Data Ware house and Data Mining

Like many industries, Banking sector is facing the challenges from change in customer demographics, global socio-political issues and understanding the consumers from millennials and generation 'z'. Online banking is also making banking industry to respond in real-time. Great recession also mandated Federal Govt to impose new compliance regulations to report to Fed Authorities. Compliance reporting will be manageable with Data Warehouse which keeps all the data at one place.

Data Warehouses provide a consolidated data view of the past transactions and is not useful without a proper strategy to convert that into information.

Mostly banking use-cases can be classified as below:

Risk: With growing terrorism and other anti-social activities, banks needs to monitor their Risk and Fraud across their internal anti–money laundering and counter–terrorist financing (AML/CTF) controls. DW provides finding the data patters to identify such activities and help setting up some of those alerts.

Performance/Profitability: With new digital technology, banks needs to be more smart data aware and analyze data from various sources. steady introduction of new and disruptive technologies are contributing to a volatile environment and expanded regulatory reach. DW will help getting the insights into such volatile environment and help bank leadership to make timely decisions.

Customer Insight: Banks are now creating predictive models to single segment customer's needs. DW in banks help continuously optimizing its cost model while striving to increase the number and profitability of its customers—and enhance the bottom line.

Compliance:

Properly designed DW will help assessing risk reporting and analytics, including risk and compliance assurance, independent model validation, quantitative assessment and back testing, risk and finance data governance and integration, and regulatory compliance monitoring and assessment (including CCAR, DFAST, OCC etc.)

Data warehousing is important for the banking sector because it helps in the following ways:

- Identify the potential risk of default and manage and control collections

- Performance analysis of each product, service, interchange, and exchange rates
- Track performance of accounts and user data
- Provide feedback to bankers regarding customer relationships and profitability

Data warehouse providers allow modern banks to address many challenges related to data quality and management. Custom or modernized DWH solutions can help you overcome these challenges faster and get quick ROI. Below are the most frequent issues that can be solved by implementing the right DWH technology.

1. **Reduced costs-** DWH solutions enable the right support to several disparate data systems which was earlier expensive while using legacy data storage systems. Data preparation time will be reduced, as well.
2. **Data consistency-** Resolve the challenges of unstructured data, minimize errors in operational and financial data, and improve data accuracy. DWH allows for organizing your data and getting a single view on your business.
3. **High scalability-** With legacy data storage systems in place, you definitely face the lack of potential for data volume growth. Modern data warehousing solutions ensure high scalability and the right business continuity.
4. **Quick access-** Among other advantages data warehouse development services can bring to your bank are seamless access to data, increased customer attraction and loyalty thanks to improved customer service.

The appropriate data management in banks ensures information quality and consistency. Data warehousing development services help banks build more effective processes for data collection, processing, storage, exchange, analysis, etc. Gathering information from various sources and converting it to valuable insights are the main objectives of DWH software.

Through the right technology, we help banks use their data as efficiently as possible. Our expertise allows us to build, upgrade, modernize DWH solutions, and make them addressing your current and future organizational needs.

DATA MINING IN BANKING AND ITS APPLICATIONS

Banking systems collect huge amounts of data on day to day basis, be it customer information, transaction details, risk profiles, credit card details, limit and collateral details, compliance and Anti Money Laundering (AML) related information, trade finance data, SWIFT and telex messages. Thousands of decisions are taken in a bank daily. These decisions include credit decisions, default decisions, relationship start up, investment decisions, AML and Illegal financing related. One needs to depend on various reports and drill down tools provided by the banking systems to arrive at these critical decisions. But this is a manual process and is error prone and time consuming due to large volume of transactional and historical data. Interesting patterns and knowledge can be mined from this huge volume of data that in turn can be used for this decision-making process. This article explores and reviews various data mining techniques that can be applied in banking areas. It provides an overview of data mining techniques and procedures. It also provides an insight into how these techniques can be used in banking areas to make the decision-making process easier and productive.

Banking as a data intensive subject has been progressing continuously under the promoting influences of the era of big data. Exploring the advanced big data analytic tools like Data Mining (DM) techniques is key for the banking sector, which aims to reveal valuable information from the overwhelming volume of data and achieve better strategic management and customer satisfaction. In order to provide sound direction for the future research and development, a comprehensive and most up to date review of the current research status of DM in banking will be extremely beneficial. Since existing reviews only cover the applications until 2013, this paper aims to fill this research gap and presents the significant progressions and most recent DM implementations in banking post 2013. By collecting and analyzing the trends of research focus, data resources, technological aids, and data analytical tools, this paper contributes to bringing valuable insights with regard to the future developments of both DM and the banking sector along with a comprehensive one stop reference table. Moreover, we identify the key obstacles and present a summary for all interested parties that are facing the challenges of big data.

Value Creation of DM in Banking by Topics

Having reviewed over 100 recent DM applications in banking post 2013, it can be generally concluded that the banking sector mainly adopts DM techniques for the following purposes:

Security and fraud detection: Big secondary data like transaction records are monitored and analysed to enhance banking security and distinguish the unusual behavior and patterns indicating fraud, phishing, or money laundering (among others).

Risk management and investment banking: Analysis of in-house credit card data freely accessible for banks enables credit scoring and credit granting which form part of the most popular tools for risk management and investment evaluation.

CRM: DM techniques have been widely applied in banking for marketing and customer relationship management related purposes such as customer profiling, customer segmentation, and cross/up selling. These help the banking sector to have a better understanding of their customers, predict customer behaviour, accurately target potential customers and further improve customer satisfaction with a strategic service design.

Other advanced supports: A few less mainstream applications focus on branching strategy, and efficiency and performance evaluation, which can significantly assist in achieving strategic branch locating and expansion plans.

In what follows, we briefly review the collected publications with respect to the areas of interests. Moreover, to serve as the one stop reference directory for all recent DM applications in banking post 2013, a summary table can be found in **Table 1**, where the literature are grouped by means of value creation, along with detailed information like data resources/regions and DM techniques adopted. Note that most implementations are applied to more than one DM technique, and some applications did not clarify the specifics due to confidential restrictions.

Sector	Key Techniques	Regions	Purposes
Security and fraud detection	classification (DT, NN, SVM, NB), k-mean clustering, ARM	Australia [12], Latin-America [29], Greece [24], Germany [32], Belgium [21], UCI Repository [15,16,17,18]	Identifying phishing, fraud, money laundering, credit card fraud, security trend of mobile/online/traditional banking.

Sector		Key Techniques	Regions	Purposes
Risk management and investment banking		classification (DT, NN, SVM, NB, LR), k-mean clustering	UCI Repository International Dataset [41,42,45], Australia [51], Iran [37], Indonesia [34], China [35], German [36,38,39,51], Taiwan [51], US [49], Canada [46]	Credit scoring, credit granting, risk management for peer-to-peer lending.
CRM	**Customer profiling and knowledge**	classification (DT, NN), k-mean clustering	Jamaica [52]	Efficiently build accurate customer profiles.
	Customer segmentation	k-mean clustering	Iran [54]	Provide sufficient customer segmentation, conduct customer-centric business strategies.
	Customer satisfaction	k-mean clustering, classification (NN)	Spain [56]	Make the most strategic investment on maintaining and enhancing customer satisfaction.
	Customer development and customization	classification (DT, NN, NB, LR, SVM), k-mean clustering	Portugal [57,58,59,60,61,62], Turkey [66,67], China [68], Taiwan [65], UCI Repository [64]	Strategic banking via direct marketing, targeted marketing, product cross/up selling.
	Customer retention and acquisition	classification (DT, NN, LR, SVM), ARM, k-mean clustering	EU [69], China [70], Nigeria [71], Croatia [73], Bangladesh [75]	Customer churn prediction and prevention, attracting potential customers and strategic future service design.
Other advanced supports		classification (NN, DT, SVM), k-mean clustering	Nigeria [77], Turkey [78,81], Canada [80], ASEAN [82], Islamic banks [83], BRICS [84], US [79]	Branch strategy, bank efficiency evaluation, deposit pricing, early warning of failing bank.

1,*