**Data Warehousing and Data Mining**

*Case Study (Summary)*

**Problem Statement**: To establish a proper decision support system to overcome the problems arise during loan operations like identifying the risk behind giving loan to the customer to minimize loan defaults.

**Data Collection:** Data of around 300 customers was taken under consideration from nationalized banks. Details like occupation, address, gross and net incomes, sanctioned loan amounts, outstanding amounts and months pending amount was taken under consideration. Once the data was collected, it was stored in MS ACCESS database and then a **data mining tool** called DBMiner was applied. This tool uses OADB connectivity to fetch data from MSACCESS database.

Loans are classified as under:

|  |  |
| --- | --- |
| **AssetType** | **MonthsOverDue** |
| Standard | 0 to 1 |
| OverDue | 2 to 6 |
| Substandard(NPA) | 7 to 24 |
| Doubtful | >=25 |

\*NPA = Customers with their loan considered as Non-Performance Assets from Risk Group.

**Analysis:** The whole data was mined using association rule and classification techniques. After applying DBMiner tool, association rules were generated and presented.

Following Rules were generated taking **Net Salary** and **Asset type** in consideration:

1. In 15% cases, if individual’s net salary is between 0-5000 that loan is considered as NPA.
2. In 6% cases, if an individual’s net salary is between 5-10000, that loan is considered as NPA.

In similar fashion, 8 rules were generated keeping various factors under consideration like (netsalary, servicetype, assettype etc)

**Conclusion:**

The whole case study is related to find out probability (rules) of the cases in which loan given to the customers is doubtful in recovery, and to sanction loan with minimum risk to the customers of the nationalized banks. Data was collected of 300 individuals and **DBMiner** tool was used to apply **association** **rules** using classification techniques which generate rules for further reference of the bank while sanctioning Loans to its customers. It was found that the rules which were generated were working on **80%** of the cases.