

Predictive Analytics Application in Banking Sector

^{#1}Priyanka Daulatabad, ^{#2}Priyanka Andhare, ^{#3}Rituja Rode,
^{#4}Riya Bhutada



¹priyanka.daulatabad@gmail.com

²priyankaandhare12@gmail.com

³ritujarode@gmail.com

⁴riyabhutada27@gmail.com

^{#1234}Department of Computer
NBSSOE, Pune.

ABSTRACT

The Banking industry is one of the largest adopters of analytics services. In today's competitive world, growing customer base and fulfilling them is considered the most challenging task. Banking has become customer-centric: knowing customers' needs and preferences are key to building customer loyalty and customer retention over the long term. Banks are increasingly utilizing analytics to gain a competitive advantage and to form conclusions and insights based on the information they have gathered through basic reporting and data collection. From this gathered information, data patterns are derived and analyzed. Predictive analytics uses these patterns and histories to determine what could happen in the future—how customers might act, what services will be profitable, and so on. Analytics allow financial institutions to target and engage customers constantly and not just when they physically enter a branch. By applying analytics to customer segmentation, customer profiles, and transaction patterns, banks have the opportunity to gain valuable insights on their customer base and can translate this knowledge into increased customer satisfaction and retention. They can also use analytics to offer customized products, services, and deals to customers based on their profiles and histories. The aim of the project to develop a Machine Learning model to perform predictive analytics on the banking dataset. The banking data set consists of details about customers like and whether the customer will buy a product provided by the bank or not with the help of data mining, we look for service pattern in a group and discover relationship in the data of the banking customer.

Keywords: Data Mining, Bank Database, Services.

ARTICLE INFO

Article History

Received: 6th May 2018

Received in revised form :
6th May 2018

Accepted: 8th May 2018

Published online :

10th May 2018

I. INTRODUCTION

Banking industry has particularly benefited from the development in digital technology. Concept of data stored at branches has given way to centralized databases. Number of channels to access bank accounts has multiplied. Banking systems have become technically strong and customer oriented with online transactions, electronic wire transfers, ATM and cash and cheque deposit machines. As number of channels has increased so is the number of transactions and the related data stored. So currently banks have huge electronic data repositories in their computing storage systems. Data has grown in terms of both dimensionality and size. huge records about their customers can be utilized to make and keep clear relationship and association with the clients so as to target them individually for offers and

schemes. Usually, the selected customers are contacted directly through: personal contact, telephone cellular, mail, and email or any other contacts to advertise the new product/service or give an offer, this kind of marketing is called direct marketing, where a single communication message is broadcasted to all customers through media such as television, radio or advertising firm. Many of the customers are not interested or do not respond to this kind of sales promotion.

Banks have realized that customer relationships are a very important factor for their success. Customer relationship management is a strategy that can help them to build long-lasting relationships with their customers and increase their revenues and profits. The challenge the bank face is how to retain the most profitable customers and how to do that at

the lowest cost. We are proposing solution for this problem using data mining technique called Predictive Analytics.

Predictive Analytics is a field in social media analytics where we try to predict the action of a user based on their characteristics. It uses historical data in order to predict future actions of a user. Marketing firms and other advertisement agencies predict whether a customer will buy a product or not, depending upon the data that is available to them about customers. Social media data have always been the most sought after data set to find the behaviour of customer and do predictive analytics.

It includes variety of techniques from predictive modelling, machine learning and data mining that analyzes existing dataset to make future predictions. Depending upon the data that is available about customers, various advertisement agencies and marketing firms predict interest of the customers for a product. Machine learning allows banks to draw intelligent conclusions from their data that reduce risks, automate processes thus improving customer engagement. Various machine learning algorithms like Naive Bayes classification algorithm, K Nearest Neighbour Algorithm, K Means clustering algorithm etc. serve as the back bone of predictive analytics.

II. PROBLEM STATEMENT

To design and implement a web application to predict customers interested in bank products, schemes, etc. and provide recommendations, notifications to end-users.

III. LITERATURE SURVEY

Title	Overview
Bank Direct Marketing Analysis of Data Mining Techniques	This paper introduces analysis and applications of the most important techniques in data mining; multilayer perception neural network (MLPNN), tree augmented Naive Bayes (TAN) known as Bayesian networks, Nominal regression or logistic regression (LR), and Ross Quinlan new decision tree model (C5.0). The performances are calculated by three statistical measures; classification accuracy, sensitivity, and specificity.
Data Mining Techniques and its Applications in Banking Sector	Analysis of data mining techniques and its applications in banking sector like fraud prevention and detection, customer retention, marketing and risk management is performed.
Real time fraud detection in the banking sector using data mining techniques/algorithm	This paper hereby addresses bank fraud detection via the use of data-mining techniques; association, clustering, forecasting, and classification to analyze the customer data in order to identify the patterns that can lead to frauds.

Implementation of a Cloud in Banking Sector	Banks should adopt a progressing evolutionary approach towards cloud computing services, examining each project based on the type of applications and nature of the data. In the long term banks will have an application portfolio mix of on-premise and cloud-based services delivered across a combination of private, hybrid, and public cloud based deployment models with the share of cloud services gradually increasing in the service mix.
---	--

IV. PROPOSED SYSTEM

The proposed system is shown in the below diagram. Customer transactional related data is taken from database and data mining pre-processing techniques are applied to these data. After some statistical analyzing, a pre-defined hierarchical class is assigned on per account data. A Classification method is applied to this data. Classification method is developed to extract rules to predict interest of users.

There will be two types of users in our system- user and admin. User will be a bank customer or it may be a new user searching for bank products. When user will get authenticated, he will search for bank products and on the basis of search history, admin will create interested user's database by using data mining approach.

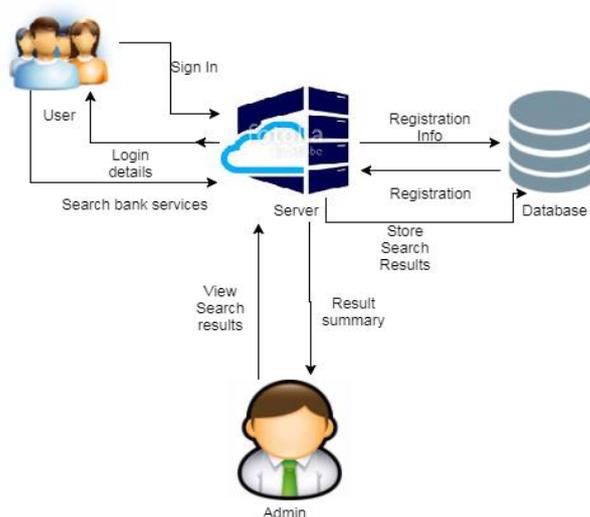


Fig.1 System Architecture

Proposed system mainly consist of two modules :-

Module 1: User

1. Registration and Validation

Users will register in bank system to get benefits from all products and scheme. We will store all personal, financial and corporate data of user. After registration a unique

username and password will be given to users. During registration, validity of data entered will be done.

2. Authentication and Login

In this module, user will get authenticated; this is very important step because only authenticated users will be allowed to use bank facilities. User will login by entering username and password, this will be checked with database values and thus user gets authenticated and can now login into the system.

3. View Services and Requests

Users can view detailed information about offers for various bank services. Eligible users will receive requests and notifications from admin about offers.

Module 2: Admin

Admin performs following functions:

1. View Users and Upload File

Admin can upload file containing user records. Attributes such as : name, email Id, password, age, income, marital status, employment type, resident type, vehicle type(2 wheel and 4 wheel). Admin can view users.

2. Check Eligibility

From the user database, eligibility check will be performed based on 3 parameters: Age, Employment Type and Income. Only these users will be eligible for bank loans (home, car, etc.).

3. Classification

Following parameters are used to calculate probability for home and car schemes respectively.

Home Loan	Car Loan
Age	Age
Income	Income
Marital Status	Employment Type
Employment Type	Vehicle (2 wheel/4 wheel)?
Resident Type	

Based on above calculated probability, users are prioritized into 3 classes :

- i. Class A - Will be frequently updated regarding loans/schemes (highest priority class)
- ii. Class B - Will be less frequently updated regarding loans/schemes (medium priority class)
- iii. Class C - Will be rarely updated regarding loans/schemes (lowest priority class)

4. Sorting

Based on user actions(i.e. history), sorting is performed. Here, count of requests sent and requests applied is recorded. Probability is calculated based on above constraints and users are further classified in classes: A,B and C.

5. Send notifications

After classifying users into classes, messages, requests and E-mails will be sent. Users can then apply for the same.

6. Analysis

Analysis is done based on number of users that have applied for a particular loan for each class. Probability is calculated and displayed in percentage for each class of each scheme.

V. CONCLUSION

The efficiency and performance of the banks are the main sources for a country's financial system. The most important aim of our project is to maintain dataset of existing customers and increase the number of customers for banking using predictive analytics. We also focus on increasing profitability, performance and efficiency of the banking sector.

REFERENCES

1. Ramesh, Ranjith. "Predictive analytics for banking user data using AWS Machine Learning cloud service." Computing and Communications Technologies (ICCCT), 2017 2nd International Conference on. IEEE, 2017.
2. Elsalamony, Hany A. "Bank direct marketing analysis of data mining techniques." International Journal of Computer Applications 85.7 (2014).
3. Chitra, K., and B. Subashini. "Data mining techniques and its applications in banking sector." International Journal of Emerging Technology and Advanced Engineering 3.8 (2013):219-226.
4. John, S. N., et al. "Realtime Fraud Detection in the Banking Sector Using Data Mining Techniques/Algorithm." Computational Science and Computational Intelligence (CSCI), 2016 International Conference on. IEEE, 2016.
5. Ghule, Sheel, R. Chikhale, and K. Parmar. "Cloud Computing in Banking Services." International Journal of Scientific and Research Publications 4.6 (2014): 1-8.