International Journal of Modern Engineering Research (IJMER)
www.ijmer.com *Pp-15-18* *ISSN: 2249-6645*

Location Prediction System for Banking Loans Using Belief Network Approach

S. Rajaprakash[1], R. Ponnusamy[2]
[1] *Research scholar,* SCSVMV University, Vinayaka Mission University, Aaruparai veedu of technology
[2]*Matha Engineering College, Chennai Tamilnadu*

Abstract: *The banking Industry in India is undergoing a major transformation due to changes in economic conditions and continuous deregulation. This continuous deregulation has made the Banking market extremely competitive with greater autonomy, operational flexibility and decontrolled interest rate in loan system. The new generation private/foreign bank are planning to invest in expanding their business into loan promotion to achieve that the banks are in need of a finding solution which can get the micro detailed survey about the prospective loan location in any part of state which is considered to be challenging task. In order to overcome this problem an attempt is made to exactly find out the prospecturners in the target location with Belief Network.*

I. Introduction

This paper "An Intelligent Loan Prediction System with Bayes Approach" deals with the measures to obtain the details on feasibility of availing loan as well the micro details involved in a particular location in prospects of loaning with loan, which helps the bank to make their decision on of starting new branches in that location. The micro details about the people's income are collected and the information is classified according to the Business type and source of income, etc. The total data specification for this paper has been collected from the leading Multinational bank namely ICICI Bank Ltd. It is having many branches at district level and wishes to expand the business. This particular bank is chosen because they are pioneer in processing the loan faster and smoother. The relevant data is collected with the help of well framed questionnaire, recording the individual views of the people in the targeted location/analyzing demands and understanding people needs. All such data are categorized as nodes and these nodes are considered as Random variable with the help of Belief network. The Belief network helps to achieve this by its joint probability distribution in Directed Acyclic Graphs (DAG). The relevant data is statistically grouped as a network with belief network in parent and child combinations and the final result will be given in ratio basis of the prospective loan availing chances in a particular targeted location.

II. Bayes Inference

Bayesian network models support efficient reasoning under uncertainty in a given domain. Reasoning under uncertainty is the task of computing our updated belief in events based on given observations on other events.

2.1 Conditional probability

The basic concept in the Bayesian treatment of uncertainty is that of conditional probability of event a is x is written as P(a/b)=x (i.e) if event b is true and everything else known is irrelevant for event 'a' then the probability of event a is x.

The following axiom gives the basis for Bayesian probability
1. For any event a, $0 \le p(a) \le 1$ $p(a)=1$ if and only if a occurs with certainty.
2. For mutually exclusive events a & b the probability that either 'a' or 'b' occurs is P (a or b)=P(a)+P(b).
3. Fundamental Rule of probability calculus
For any two event 'a' and 'b' the both a&b.
(i.e)P(a,b)=P(a/b)P(b)=P(b/a)P(a). P(a,b) is called the joint probability of the event a & b.

2.2 Bayes Rule

Generalizing the above rule 3 to random variables X and Y we get the fundamental rule of probability
P(X,Y)=P(X/Y)P(Y)=P(Y/X)P(X). (1)

$$P\left(\frac{Y}{X}\right) = \left(\frac{P\left(\frac{X}{Y}\right)P(y)}{P(X)}\right) \quad (2)$$

$$P(Y/X) = \frac{P\left(\frac{X}{Y}\right)P(Y)}{P\left(\frac{X}{Y=y1}\right)P(Y=y1)+\cdots+P\left(\frac{X}{Y=yn}\right)P(Y=y_{||Y||})} \quad (3)$$

That is , the denominator can be derived from the numerator in (2) Furthermore, the denominator is obviously the same for all states of Y.

2.3 Chain Rule

For a probability distribution, P(X), over a set of variables $X=\{x_1 \ldots, x_n\}$, we can use the fundamental rule repetitively to decompose it into a product of conditional probabilities

$$P(X) = P(x_1 \mid x_2,...x_n)P(x_2,...x_n)$$
$$= P(x_1 \mid x_2,...x_n)P(x_2 \mid x_3...,x_n)...P(X_{n-1} \mid p_n)P(x_n)$$
$$= \prod_{i=1}^{n} P(x_i \mid x_{i+1},...x_n)$$

2.4 Reasoning under Uncertainty

A probabilistic interaction model between a set of random variables may be represented as a joint probability distribution. Considering the case where random variables are discrete, it is obvious that the size of the joint probability distribution will grow exponentially with the number of variables as the joint distribution must contain one probability for each configuration of the random variables. Therefore, we need a more compact representation for reasoning about the state of large and complex systems involving a large number of variables. To facilitate an efficient representation of a large and complex domain with many random variables, the framework of Bayesian networks uses a graphical representation to encode dependence and independence of relations among the random variables. The dependence and independence relations induce a compact representation of the joint probability distribution. By representing the dependence and independence relations of a domain explicitly in a graph, a compact representation of the dependence and independence relations is obtained.

2.4 Construction of Belief Network

- Choose the set of relevant variables that describe the domain
- Choose an ordering for the variables
- Choose a variable x and add a node for it
- Set parents to some minimal set of existing nodes such that the conditional independence property is satisfied
- Define the conditional probability table for the network
- This network should be DAG.

III. Current System Design

In this paper the loan prediction has been made with Belief network by considering nodes as random variable. The variables are considered to be random. They are chosen with the feasibility of location prediction which has higher chance in making the location prediction process as a successful one. The variable are connected in the form of DAG, with each node having a conditional probability table (CPT) that quantifies the effects that the parents have on the node. The parents of a node X are all those nodes that have arrows pointing to X. The same above process is modified and adapted as described below.

In this paper the attributes are classified into two major categories
1. Major nodes
2. Vital node

The Major nodes:- 'Agriculture','Business','salaried' and 'others' are considered as major attributes (hence termed as major nodes) which helps to form a relevant parent and child node with help of Belief network.

The Vital node: - The 'L' node is considered as vital node because it use to find the location prediction of uncertainty involved in loan prediction process with belief network. The loan prediction belief network explained above in targeted location is illustrated in Figure 1.

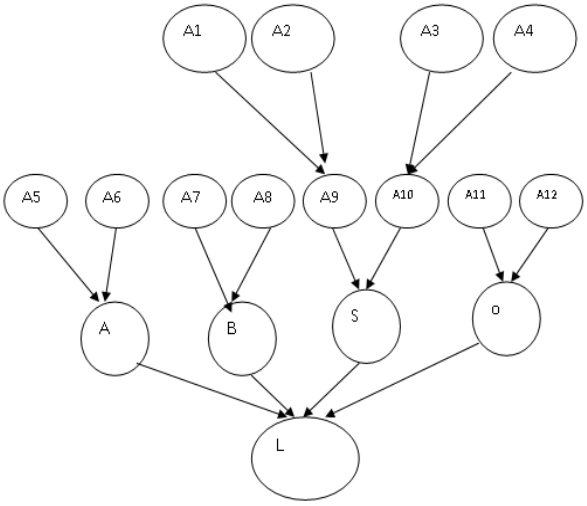

Figure 1

3.1 Data Source

The Data specification for this paper is collected from the Multi National Bank namely ICICI Bank Ltd Kanchipuram, Tamilnadu. The Data's were actually collected by the bank in order to expand and promote their business into loan section. From the data collected with the help of the bank, the data is grouped in two main categories Loan availing and Loan non availing

3.2 Data Process

The process adopted to collect the Data is Targeted location survey

- Questionnaire to people
- Recording individual views
- Analysis the people needs
- Organizing and recording group interviews of individual business, agriculture, salaried and other type of people involved
- Finding the interest level of people in availing the loan facility

3.3 Data Analysis

The Data gathered in the above manner is subjected to research with the help of this paper "An Intelligent Loan Predication with Bayes Approach" that is Belief network.

3.4 Research Methodology

The real time Data's are further analyzed and statistically grouped with their relevant matching criteria and formed in the manner of a network called belief network.

3.5 Formation of Belief network

The Area of business or source of Income is classified as Agriculture, Business, Salaried, Others. The ultimate reason for above collection is to find the loan availing ratio which is also considered in this paper.

The Belief network is formed by assuming the following factors as Nodes
Node Representation
A: Agriculture B: Business
S: Salaried O: Others A1: Professional A2:Others
A3:Professional A4:Others A5:Seasional crop A6:Non seasonal crop A7:Small scale A8: Large scale
A9:Government A10:Private A11:NGO(non Government Organization)
A12: Charitable Trust L: Loan

Each node is first classified with it relevant matching criteria and forming the Conditional probability table and calculated the value.
Here $P(A_1)=0.6$ (availing loan as professional) $P(A_2)=0.5$(availing loan as other category)

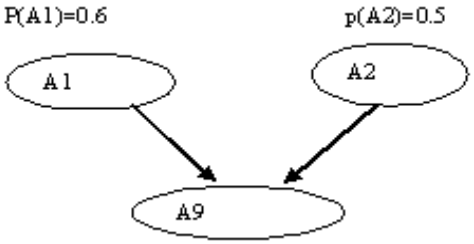

Figure 2

Table 1 Conditional Probability Table

A_1	A_2	$P(A_9/A_1A_2)$
T	T	0.7
T	F	0.5
F	T	0.4
F	F	0.002

$$P(A_9) = P(A_9\, A_1\, A_2) + p(A_9\, A_1\, A_2^1) + p(A_9\, A_1^1\, A_2) + p(A_9\, A_1^1\, A_2^1)$$
$$= p(A_9 / A_1\, A_2)P(A_1\, A_2) + p(A_9 / A_1\, A_2^1)P(A_1\, A_2^1) + p(A_9 / A_1^1\, A_2)P(A_1^1\, A_2)$$
$$+ p(A_9 / A_1^1\, A_2^1)P(A_1^1\, A_2^1)$$
$$= 0.21 + 0.15 + 0.08 + 0.0004 = 0.4404$$

Here the calculated value of $P(A_9)$ is 0.4404 from the collected data .Similarly we can find the following values

$P(A_1) = 0.6$ $P(A_2) = 0.5$ P(A3) = 0.5 $P(A_4) = 0.35$
$P(A_5) = 0.7$ $P(A_6) = 0.8$ $P(A_7) = 0.7$ $P(A_8) = 0.8$ $P(A_9) = 0.44$ $P(A_{10}) = 0.343$ $P(A_{11}) = 0.4$ P(A12)=0.45 P(A) =0.5592 P(B) =0.448
P(S) = 0.180 P(O) = 0.277
P(L) = 0.312

IV. Simulation Result

The vital node L is derived with the valu 31%with the help of belief network. This percentage decides on the expansion of the bank. The Loan availing chances of the target location for example in this paper is P(L)=0.31 which means this that location has lower prospectus and chances of loan availing , hence the bank need not venture in the business into this area. If the percentage is lesser than the 50%then it is not advisable for the bank to venture in new business into this Area.

V. Conclusion

In paper, an attempt is made to help the bank to find out the venturing and expansion of their new business into a targeted location with the help of statistical information by the belief network. The information given in this paper is only a simple sample, when the belief network is fully expanded and supported with huge mass of data then the predicted result will be accurate. The future extension of this paper will be on finding the closely netted matching data criteria and finding the ratio of not only the interest level in availing the loan, but also the repaying capacity of the loan of an individual will also be expected to be predicted. More than the targeted location the results can be used to find the Targeted Zone as well as the success ratio of it. This work not only considers macro or even the micro view of a single person loan availing chances but also help us to get a better results. The belief network can be further tested with complex data to create a huge and larger vision and to bring out the results in any sector of the business for finding the success ratio of targeted projects. This paper helps the bank to find out the venturing and expansion of their new business into a targeted location with the help of statistical information by using the belief network.

References

[1] Astrid A. Dick , "Demand Estimation and Consumer Welfare in the Banking Industry", Journal of Banking & Finance 32 (2008), pp. 1661–1676.
[2] Richard S. Barr, Lawrence M. Seiford, Thomas F. Siems, "Forecasting Bank Failure: A Non-Parametric Frontier Estimation Approach", Recherches Economiques de Louvain, Vol. 60, No. 4. (1994), pp. 417-429.
[3] Uffe B.Kiaerulff ,Anders L.Madsen "Probabilistic Network- An Introduction to Bayesian Networks and Influence Diagrams" may-2005-Aalborg University
[4] G.H.Bakir, T.Hofmann,B.scholkopf, A.J.Smola, and B.Taskar. Predicting structured Data. Cambridge: MIT press,2007
[5] Jeasen,F.V.(1996). An introduction to Bayesian Networks, UCL Press , London
[6] L.C.van der Gaag and P.R.de Wall. Multi-dimensional Bayseian network classifiers. In third eruopean conference on Probabilistic Graphical Models. Pages 107-114,2006
[7] W.Cheng and E.Hiilermerier. Combining instance-based lerarning and logistic regression for multilabel classification. Machine Learning, 221-225,2009
[8] Laurizen, S.L.& Jensen, F.(2001). Stable Local with Mixed Gaussian Distributions, statistics and Computing 11(2): 191-203.
[9] J.pearl. Probabilistic Reasoning in Intelligent system: Networks of Plausible Inference. Morgan Kaufmann Publishers, 1988.
[10] P.Lacas. Bayesian model-based diagnosis. International journal of Approximate Reasoning, 99-119,2001