3rd International Conference on "Latest Innovations in Science, Engineering and Management"
The International Centre Goa, Panjim, Goa (India) (ICLISEM-17)
19th February 2017, www.conferenceworld.in ISBN: 978-93-86171-17-7

A REVIEW OF USING DATA MINING TECHNIQUES FOR EFFICIENT DECISION MAKING IN HEALTH SECTOR AND BANKING SECTOR

Jaspreet Singh Phul[1], Gaurav Verma[2], Sonia Vatta[3]

[1,2] Research Scholars, School of CSE, Bahra University, Waknaghat (India)

[3]Assistant Professor, School of CSE, Rayat Bahra University (India)

ABSTRACT

I*t is becoming common knowledge in health sector that prediction of services required or can be provided to patients can be an important attribute for development of existing health care system. Identification of these services truly depends on the historical data of such patients. Once identified, it can aid into accurate deployment of services. Further it can aid in reduction of death rate caused by delay in treatment. All this can be achieved by preprocessing huge data sets and getting meaningful information out of them. Such reservoir of information can act as a knowledgebase, in turn acting as bedrock for a decision support system, designed truly for the cause. In banking sector,* *It has evolved social knowledge in business where predicting of services required or can be provided to customers is an important strategy for business development. One in while identified, these customers can be captioned with proactive services campaigns. These proactive marketing campaigns usually include the initiating for new services to entice the customer into coming up with the services. Numerous researchers have presented algorithms for mining potential info system from multi-relational databases of various applications like, data financing and healthcare system. In financial data, account has both static characteristic in relation "account" and dynamic characteristic in relations "permanent order" and "transaction" has been given. Relation "client" describes characteristics of persons to compute with the accounts.*

Keywords: Health System,Banking Sector,Multi-Relational,Financial,Data Mining.

Tool used: SPSS Clementine

I DATA MINING TECHNIQUES

- **k-means clustering** : Clustering is one of the basic data processing. It is widely used in solving problems of customer segmentation and determination of swindling acts. In clustering applications we end up performing three tasks:

1. Separation of data sets into sections within the clusters,
2. 2. Verification of results of clustering,

3. 3. Interpreting the clusters.

Objective in models of clustering, bases on the fact that the elements of the clusters, resemble each other very much, but have characteristics that are present in clusters having a rather different aspects. Records present in database, are divided into this different clusters. In the K-means algorithm, K value can be determined according to problem or it can not be determined. Like squared error criterion, there is need to have a clustering criterion. The K-means algorithm starts with random selection of an object that will represent every cluster. Each of remaining objects is assigned to a cluster and the clustering criterion is used to compute average of the cluster. These averages are used as new cluster averages and each of the objects are assigned again to the cluster that resembles itself most. These clusters are computed and until no change is observed in the clusters and no change fall under the desired error level, this cycle is continued.

- **C4.5 and C5.0 Algorithms**: The most widely used decision tree algorithm is the C4.5 algorithm which is the develop state of ID3 algorithm that was proposed in 1986 by Quinlan. The C5.0 algorithm is the develop state of C4.5 and it is used especially for large data sets. To increase the correctness for the C5.0 algorithm, the boosting algorithm is used and therefore they are also known as boosting trees. The C5.0 algorithm is more rapid as compared to C4.5 and uses memory in a more productive manner [14]. Even to the results of both of the two algorithms are the same, the C5.0 as form makes it possible to come out with a smoother decision trees.
- **CART Algorithm**: It has the nature of being the continuation of the decision tree of Morgan and Sonquist titled AID (Automatic Interaction Detection) and was proposed by Breiman and others in 1984. CART algorithm accepts both numerical and the nominal data types as input and predicted variables; can be used as a solution in classification and regression problems. CART decision tree, has unique dual form divided into a structure. As branching criteria, CART tree benefits from Gini index, without any stopping rule at the phase of its structuring, it is continually divided and grows. In the state where a new branching will not be realized, a cutting out from top in the direction of root is started. The probable most successful decision tree is subjected to assessment with a test data independently selected after each cutting offs and efforts are made to make determinations [15].

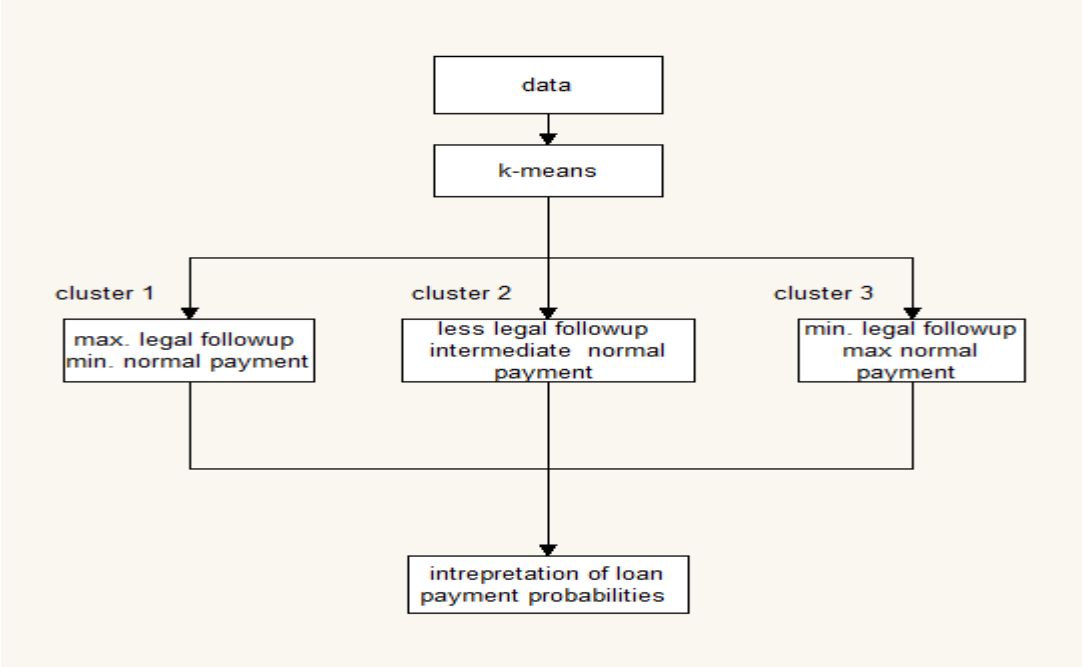
II METHODOLOGY

Data

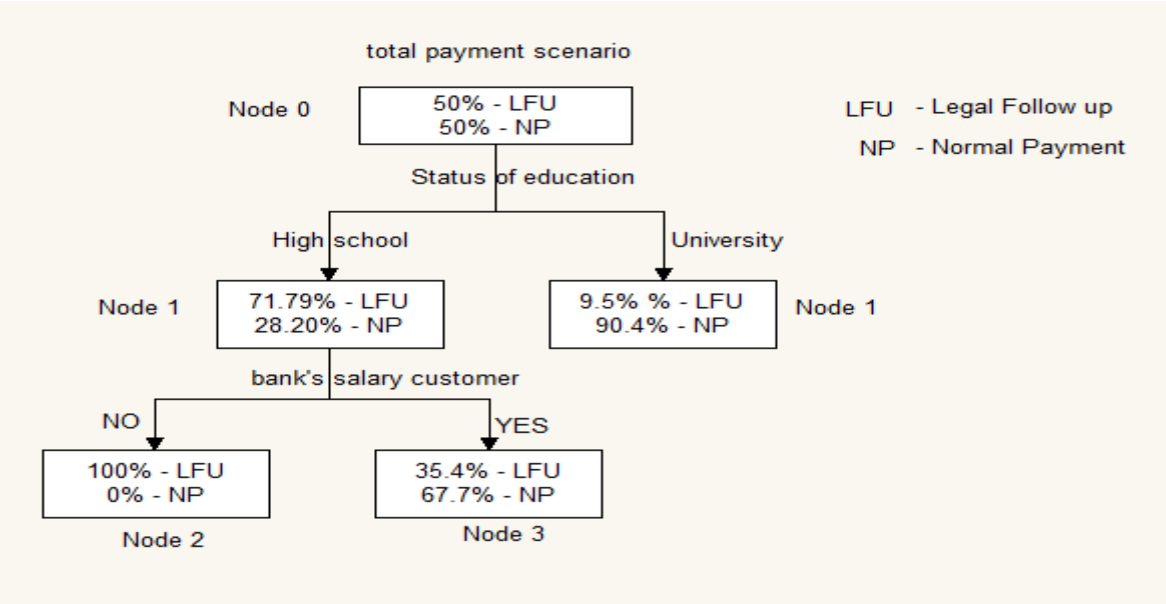
Within the scope of the study, data containing customer numbers and information about the status of credit paybacks belonging to the credit customers of the branch where the application is going to be made were secured from the operating within the structure of General Directorate. Information about gender, marital status, age, monthly income, income by spouse, status of education, owning house and car, having children, being a customer who receives his salary form the bank, the way of work were reached by using customer numbers and the existing system in the bank for making examinations.

K-Means Methodology

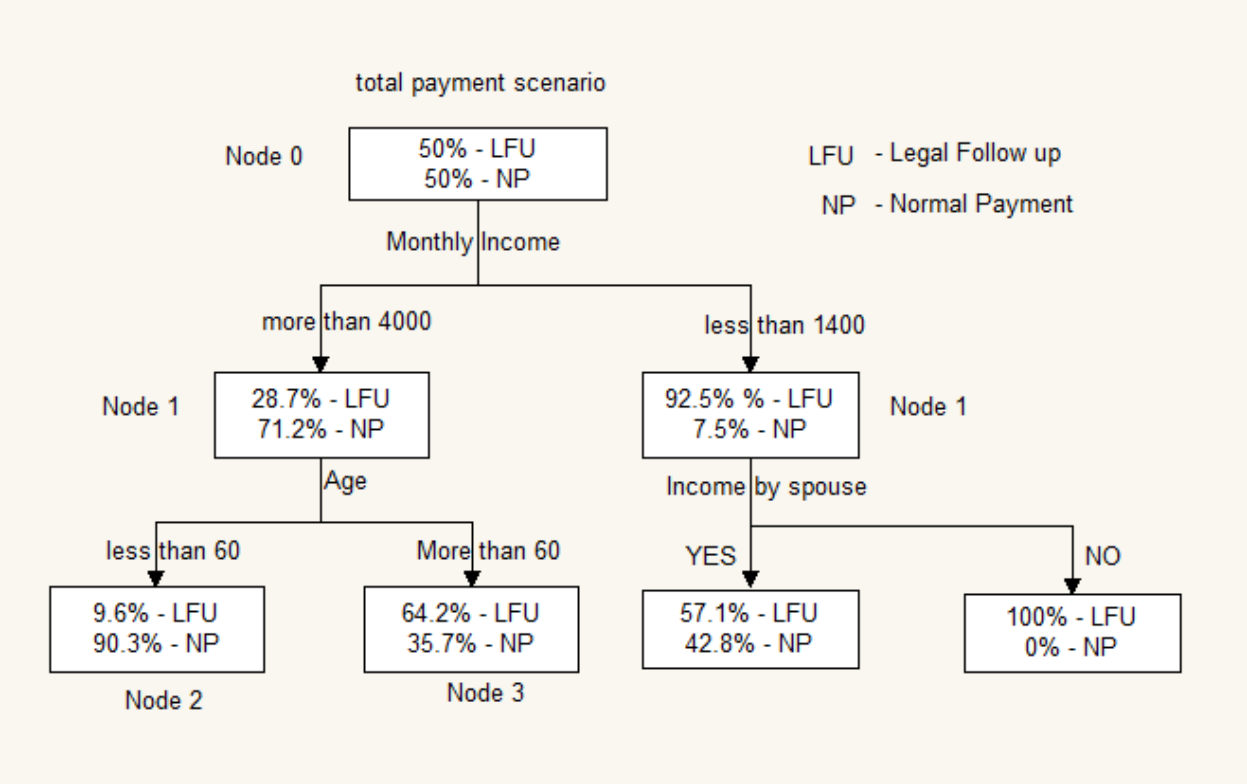
 ISBN: 978-93-86171-17-7



C5.0 Algorithm



C&RT Algorithm



III RESULTS OBTAINED

Techniques	C5.0	C&RT
Accuracy	88.75	86.25
Error	11.25	13.75

REFERENCES

[1] My Meta Process in banking, available online: http://www.mybank.in/askanything.php, last accessed on 27 July, 2015.

[2] Simple Data Collection, http://www.ihi.org/resources/Pages/Tools/SimpleData Collection Planning.aspx, last accessed on 28 December 2015.

[3] Systems of banking, https://www.bankingsystems.com, Availabe online, last accessed on 27 December 2015.

[4] What are the Advantages of Core Banking Processes, available online from: http://www.Banking.gov/providers-professionals/faqs.html, last accessed on 26 June, 2015.

[5] Suhardi, Rahardian Dwi Juliarto, "Validation of an IT Value Model for Branchless Banking," International Conference on Electrical Engineering and Informatics, August 10-11, 2015.

[6] Electronically controlled banking processes, available online from: http://www.professio nals.com /directory/banking, last accessed on 27 July, 2015.

[7] Paul Alpar & Moshe Kim, "A Microeconomic Approach to the Measurement of Information Technology Value," Journal of Management Information Systems, vol.14, no.12, 21 Dec 2015, pp.765-778.

[8] Fintellix.com, "Unified Metadata: Key to reliable, consistent information exchange between Banks and Regulators," White paper, 2015.

[9] J. Gopalkrishnana, Prof. B.B Mishra, "A study on the innovative business model at state bank of India: India's largest bank," Interdisciplinary Journal of Contemporary Research in Business, vol. 7, issue. 1, 2015, pp. 429-434.