveraged the technological revolution which popularized laptops and mobile phones and took the Internet closer to the masses. Beginning with online banking, mobile banking and remote banking, technology has helped banks minimize dependency on physical branches and reach out to broader customer base using virtual banking resources.

4. ADVENT OF BIG DATA

Big data can be formulated as

$$\text{BIG DATA} = \text{Transaction} + \text{Interaction} + \text{Observations}$$

Gartner defines big data as high volume, speed and variety information assets which require cost-effective, creative forms of information processing for enhanced insight and decision-making. Every day, 2.5 quintillion bytes of data are generated.

The volume of data is exponentially increasing from terabytes to petabytes, exabytes and now zettabytes

According to IBM, 80 percent of today's data captured is unstructured from sensors used to gather climate information, posts to social media sites, digital images and videos, purchase transaction records, and GPS signals for cell phones, to name a few. Big data is all those unstructured data. Big Data is a series of large and complex data sets that are difficult to handle using on-hand information management systems or conventional applications for data processing.

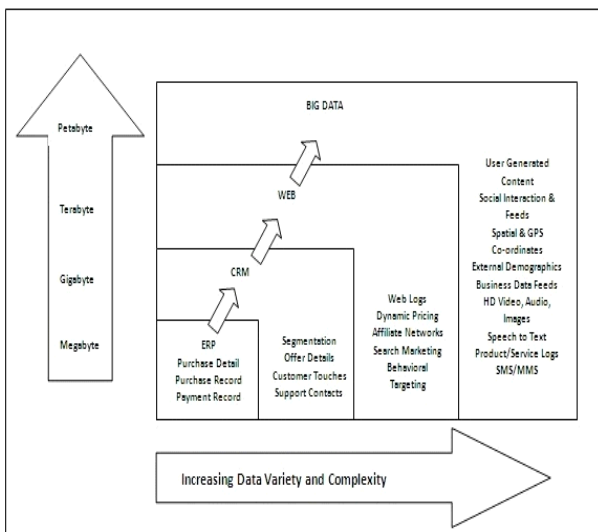


Figure 3 BIG DATA= Transaction + Interaction + Observation

Generation of Big Data

- Social Media and Networks
- Scientific Instruments
- Mobile Device and Laptops
- Sensor Technology and Networks

The payoff from using big-data analytics to analyze banking transactions is huge. The quantity of successful case studies continues to build, reinforcing wider research suggesting that when firms inject data and analytics deep into their banking operations, they can deliver efficient sales and higher profits. The new approach of data-driven sales, more in-depth knowledge on customer preferences, stronger forecasts and shorter decision-making periods leads companies to follow this model more quickly.

2019 This Is What Happens In An Internet Minute



Figure 4 Sources of Big Data in 2019

5. APPLICATION OF BIG DATA IN BANKING SECTOR

Application of Big Data in Banking sector are as follows,

- Risk Management
- Fraud Detection
- Customer Contentment
- Optimization of Business Processes
- Examine customer feedback
- Detect when a customer is about to leave
- Sentiment Analytics

Risk Management

Establishing a comprehensive system of risk management is of utmost importance to banking organizations or else they have to suffer from huge losses in revenue. Organizations have to keep innovating new things to keep alive in the competitive world and increase their profit as much as they can. Big Data Analysis helps businesses to identify danger in real time and, obviously, save the consumer from possible fraud.

Fraud Detection

The fast-growing digital world provides us with multiple benefits but also gives rise to various forms of fraud. The personal data are now more vulnerable than ever to cyber-attacks and it's the greatest obstacle a banking organization faces. Using Big Data Analysis and certain Machine Learning Algorithms, companies can now identify fraud before it can be

placed in. It is achieved by detecting common trends of user spending, anticipating irregular user behaviors, etc..

Customer Contentment

Considering the high amount of risk involved when dealing with the banking firms, one of the most challenging tasks for them is to ensure a customer satisfaction. Customer retention is a life-long journey for the banking companies, from ensuring the safety of their transactions to supplying them with the most appropriate and beneficial deals. Even more important than ever is the data they receive from their clients. Analyzing the data from their customers based on various factors allows them to target their customers even better. **Optimization of Business Processes**

Big data, when combined with machine learning, can help banks analyze internal processes and take steps to optimize them. This way they can reduce operating costs significantly

Examine customer feedback

Customer feeling can be gathered from various social media websites in the text type. Once these emotions can be recorded, they can be divided into positive and negative, and they can be used to deliver services to consumers by applying different filters. Detect when a client is about to leave. As we know, acquiring new customers costs more than retaining their old customers. If the bank takes care of consumer needs by knowing the issue; attention needs to be paid to finding a solution.

Sentiment Analytics

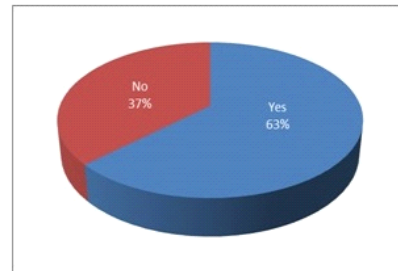
For marketing purposes, banks must constantly track what consumers are saying. Banks need to recognize who are the main consumers and strengthen those gaps to maximize efficiency and quality by obtaining feedback.

6. DATA ANALYSIS

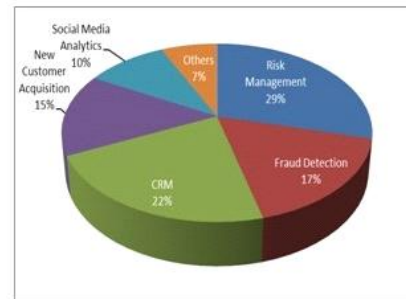
Result and Findings

The data was analyzed in one of the Multinational Bank during the Internship. The Bank workers working in the IT departments were asked questions.

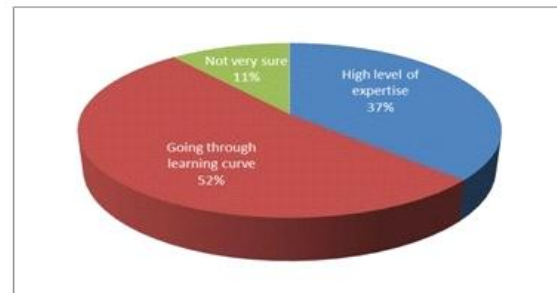
1. Do you use Big Data Analytics for your Banking operations?



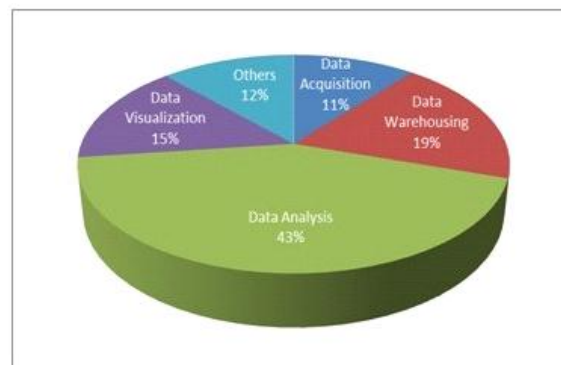
2. What are the main purposes for which you use big data analytics?



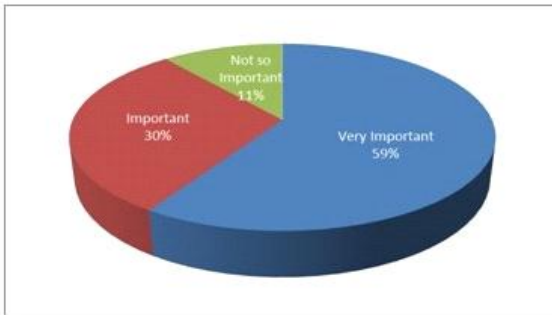
3. How well trained are your personnel to leverage the power of big data analytics?



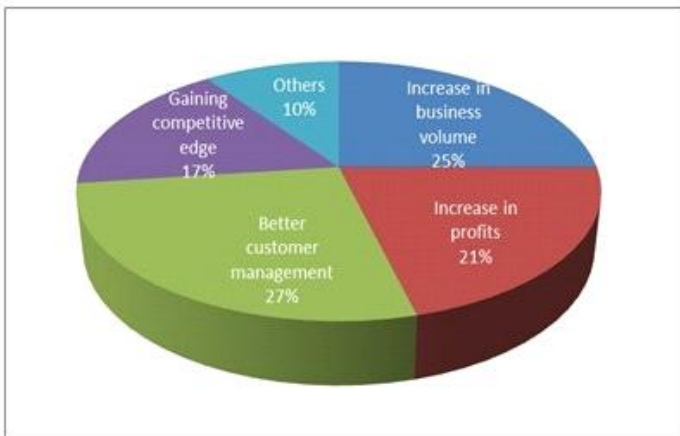
4. What are the main challenges that you face in utilization of big data analytics?



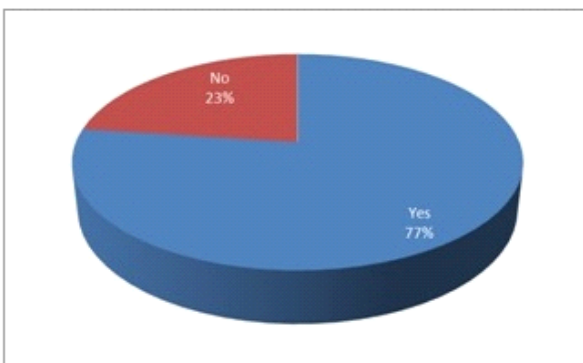
5. What is the importance of big data analytics for formulating strategies in your banking operations?



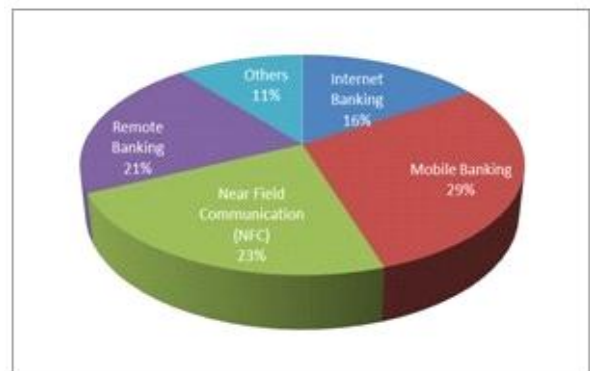
6. What benefits are you deriving by using big data analytics?



7. Has big data analytics been effective to provide meaningful insights into customer behaviour?



8. What do you think will be the future drivers of banking operations?



7. FUTURE OF BIG DATA IN BANKING SECTOR

As you can see, there's so much use of how big data in banking is used. But, barely scratched the surface of all those attempts. Big data's full capacity also needs to be harnessed in banking.

62 per cent of banks believe that big data is crucial to their success, according to Global Transaction Banking's whitepaper. Yet, only 29 per cent of them report having sufficient business value from their results.

If banks want to remain relevant and competitive, they need to rethink their operations and adopt data-driven approaches. Plus, big data will help you develop and expand your business in the banking sector.

8. Conclusions

The Age of Big Data is upon us. Organizations need to understand what Big Data is like and how to use it. The advantages and benefits are too great for companies to disregard.

The combination of various data sets, such as company data, public data and social data, would provide even more information.

There are more ways that banks and different monetary establishments have begun to capture client connected knowledge for sentiment analysis, ranging from social media websites to varied marketing research channels.

The impact of Big Data on society will be big, but it remains to be seen how society will impact Big Data.

9. REFERENCES

Impact of Big Data Analytics on Banking Sector: Learning for Indian Banks, Utkarsh Srivastava1, Santosh Gopalkrishnan2(2015)

Impact of Big Data Analytics on Banking Sector, Abhinav kathuria (2016)

An Empirical Study of Big Data Analytics In Indian Banking Sector, Mr.RAMPRAVESH GOND, Dr. RASHI GUPTA (2017)

Big Data in Banking: Opportunities and Challenges Post Demonetisation in India, Ayesha Anam Siddiqui¹, Dr. Riyazuddin Qureshi² (2018)

Digitalisation and Big Data Mining in Banking, Hossein Hassani ¹, Xu Huang ² and Emmanuel Silva ³(2018)

Cyfuture Blog - <https://cyfuture.com/blog/big-data-in-banking-advantages-challenges-and-applications/>

Eastern Perk Blog - <https://easternpeak.com/blog/big-data-in-the-banking-industry-the-main-challenges-and-use-cases/>