JCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-3, Special Issue-1 E-ISSN: 2347-2693

Loan Customer Analysis System using Column-wise Segmentation of Behavioural Matrix (CSBM)

Rakesh Kumar Mandal*
Department of Computer Science & Application
University of North Bengal, India
rakesh_it2002@yahoo.com

N R Manna
Department of Computer Science & Application
University of North Bengal, India
nrmanna12@gmail.com

***Abstract*—In order to approve bank loans, modern day researchers and bankers are involved in different types of work related to analysis of the behaviour of the loan applicants. Customer data are collected to analyze their behaviour which may predict the possibility of repayment of the EMIs. In this paper, a new approach has been made to make this process fast. The data used to analyze customer behaviour are actually behavioural patterns. Artificial Neural Networks (ANNs) are very good tool to train a system for known patterns and which can later be used to identify unknown patterns. Two dimensional binary pattern matrixes are formed considering different behaviour of customers from different views. Matrixes are further segmented column-wise and each column is presented to Perceptron (ANN) in order to train the ANN for the known patterns. Later on, unknown patterns of customer behaviour can be presented to the net to reach to the decision to provide loan to a customer or not.**

Keywords- Loan, Customer, ANN, Column-wise Segmentation, Perceptron, Behavioural Pattern

I. INTRODUCTION

Banks provide loans to its customers and need proper recovery of the loan in time. A few customers take the opportunity of not paying back the EMI intentionally after taking the loans, or fail to repay due to some financial constraints. It becomes very significant for the financers to analyze the behaviour of the customers before providing the loans. Artificial Neural Networks (ANN) can be used to develop models to predict the intention or inability of the customer to payback. These ANNs are actually initially trained to learn the behavioral patterns of different types of customers and help the bankers to take decisions to approve the loans. The bankers, in order to analyze the customer behaviour, use various approaches for collecting customer data, preparing questionnaires and survey, study patterns etc. The approach has been made here to study the behavioral patterns of the loan applicants.

The overall program is divided into four parts: forming the Customer Behavioral Matrix (CBM), segmenting the behavioral matrix, forming column vectors and training the net. Binary vectors are generated by studying the customer behaviour. These vectors can be used to train an Artificial Neural Network (ANN) designed for the purpose. ANN-based techniques are found very helpful to analyze credit risk in financial and banking industry [6]. The trained ANN can later verify an unknown customer's behaviour from his individual behavioural matrix pattern. Decision of loan disbursement may be made accordingly.

A segmentation approach to recognize handwritten patterns have already been made through Column-Wise Segmentation Technique (CSIM) [1]. Perceptron learning method has been used to train different ANNs [2, 3, 4,

and 5]. Various applications of Information Technology are discussed to mitigate the problems of loan fraud in Nigeria [7].

In this paper, an attempt has been made using Column-Wise Segmentation Technique (CSIM) to train the ANN for some known customer behaviour and later take decisions for unknown customer behaviour. Finally, testing has been done by providing unknown behavioral patterns taken from different individuals.

II. METHODOLOGY

The overall methodology adopted in this work is divided into mainly four phases. Phase I discusses on the preparation of input matrix taking into consideration various types of behaviour of the loan applicant by collecting some related customer data. Phase II discusses on the designing of an ANN. Phase III discusses on the training of the ANN using patterns those are considered favorable to approve the loan. Finally, Phase IV is related to testing the ANN, considering some matrixes obtained from different customer's behaviour.

A. Preparing the Input

Inputs of ANN are actually two dimensional binary pattern matrices named as CBM, which are prepared by collecting some customer data mainly through questionnaires. As an example, the single input matrix (CBM) has been prepared through the following set of questions. The questions may be prepared taking into consideration the previous experiences of the Bank authority and the current market expenses.

Example:

- *Customer Information Regarding Income*

QR1 (a) Is the applicant salaried?
QR1 (b) Is the applicant in business?
QR1 (c) Is the applicant's income ≥ Rs. 40000/-?
QR1(d) Is the applicant's age>45?

- *Customer Information Regarding Family Background*

QR2 (a) Is the applicant's 50% or more family members are earning?
QR2 (b) Is the applicant's 50% or more family members are medically fit?
QR2 (c) Is the applicant's 100% family members are adults?
QR2 (d) Is the applicant's 75% or more family members are literate?

- *Customer Information Regarding Assets*

QR3 (a) Is the applicant having any fixed deposits?
QR3 (b) Is the applicant having any insurance policy?
QR3 (c) Is the applicant having any property value>Rs. 1000000/-?
QR3 (d) Is the applicant having any other property value>Rs. 500000/-?

- *Customer Information Regarding Credit Cards*

QR4 (a) Is the applicant having any credit card having credit limit > Rs.50000/-?
QR4 (b) Is the applicant paying the credit card bill regularly for last one year?
QR4 (c) Is the applicant last six months purchase is more than Rs.50000/-?
QR4 (d) Is the applicant paid any penalty for anytime for not paying of bills?

Taking into consideration the above questionnaires a two dimensional behavioral matrix (CBM) is prepared where each group of questions is considered as a row of the matrix. The answers are considered as binary inputs where 'Yes' is considered as binary '1' and 'No' is considered as binary '0'.

The above question set may generate the following matrix depending on the answers yes / no. of each question:

1	0	1	0
0	1	1	1
1	1	0	1
…	…	…	…
…	…	…	…

The answers of the training question set are decided randomly in such a manner that generally, a customer having this type of characteristics pattern is supposed to pay back the loan in time and this behaviour is considered as favourable pattern to approve the loan.

B. Architecture of CSBM-net

An ANN has been designed which can be trained using the matrix (CBM) and tested for matrices of same dimension which are generated by the behavioral patterns of the loan applicants. Figure 1 displays the ANN consisting of two layers of neurons and named as CSBM-net. The first layer is the input neuron layer and the second layer is the output neuron layer. Input neuron layer consists of sixteen neurons and is divided into four segments, where each segment is a column of the matrix (CBM). There are four neurons in the output layer. Each column segment is connected to one individual neuron in the output layer forming four perceptrons [3]. There is one weight layer which is sandwiched between two neuron layers. There are sixteen weight elements in the weight vector.

C. Training the ANN

All the elements in the weight vector are initialized to zero. The training starts by presenting the vector generated from CBM. Before presentation, some preprocessing has been done to the input matrix by converting it into a bivalent matrix for better learning [3]. A bivalent matrix consists of only 1s and -1s, where the 0s of the binary matrix are replaced by -1s. Each column of the CBM is presented to corresponding segment of the ANN for training.

Training is done by applying perceptron learning rule which is given as follows:

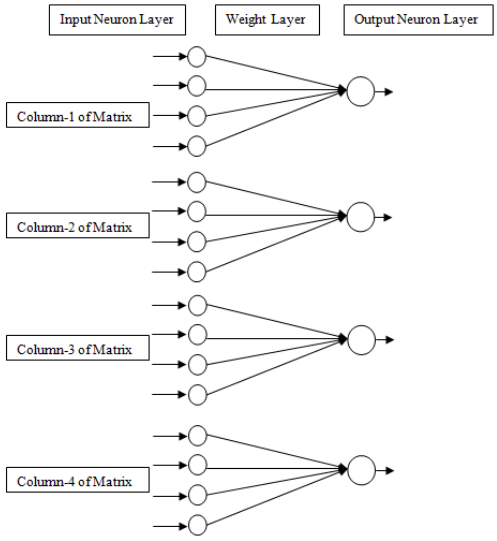

Figure 1: Architecture of CSBM-net

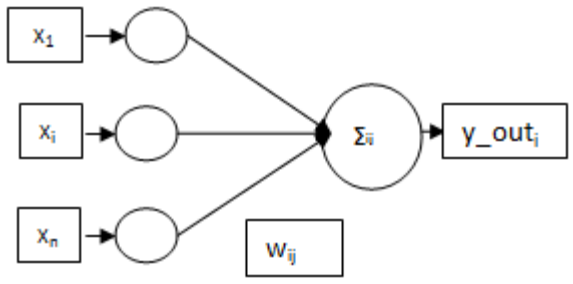

Figure 2: A basic Perceptron

Perceptron Learning Rule

Figure 2 displays the architecture of a perceptron. Where, 'y_j' is the output of the j^{th} perceptron which is calculated by applying the 'f' called the activation, which is a function of the net output, y_out. So,

$$y_j = f(y\_out_j) \quad \text{and}$$
$$y\_out_j = \sum_i x_i w_{ij} \qquad \text{Eq. (1)}$$

where, x_i is the i^{th} element of input vector and w_{ij} is the weight between the i^{th} element of the input vector and j^{th} element of the output vector. The function $f(y\_out_j)$ takes the following values depending on the values of y_out_j.

$$f(y\_out_j) = \begin{cases} 1 \text{ if } y\_out_j > \theta \\ 0 \text{ if } -\theta <= y\_out_j <= \theta \\ -1 \text{ if } y\_out_j < -\theta \end{cases}$$

Here, θ is the threshold value taken at random. For each training input, the ANN calculates the response of the output unit. The ANN determines whether an error occurred for this pattern by comparing the calculated output with the target value. If an error occurs for a particular training input pattern, the weights are changed according to the formula:

$$w_{ij}(new) = w_{ij}(old) + \alpha t_j x_i \qquad \text{Eq. (2)}$$

where, α is the learning rate, the value of which is taken at random, t_j is the target value, which is the output expected from the ANN. The output y_j produced by the ANN is compared with t_j and the difference leads to the modification in weight given by equation (2). The process continues until y_j becomes equal to t_j. Weights obtained at that point are the final and standard weights. The target values of the output neuron layer are set to all 1s. The ANN is trained until all the neurons produce 1s. The standard weights are considered for testing.

D. Testing the ANN

The ANN is tested by presenting some CBM matrices generated through questioning some loan applicants. While testing, it is decided that if three or more output neurons produce 1s, then that vector is considered successful and the applicant responsible for creating the CBM matrix is considered for loan approval.

III. RESULT ANALYSIS

The ANN model has been trained and tested for demo data and found successful. Table 1 displays the results produced by presenting ten sets of questionnaires to the ANN.

Table 1: Result Analysis of CSBM-net

Question	Col-1	Col-2	Col-3	Col-4	Approval
Set-1	1	0	1	1	Yes
Set-2	1	1	0	1	Yes
Set-3	1	1	1	1	Yes
Set-4	1	0	0	0	No
Set-5	1	1	0	1	Yes
Set-6	1	1	1	1	Yes
Set-7	1	0	0	1	No
Set-8	0	0	0	0	No
Set-9	1	0	1	1	Yes
Set-10	1	1	1	0	Yes

IV. DISCUSSION

The method developed and discussed here has been compared with the already available ones in this domain and it was found that most of the AI methods are rule based and lengthy. For small databases, rule based approach is

comfortable but for a vast range of loan applicants, rule based systems become cumbersome. ANN models can be trained with few matrices but can be tested for huge amount of data.

V. CONCLUSION

This method is a unique approach of using ANN for loan approval. In future, this approach can be used online. Real time data can stop loan frauds in financial institutions and will be a very helpful tool for the bankers.

Here, a sample question set has been prepared in which each set consists of four questions. Actually, each question set may have as much number of questions as required for the problem. In such cases, the CBM matrix can be prepared by padding up the unfilled matrix elements by a third value such as '-1'. It will be better to assign the answer of the questions as 1, -1 and 0 for yes, no and the padding elements, respectively, as there is no contribution due to padding elements. Of course, equation 1 may have to be modified suitably for fixing new weights of the net.

REFERENCES

[1] Rakesh Kumar Mandal and N R Manna, "Hand Written English Character Recognition using Column-wise Segmentation of Image Matrix (CSIM)", WSEAS Transactions on Computers, Volume 11, Issue 5, May 2012.

[2] G.N. Swamy, G. Vijay Kumar, "Neural Networks", Scitech, India, 2007.

[3] L. Fausett, "Fundamentals of Neural Networks, Architectures, Algorithms and Applications", Pearson Education, India, 2009.

[4] Apash Roy and N R Manna,"Character Recognition using Competitive Neural Network with Multi-scale training", UGC Sponsor National Symposium on Emerging Trends In Computer Science (ETCS 2012) on 20-21 January 2012, pp 17-20.

[5] Apash Roy and N R Manna, "Competitive Neural Network as applied for Character Recognition " - International Journal of advanced research in Computer science and Software Engineering, Volume 2, Issue 3, 2012, pp 06-10.

[6] V Moonasar, Credit Risk Analysis using Artificial Intelligence: Evidence from a Leading South African Banking Institution. Available: www.academia.edu/502093/credit_risk_analysis_using_artificial_intelligence_evidence_from_a_leading_South_African_banking_institution.

[7] Ifeyinwa Ajah, Chibueze Inyiama, Loan Fraud Detection And IT-Based Combat Strategies, Journal of Internet Banking and Commerce, 2011, Vol. 16 No. 2, pp 1-13. Avialable at: www.arraydev.com/commerce/JIBC/2011_08/Ajah.pdf