Topic:- Big Data Analytics (Bank Sector)

Kunal Vijay Ambre(C21002)

Aditya Arun Deshpande(C21027)

Deccan Education Society (Navinchandra Mehta Institute of Technology and Development)

ABSTRACT:

The arrival of new technologies, devices, and communications, the amount of data produced by mankind is growing rapidly every year. This gives rise to the new era of big data. The term big data comes with the new challenges to input, process and output the data. This Paper focuses on limitation of traditional approach, the techniques available to manage the data and the components that are useful in handling big data. One of the approaches used in processing big data is Hadoop framework, the paper presents the major components of the framework and working process also intorduces some of the challenges on Big Data.

For monitoring financial sectors or their networks Big data analytics helps to detect fraud transactions. Big Data analytics ensure exchange commissions, trading market, financial or bank sector that no fraudulent or any illegal activity or trading happens while Monitoring the markets or any stock market. Budgetary foundations assemble and get to investigative understanding from huge volumes of unstructured information so as to settle on sound money related choices. Enormous information examination enables them to get to the data they need when they need it, by killing covering, excess devices and frameworks. Bank are creating huge amount of the day by day. The speed of generating data is faster than our computer processor so handling of such large amount of data became difficult to handle for our system. In Banking Sector doing the things in old ways is riskier in nowadays. The bank sector must Adopt the new Technology if they want to succeed. Adopting the Big Data Analytics into new Banking sector or Existing one will be the key element of surviving and rapidly growing in digital environment of business' the last few Banking Sector has heavily worked and Invested to modernizing their mobile banking services and providing offers.

Keyword:- Big data Analytics, Big data application & Challenges, Digital bank & fraud detection, Tool used in big data.

INTRODUCTION

Big Data & Analytics

The Big Data name itself define that big data. We can say that the big amount of data which are structure or unstructured in nature[13]. This type of data is Gigabytes or terabytes or may be more than zettabytes. In Big Data the Structure data get processed easily. The structure data are having knew attribute ties like integer, character, decimal and. But unstructured data are not easy to processed[13. The structure is in tables, rows, column but the unstructured data are in the form of audio, video , image, blogs, emails, tweets, forums may unformed data which are difficult to

processed. The Analytics is the major role is to processed this type data and took our something meaningful data form those data. In a basic manner Big information is structure and unstructured information which move all through inside or outside of association firewall. the Analytics is a processed in which a few Models are construct utilizing scientific and factual algorithms. This model is predictive in nature and have the knowledge data and this data permits to make a move.

The unstructured information like emails, web clicks, PS signal, data from ATM's, social media post thus on. Big information is simply enormous measure of genuine information from different sources. Enormous Data Analytics is only instruments or we can say that the procedures of cleaning the knowledge from huge measure of information This information will be ordered by high speed, huge measure of information or extraordinary variety. Big Data Analytics plan to inferring amendment and conclusion. This task is finished utilizing apparatuses like Hadoop, R, Python and SAS and so on.

The financial division is the segment of the economy committed to the holding of budgetary resources for other people, contributing those money related resources as influence to make more riches and the guideline of those exercises by government organizations. There is hearty interest in Indian bank division it increments in working populace and developing discretionary cashflow will raise the interest of banking and related services. Rural banking is likewise expected to increment late on. Healthy business fundamentals can be reflected through high interest margins. Banking segment in India has been changed totally. Governments from all around the globe and organizations of all sizes are effectively getting connected with enormous information ventures since this information has a story to advise that can possibly change business tasks to improve things.
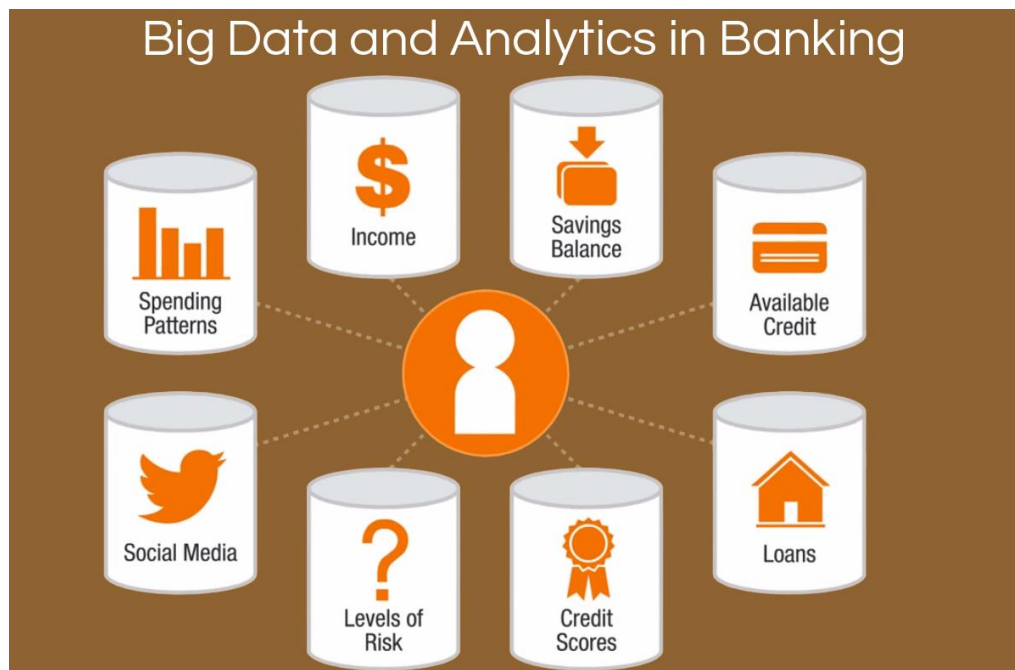
Big data analytics in digital bank & fraud detection:

The huge information upset occurring in and around 21st century has discovered a reverberation with banking firms, considering the important information they've been putting away since numerous decades.

Source:- https://arunkottolli.blogspot.com//2018/07/big-data-analytics-for-digital-banking.html
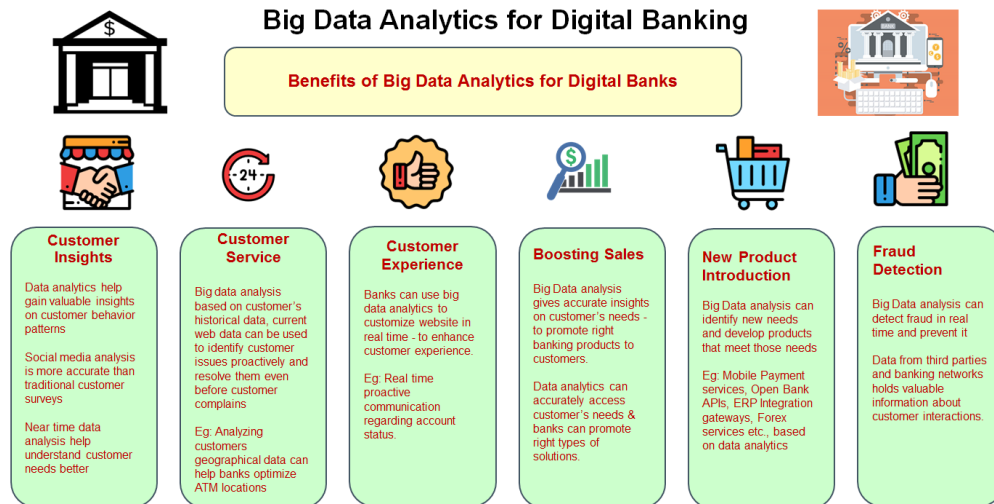
Numerous people using their smart phone to access many online services. The mobile banking which is digital banking increasingly Various banks are advancing compact money related organizations which license bank customers to check balance in their very own record, to move resources among records and make online portions wherever and at whatever point by simply using flexible monetary applications presented on their mobile phones.

Shockingly, versatile malware is rapidly expanding in recurrence and advancement in the previous couple of year and has caused an assortment of harms including spilling of delicate monetary information, budgetary misfortune and identify robber Advanced Banking has been created as a powerful and helpful channel for monetary foundation to disseminate their administrations to client.

This computerized bank become easily accessible to client through shrewd gadgets. Portable malware is rapidly expanding in caused an assortment of harms including spilling of touchy money related information, monetary misfortune and recognize robbery. There are customer insights, services, experiences, sales and new products information in digital bank.int customer insights data analytics help to gain more valuable values on customer behaviour patterns. This data is gathered from the various source like social media and their time spending on online services based on that they analyzing the accurate data form different customer surveys. This analysis helps the bank to understand the cutomers need better.

In the customer services there is data analysis is done which is based in the customer past data and based on this this the organization find the customer issues and they solve them before customer approached the bank for the problem.

The data analytics help the bank to enhance their services through the customer data analysis. This analysis provides customer a better customer experience. The big data analysis give the accurate valuable insight on the customers requirement and need which tends to better banking product to the customers.

The sales and the new product are introducing based this type of analysis. This type of data helps the bank to improve their product like there payment gateway services based on this analysis. The fraud detection is done in real time and prevent it using this analysis. This fraud is happening when people used open network to transfer the fund from one account to another account so it is possible that some can damaged the network and takes your payment information. So, through this analysis real time fraud detection is happen and prevent also.

There are many threats to the digital banking which is malwares and many viruses. Some malware which affecting computer and online service on the smart phone. The cybercriminal has been refining these viruses and malwares to target the mobile devices for acces the customer transaction information. This threat come for mt he third party application which secretly tamper an existing bank application which already installed in mobile devices.

The protection against this threat is having many security mechanisms such as authentication, connection encryption, authorization, antivirus application can enhance the security threats. Mobile banking, NFC Payments, Online Budgetary devices, Social Media Banking, Video visit warning, and so forth are taking banking administrations to the following dimensions. These activities are making banking administrations more services and simple to get to[18].

Application of big data in banking sector:

Tremendous Data Analytics Applications are another arrangement of programming applications that impact enormous scale data, which is ordinarily too immense to even consider evening think about fitting in memory or in fact, even on one hard drive, to uncover noteworthy getting the hang of using gigantic scale parallel-planning establishments. This big data comes from various running resources, email, twitter during sports event, stock market or the data from any other rapidly growing data-intensive software system.

Security and privacy issue:

Gigantic data gives understanding into various baffling parts of individual's life including their lifestyle, needs, and tendencies of their customers so it is straightforward for banks to tweak organizations to the necessities of each individual.

With quick advancement in development, frameworks organization and cost decline away devices, data uprising has happened naming it as Big data. Tremendous data is suggested as one in which huge measure of data can be assembled, secured and separated at reasonably ease. Numerous associations like portable organization, protection organizations, banks offering Loans and so on gather the individual data of an individua like telephone number, addresses, email data and so on from different sources and use it for their own uses, making issues the general people & customer. The customer even doesn't know the how their data are distributed in different organization. Despite the fact that the rise of huge information give huge benefits, yet has likewise activated protection concerns.

Detecting financial fraud:-

Banks create a lot of information: paper records with marks, financial records, versatile banking, credit and platinum cards, loans, etc. Today numerous information originates from the Bank's contacts with clients in online mode. Such information like messages, visit logs, sustains, posts, web logs and semi-organized information such as client surveys.

Some of fakes in the banking are in the web-based banking (charge card, web or portable exchanges) where the fraudster plays out the exchanges with a similar code or indication of client. The bank, so as to prevent the achievement of such fake exchange while running you should give a client profile dependent on the narrative of its money related exchanges.

Through the investigation of the considerable number of exchanges made by clients you can make a "client profile" and its relations with different reporters and installment strategies utilized. Each new exchange is contrasted with the profile to distinguish if an activity fits into the typical ones and if not, to report doubt of fraud.

Marketing and customer:

An enormous Indian bank understood that their promoting efforts were incapable and it required a long investment to assess the aftereffects of the showcasing effort. Further, it set aside more effort to introduce these discoveries to the senior administration of the bank. Bank chose to apply examination by associating all channels of correspondence with the client to an incorporated framework. Advertisers have started to utilize facial acknowledgment programming to figure out how well their publicizing succeeds or comes up short at animating enthusiasm for their items.

Banks can gather information on item highlights, input and call focus data to take choices identified with item structure. Versatility displaying for evaluating can be estimated for items to decide the value request relationship. The interest forecast can help banks to decide spending assignment and branch level advertising choices

Tools used in Big data:

Huge information manages numerous information types like structure, unstructured and furthermore volume of information is excessively huge so to investigate these information we need some device like Hadoop, NoSQL, MapReduce, PythonBig information manages numerous information types like structure, unstructured and furthermore volume of information is excessively enormous so to break down these information we need some device like Hadoop, NoSQL, MapReduce, Python, R-Language, WEKA and KNIME, R-Language, WEKA and KNIME.

Hadoop :

Hadoop is an open source programming, that stores and procedure enormous information, Hadoop manages and handle enormous data using cluster approach.this frame work is build on java.Hadoop is processed large amount of data or complesx dataon different operating system like windows, llinux and unix based system.

NoSQL:

NoSQL stand for not only Sql.this tool is used for handling structure as well as unstructured data. Means information can be gotten to by SQL inquires just as a few new methods can be utilize. NoSQL is open source programming and uses DaaS (Database as a Service). As Graph databases are progressively prominent answer for huge information
NoSQL is the correct application who oversees it productively.

MapReduce:

Mapreduce is used to processed large amount of data very easily. MapReduce can be ordered in two sections Guide changes over the information into another sort of information and partitions into information tuples like key/esteem pair and Reduce takes contribution from Guide and joins information tuples into little set of tuples.

R-Language:

R is customizing language for measurable values, complex information and graphical data. Powerful information dealing with and capacity should be possible by R-language. R gives graphical offices to information examiner furthermore, has numerous device to perform information investigation.

Python:

Python is a broadly useful programming language which empowers software engineers to compose less lines of codes and make it progressively intelligible. It has

scripting highlights what's more that utilizations many propelled libraries, for example, Numpy, Matplotlib, and Scipy which makes it helpful for logical registering. Python gives countless libraries to take a shot at Big Data.

Spark:

In the event that you have huge information which may work better in gushing structure (constant information, log information, API information), at that point Apache's Spark is an incredible device. PySpark, the Python Spark API, enables you to rapidly get ready for mapping and lessening your dataset. It's likewise inconceivably well known with AI issues, as it has some worked in calculations.

Big Data Analytics Challenges:

While implementing the big data analytics the segmentation from the Data cluster is the big challenge. Using clustering through a straightforward point and snap exchange, clients can consequently discover bunches inside information dependent on explicit information measurements. Using clustering (K-implies calculation) through a straightforward point and snap exchange, clients can consequently discover bunches inside information dependent on explicit information measurements. With grouping, it is then easy to recognize and address bunches by client type, content reports, items, quiet records, click way, conduct, obtaining designs, and so on. At whatever point new innovations develop, they meet with new difficulties in every one of the angles. Once the useful difficulties are set up, the following family is the specialized challenges. Enormous information faces numerous specialized difficulties which are on the roadway of the examination.

Failure handling:

Frameworks can be contrived so that the likelihood of disappointment must fall inside the allowed limit. Adaption to internal failure is a specialized test in enormous information[1]. At the point when a procedure began it might include with various system hubs and the entire calculation procedure winds up awkward. Holding check focuses and fixing the limit level for procedure restart if there should arise an occurrence of disappointment, are more noteworthy concerns.

Data heterogeneity:

Huge information manages unstructured, semi-organized what's more, organized information[1]. Connecting unstructured information with organized information, changing over information form one structure into another required structure needs a great deal of research.

Data quality:

Quality Colossal measure of information relating to an issue is without a doubt a major resource for both Business just as IT pioneers[1].

Indian banking sector growth:

The Indian banking system consists of 12 public sector banks, 22 private sector banks, 46 foreign banks, 56 regional rural banks, 1485 urban cooperative banks and 96,000 rural cooperative banks in addition to cooperative credit institutions As of September 2021, the total number of ATMs in India reached 213,145 out of which 47.5% are in rural and semi urban areas.
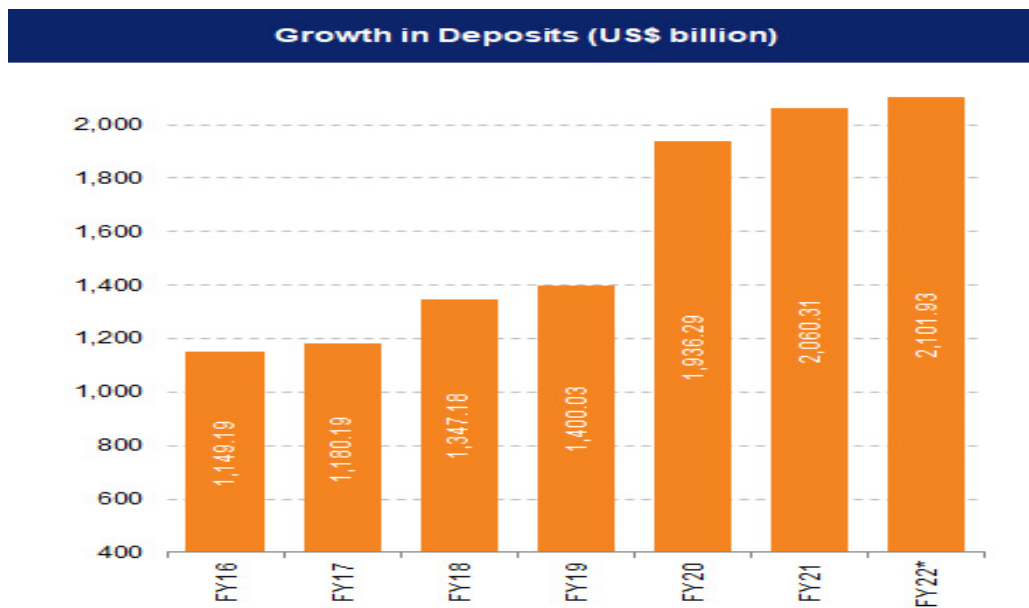
In 2020-2022, bank assets across sectors increased. Total assets across the banking sector (including public and private sector banks) increased to US$ 2.67 trillion in 2022.

In 2022, total assets in the public and private banking sectors were US$ 1,594.51 billion and US$ 925.05 billion, respectively.

During FY16-FY22, bank credit increased at a CAGR of 0.62%. As of FY22, total credit extended surged to US$ 1,532.31 billion. During FY16-FY22, deposits grew at a CAGR of 10.92% and reached US$ 2.12 trillion by FY22. Bank deposits stood at Rs. 173.70 trillion (US$ 2.12 trillion) as of November 4, 2022.

According to India Ratings & Research (Ind-Ra), credit growth is expected to hit 10% in 2022-23 which will be a double- digit growth in eight years. As of November 4, 2022 bank credit stood at Rs. 129.26 lakh crore (US$ 1,585.09 billion).
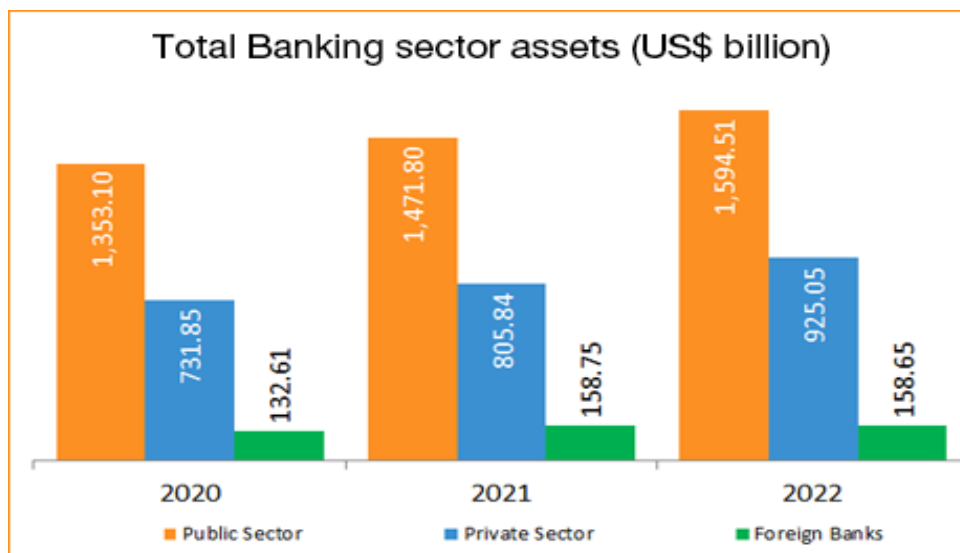
Non-food bank credit registered a growth of 17.6 per cent in November 2022 as compared with 7.1 per cent a year ago on the back of robust credit demand from the segments such as services, industry, personal, and agriculture and allied activities, according to RBI's statement on Sectoral Deployment of Bank Credit.

Growth in Deposits (US$ billion)

FY16	1,149.19
FY17	1,180.19
FY18	1,347.18
FY19	1,400.03
FY20	1,936.29
FY21	2,060.31
FY22*	2,101.93

Source:- https://www.ibef.org/assets/images/banking_growth_in_deposits.jpg

Industry Contacts

- **Indian Banks Association**
- **Reserve Bank of India**
- **Institute for Development and Research in Banking Technology**

Total Banking sector assets (US$ billion)

	Public Sector	Private Sector	Foreign Banks
2020	1,353.10	731.85	132.61
2021	1,471.80	805.84	158.75
2022	1,594.51	925.05	158.65

Source:- https://www.ibef.org/uploads/industry/banking_cluster.jpg

Conclusion:

This paper is basically a critical analysis of numerous research papers according to my understanding. I have clubbed Big Data and Analytics from different research paper, blogs etc. To discuss its potential in the Banking Sector which is in boom nowadays. I have also covered big data analytic in Digital and Fraud Detection Schemes. After that, i have also discussed the challenges and rectifiable issues seen in Big Data as an analytic tool. It is clear from the research that Big data has huge applications in other sectors such as Healthcare, Agriculture, Food, to name a few apart from the aforementioned. Through this review, my aim was to give a generic introduction to Big Data's use in the field of Banking. It is pretty visible that with the growing data generation seen every day, Big Data is proving itself to be a powerful analytics field which can be used to our benefit as discussed and much more in the coming future.

Reference:

1. TRANSFORMATIONAL ISSUES OF BIG DATA AND ANALYTICS IN NETWORKED BUSINESS
   https://pdfs.semanticscholar.org/ofc1/5b775475b97c5e29d1422af5df6afef*e44.pdf

2. .Examining Security Risks of Mobile Banking Applications through Blog Mining Wu He, XinTian & Jiancheng Shen Old Dominion University, Norfolk, VA, USA
   https://www.researchgate.net/profile/Xin_Tian17/publication/282376655_Examining_security_risks_of_mobile_banking_applications_through_blog_mining/links/5718508ae986b8b79eaf2/Examining-security-risks-of-mobile-banking-applications-through-blog-mining.pdf

3. Fintech reloaded – Traditional banks as digital ecosystems With proven walled garden strategies into the future
   https://www.dbresearch.com/PROD/RPS_EN-PROD/PROD0000000000451937/fintech_reloaded_%Do_Traditional_banks_as_digital_ec.PDF

4. Investigation into Big Data Impact on Digital Marketing K. Grishikashvili, S. Dibb, M. Meadows Business School Department of Strategy and Marketing Open University, Uk
   https://www.cmdconf.net/2014/pdf/24.pdf

5. Customer relationship management with big data enabled in banking sector Muhammad Anshari*, Syamimi Ariff Lim E Government Innovation Centre, Universiti Brunei, Darussalam, Brunei
   https://jsrad.org/wp-content/2016/Issue%204,%202016/1jj.pdf

6. Big Data in Banking: Opportunities And Challenges Post Demonetisation in India Ayesha Anam Siddiqui1 , Dr. Riyazudding Qureshi21 Department of Management Sciences & Computer Studies, Maulana Azad College of Arts, Science And commerce, Aurangabad(MAH), India2Department of Management Sciences & Computer studies, Maulana Azad College of Arts, Science And Commerce, Aurangabad(MAH), India
https://www.researchgate.net/profile/Riyazuddin_Qureshi/publication/313836902_Big_Data_In_Banking_Opportunities_And_Challenges_Post_Demonetisation_in_India/links/58a87aaba6fdcc0e0790fd49/Big-Data-In-Banking-Opportunities-And-challenges-Post-Demonetisation-in-India.pdf-Post-Demonetisation-in-India.pdf