



Big Data in Banking – Case Studies & Applications

UNIT 1

Applications of Big Data in Banking Sector

- Big Data is renovating the world and it has left no industry untouched with its enormous benefits. It has emerged as a ***lifeguard*** for the Banking Industry.
- Saved a ***lot of revenues*** from the banking firms so far.
- These are some applications of Big Data in Banking sector-
 - Risk Management
 - Fraud Detection
 - Customer Contentment

Risk Management



Establishing a robust risk management system is of utmost importance for banking organizations or else they have to suffer from huge revenue losses.

To stay alive in the competitive world and increase their profit as much as they can, organizations have to keep innovating new things.

Through Big Data Analysis, firms can detect risk in real-time and apparently saving the customer from potential fraud.

UOB Singapore

- The United Overseas Bank (UOB) Limited, the 3rd-largest bank in SouthEast Asia, has leveraged Big Data to direct ***risk management***, the biggest area of concern for any banking organization.
- UOB took a gamble with employing a risk management system that is based on Big Data.
- Calculating the ***Value at Risk (VaR)*** is a time-consuming effort, usually taking up to ***20 hours***.
- Through its Big Data risk management system, UOB was now able to do the same task in just a ***few minutes*** and with the aim of doing it in ***real-time*** pretty soon.

Fraud Detection

The rapidly growing digital world is furnishing us with numerous benefits but on the other hand, gives birth to various kinds of ***FRAUDS*** as well.

Our personal data is now more vulnerable to ***cyber attacks*** than ever before and it is the biggest challenge a banking organization faces.

Employing Big Data Analytics with some ***ML Algorithms***, organizations are now able to ***detect frauds*** before they can be placed.

This is done by identifying ***unfamiliar spending patterns*** of the user, predicting unusual activities of the user, etc.

Danske Bank

Danske Bank, with a customer base of more than 5 million, is the largest bank in Denmark.

The bank was struggling with its fraud detection methods having a very low percentage i.e. only a **40%** fraud detection rate and managing up to **1200** false positives per day.

With Teradata, a leading database & analytics service provider company, they employed some advanced Big Data analytics for improving their fraud detection techniques.

The bank saw a **60%** reduction in false positives, expecting it to soon reach an **80%** mark and an increase in the true positive rate by **50%**. They also observed a massive operating profit of **\$70 million in 2018**.

Customer Contentment

Considering the high amount of risk involved when you deal with the banking firms, to ensure the ***satisfaction*** of a customer is one of the ***most challenging*** tasks for them.

From ensuring the ***safety*** of their transactions to providing them the most relevant & beneficial offers, ***customer retention*** is a lifetime journey for the banking firms.

The data that they collect from their customers is now more important than ever. Analyzing their customer's data on the basis of different parameters helps them in ***targeting*** their customers in a much better way.

JPMorgan Chases Big Data

JPMorgan Chase & Co. is the largest bank in the United States and the 6th-largest in the world.

Also, it is the world's **most** valuable bank in terms of market capitalization. With a customer base of over 3 billion, the AMOUNT of data it generates is UNIMAGINABLE including a vast amount of credit card information and other transactional data of its customers.

They have adopted Big Data technologies, mainly **Hadoop**, to deal with this data.

By employing Big Data Analytics, they are now able to generate **insights** into customer trends and the same reports are offered to its clients.

They are able to **analyze a customer individually** and these reports are generated within seconds.

Big Data Extricating The Bank of America

Bank of America is one of the largest banks in the United States with a customer base of around **70 million**.

In the year 2008, they realized that their customer base was ***declining*** at an alarming rate as they saw their customers shifting towards smaller banks.

Big Data Analytics came to their rescue. Through ***analyzing*** their customer's data from a variety of sources such as their website, call center logs and personal feedbacks, they ***discovered*** that their ***end-to-end cash management system*** was too stiff for the customers as it hindered their freedom to access trouble-free & flexible cash management system.

Though smaller banks were offering an effortless solution to it. Ultimately, they decided to end their all-in-one offering.

Soon in the year 2009, as a solution to these problems, they launched a website that was a ***more flexible*** online product, **CashPro Online**, and its mobile version, CashPro Mobile later in the year 2010.

This was developed with an aim to provide their customers with a 1-stop solution for all the services they offer.

Credit Card Fraud Detection

- It has become a challenge for Credit card companies to identify whether the requested transaction is fraudulent or not.
- A credit card transaction hardly takes **2-4 seconds** to completion. So, the companies need an innovative solution to identify the transactions which may appear as fraud in this small time and thus protect their customers from becoming its victim.
- An **abnormal number** of clicks from the same IP address/ a pattern in the access times — although this is the most obvious and easily identified form of click fraud, it is amazing how many fraudsters still use this method, particularly for quick attacks.
- They may choose to strike over a **long weekend** when they figure you may not be watching your log files carefully, clicking on your ad repeatedly so that when you return to work on Tuesday, your account is significantly depleted. Part of this fraud might be unintentional when a user tries to reload a page.
- Again, if you have made any transaction from **Mumbai today** and the very next minute there is a transaction from your card in Singapore. Then there are chances that this transaction may be fraud and not done by you.
- So, companies need to process the data in **real time** (Data in Motion analytics DIM) and analyze it against individual history in a very short span of time and identify whether the transaction is actually fraud or not. Accordingly, companies can accept or decline the transaction based on the severity.
- To process the **data streams** we need streaming engines like **Apache Flink**. The streaming engine can consume the real-time data streams at very high efficiency and process the data in low latency (without any delay).

Customer Churn Analysis

Churn analysis is the calculation of the *rate of attrition* in the customer base of any company. It involves identifying those consumers who are most likely to discontinue using your service or product.

The best way to manage these issues will be to *predict* the subscribers who are likely to churn, *well in advance* so that business can take required measures to mitigate it and win-back the customers or re-activate the sleeping base.

To find out the *root cause* of Customer churn companies need to analyze huge data in the range of TBs to PBs

Companies need to go through billions of customer complains which are stored for years and get them resolved with immediate effect.

Data from social media, where users write their opinion about the product, can identify if customers are liking their products or not.

Let us take an example of Call Centre Analysis. Here the data used is Call Log & Transactional data. Many banks are integrating this call centre data with their transactional data warehouse to reduce churn, and increase sells, customer monitoring alerts & fraud detection.

Apache Flink offers an opportunity to tap into the many internal & external customer interaction and behavioral data points to detect, measure and improve the desired but illusive objective of consistent & rewarding Customer Experience success.

Sentiment Analysis



A basic task in sentiment analysis is ***classifying the polarity*** of a given text at the document, sentence, or feature/aspect level — whether the expressed opinion in a document, a sentence or an entity feature/aspect is ***positive, negative, or neutral***.



Advanced, “***beyond polarity***” sentiment classification looks, for instance, at emotional states such as “angry,” “sad,” & “happy.”



Here, ***language*** is processed to identify and understand consumer ***feelings & attitudes*** towards brands or topics in online conversations ie, what they are thinking about a particular product or service, whether they are happy or not with it, etc.



For example, if a company is launching a new product, it can find what its customers are thinking about the product.



Whether they are ***satisfied with it or not*** or they would like to have some modifications in it can be found out using Big data by doing sentiment analysis i.e, using sentiment analysis we can identify users’ opinion about the same.



Then the company can ***take action*** accordingly to modify or improve the product to increase their sales and to make customers feel happy with their product.

Sentiment Analysis

- A large airline company started monitoring ***tweets*** about their flights to see how customers are feeling about upgrades, new planes, entertainment, etc. Nothing special there, except when they began feeding this information to their customer support platform and solving them in real-time.
- One memorable instance occurred when a customer tweeted ***negatively*** about lost luggage before boarding his connecting flight. They collect the tweets (having issues) and offer him a free 1st class upgrade on the way back.
- They also tracked the luggage and gave information on where the luggage was, and where they would deliver it. Needless to say, he was pretty shocked about it and ***tweeted like a happy*** camper throughout the rest of his trip.
- With ***Hadoop***, you can ***mine*** Twitter, Facebook and other social media conversations for sentiment data about you & your competition, and use it to make targeted, real-time, decisions that increase market share.
- With the help of ***quick analysis*** of customer sentiment through social media, company can immediately take decision & action and they need not wait for the sales report (which might take 6 or more months also) as earlier to run their business in a better manner.

Income Tax Department to Scrutinise Bank Accounts

On 8th Nov-16, our PM announced that from the same day the Rs.500 & Rs.1000 currency notes would be discontinued to track black marketers and the black money they carry.

He also announced that anyone having money in these denominations can get them exchanged or deposit them in their accounts with some limitations that were also declared.

The question was – How did the government or IT department track the black money that has been deposited and how they segregated black money holders from genuine tax payers?

With more than 1.35 Billion populations and 100s of millions of bank accounts, it is a big question that how IT department will find out discrepancies?

Income Tax Department to Scrutinise Bank Accounts

It is practically ***impossible*** for IT officials to manually compare tax/account data with the data that it is collecting from banks.

Big Data Analytics is being used to look into the ***inconsistencies*** which will then be thoroughly inspected by tax officials.

The Big Data tools won't be just used to collect details about the money being deposited by individual but it will also be used to gather ***information*** like tax paid by the individual in past years, corporate tax that his company has paid, number of employees he has in his company and if they have also deposited any money recently and other tax related data.

If discrepancies are observed, these Big Data tools will raise ***red flags***/alerts. This would help tax officials to finally come to conclusion regarding his genuineness or be a black marketer.

As Big Data can handle ***huge volumes*** of data in the range of PBs or 100s of trillions of records, it will help IT professionals to scrutinize the Transactions / Account details / Tax information efficiently and quickly.

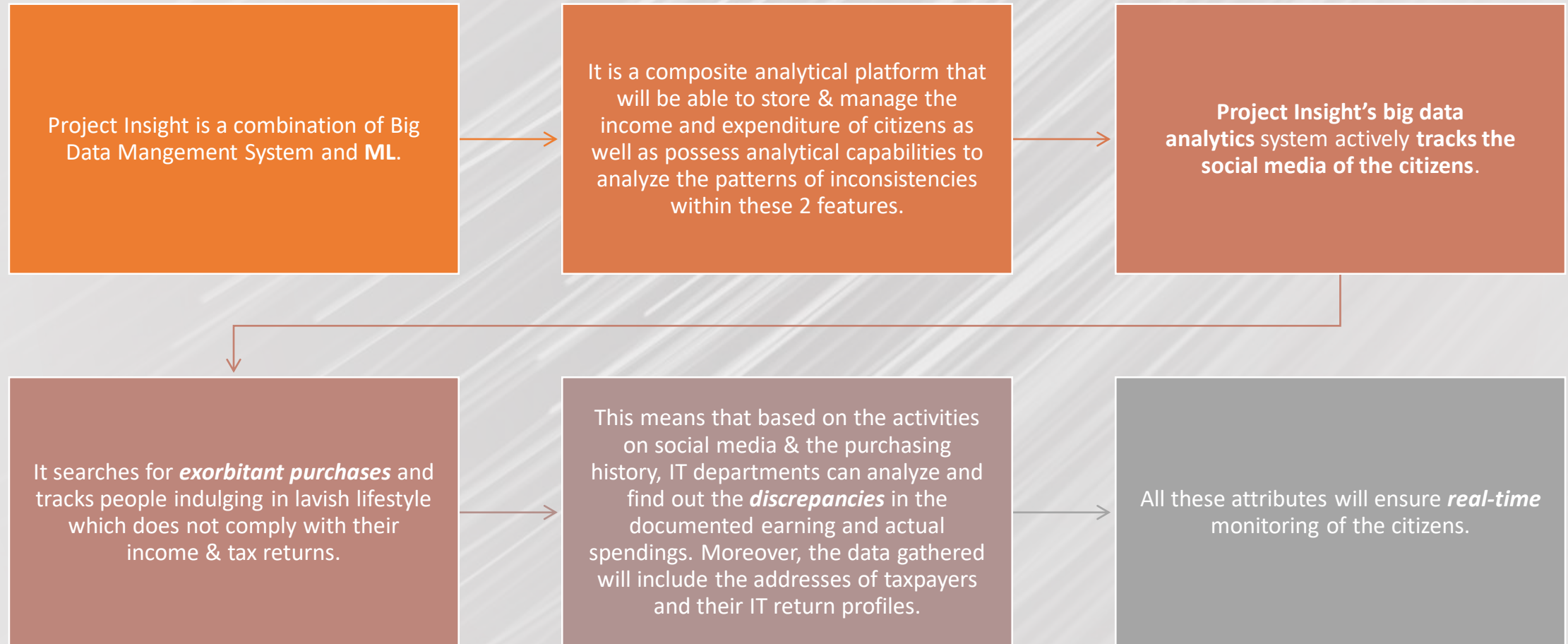
IT Department Use Big Data to Catch Tax Evaders

- In India problems like failure to pay tax dues, submission of false tax returns, inaccuracy in financial statements, lack of income reports, storing wealth in foreign countries is becoming an increasing trend.
- Many times, people purchase products or indulge in a lavish lifestyle beyond their monetary means as per their tax returns.
- This does not often come up on the radar of tax officials as it is not possible to analyze all of the information on a manual level.
- However, with the emergence of **Big Data & Data Analytics technologies**, it is now possible to analyze all of the large scale information.
- The 2016 operation to curb the usage of black money provided a key revelation to the IT Department to upgrade their tax collection procedures and utilize an efficient methodology to track tax evaders. *This led to the inception of **Project Insight**.*

Income Tax Department Will Use Big Data to Catch Tax Evaders

- India is now one of the several countries that use Big Data to track tax evasion. In the United States, IRS has been using Big Data on its phone surveillance records, tracking social media accounts and using extensive **data mining** to develop analytical algorithms for identification of tax compliance issues.
- Similarly, the **UK Government had implemented big data analytics in their income tax segment** and observed an increase of \$ 5.4 billion in tax return revenue.
- With a rapid expansion in data, it is becoming increasingly **difficult to analyze all** of it through traditional methodologies. Moreover, a country of the **huge population** like India cannot only rely on traditional methods of survey.
- Many of the analytical problems could not be solved before due to **high expenses**. Also, the analytical platforms could not handle such large volumes of data. Moreover, most of the data collected were not present in the electronic form.
- All these problems made it difficult to manage & analyze large amounts of data. Therefore, we need Big Data to analyze an **astronomical amount of data** that is generating every day.
- In the light of the above issues and the limitations in the current technology, the IT Department has introduced Big Data Analytics to monitor & track the social networking profiles of people and find patterns of inconsistencies between their income and spending.

What is Project Insight?



What is Project Insight?

There is also a **business intelligence** platform that includes large scale data warehousing and analysis of data. It will help the income tax officials to track high-value transactions and inhibit the flow of black money.

The government allied itself with **L&T Infotech** to implement this project. It combines the databases of IT returns, TDS & TCS statements, IT forms and forms from financial institutions of the country.

Also, as part of Project Insight, a Compliance Management Centralised Processing Centre handles verification, management of campaign and the generation and follow-up of letters & notices.

The IT Department has introduced a **360 Degree vision** profiling that allows them to identify non-filers of tax through social media monitoring.