Estimation of Credit Risk for Business Firms of Nationalized Bank by Neural Network Approach

Ms. A. R. Ghatge 1 and, Mr. P. P. Halkarnikar 2
1Department of Technology, Shivaji University, Kolhapur
E-mail: amruga12@gmail.com
2D.Y.Patil College of Engineering and Technology, Kolhapur
E-mail: pp_halkarnikar@rediffmail.com

Abstract—Financial credit risk assessment has gained a great deal of attention. Many different parties have an interest in credit risk assessment. Banking authorities are interested because it helps them to determine the overall strength of the banking system and its ability to handle adverse conditions. Due to the importance of credit risk analysis, many methods were widely applied to credit risk measurement tasks, from that Artificial Neural Network plays an important role for analyzing the credit default problem. Artificial neural networks represent an easily customizable tool for modeling learning behavior of agents and for studying a lot of problems very difficult to analyze with standard economic models. ANN has many advantages over conventional methods of analysis. According to Shachmurove (2002), they have the ability to analyze complex patterns quickly and with a high degree of accuracy. The focus of this paper is to determine that a neural network is a suitable modelling technique for predicting the business firm loan is satisfactory or not. This paper shows that an ANN approach will classify the applicant as a default or not and predict a credit default allowance amount more closely aligned with the credit default expense incurred during the fiscal period than traditional management approaches to estimating the allowance. The results show that credit risk evaluation using Back propagation neural network and expert evaluation have the very good consistency.

Keywords—Financial Analysis, credit risk, NN algorithm, ROC Analysis

I. INTRODUCTION

Banking industry deals with capital and risk for making profit. The bank success is directly pertaining to its capability of controlling and managing related risks. Banks are exposed to different kinds of risk, but the most challenging risk which can cause a bank to failure is credit risk. The performance of loan contracts affects profitability and stability of a bank growth and development. The extent to which a borrower uses the credit facility efficiently, greatly impacts the firm’s repayment ability and performance, which by large, affects the lending institutions. Credit risk is the loss of bank’s profit, since the customer does not adhere to his or her loan refund commitment. Banks should control credit management thoroughly. Financial institutions are facing the problem of credit proposals because of continuous changes in the business environment, credit regulations, marketing strategies and the competition amongst them. The objective of credit scoring is to help credit providers quantify and manage the financial risk involved in providing credit so that they can make better lending decisions quickly and more objectively. It is very hard to design or even to use a banking loan model that is capable of tracing all possible and real responses and interactions. Before sanctioning /granting loans bank have to take various precautions such as performance of the firm by analyzing last year’s financial statements and history of the customer.
Various statistical and machine learning techniques have been used to model company credit and bond ratings in the past. Various work had been carried out and it shows that the machine learning techniques is having better performance than statistical techniques. Artificial neural networks have been successfully used in a variety of business fields including marketing, accounting, management information systems and production management. Most of the studies have used neural networks for predicting future default in the banks by selecting good accounts. System also controls bankruptcy, frauds in the credit area. The researchers have used neural networks in modeling market response collective behavior, telecommunication and real estate evaluations and decision making. Artificial neural networks have successfully provided effective credit evaluations for supporting granting loans.

The objective of this study is to determine that the utilization of ANNs is a valid and useful mechanism for analyzing credit risk and estimating the allowance for credit default. The applicant has been divided into two categories: good credit and bad credit. A good credit customer is likely to repay the debt whereas a bad credit customer is likely to default. An analysis of credit default can provide, indirectly, an indication of whether or not a company’s credit granting policies are proper.

**II. Literature Overview**

With the rapid growth of credit industry in last few decades, credit evaluation of loan applicants becomes an important issue not only because of the urgent demand of bankers, but also due to cash flow and collections. Credit risk is the oldest form of risk in the financial markets and the most important risks for banks in terms of the size of potential failures. The main objective of credit scoring is to decide whether or not to grant credit to an applicant. Most of the credit scoring models were established using either statistical methods or artificial neural networks. There is a wide range of techniques used to solve credit risk problem.

The paper of Roli Pradhan, K.K. Pathak, V.P. Singh [1] proposed that the back propagation neural networks algorithm has been used to forecast the Z Score for the firms using back propagation algorithm.

The paper of Donald Thomas Joyner [2] proposed the findings indicate that neural networks over the balance of the time are better predictors of a company’s ending allowance for bad debt than regression.

The paper of Simon James Frank [3] concluded that back propagation artificial neural networks, using publicly available rating and financial data, that are suitable tools for the prediction of future company credit ratings (or changes in ratings) for manufacturing companies in the United States of America. They also proposed that the use of public information also illustrates that it is a relatively cost effective alternative to relying on expensive proprietary information and expert knowledge.

The paper of Qeethara Kadhim al-shayea Ghaleb awad el-refae [5] suggested the study which predicts the bank insolvency before the bankruptcy using artificial neural networks, to enable all parties to take remedial action. They also make use of backpropogation algorithm and what are the results after applying Feed-forward back propagation neural network methodology to predict financial distress based upon selected financial ratios show abilities of the network to learn the patterns corresponding to financial distress of the bank. The percent correctly classified in the training sample by the feed-forward back propagation network is approximately 91 percent.

The paper of Bakpo, F. S. and Kabari, L. G [6] discussed with an architecture that uses the theory of artificial neural networks and business rules to correctly determine whether a customer is good or bad. They proposed an approach that allows for using different rules within the same data set, and for defining more accurately clients with high risk.
III. Model Design and Methodology

In this paper, a two step methodology has been adopted. The part A provides the steps formulated for the prediction of financial ratio, followed by part B which lists the steps followed for the prediction of financial ratios using artificial neural networks.

1) Part A: Formulation of Ratios

The basic ratios are formulated from details mentioned in published statements like balance sheet, cash flow statements, yearly details of banks, profit and loss statements obtained from nationalized bank. The table shows related ratios:

<table>
<thead>
<tr>
<th>RISK FACTORS</th>
<th>VARIABLES</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leverage and Solvency Indicators</td>
<td>Capital</td>
<td>Net means of the firm/borrower i.e. his own funds</td>
</tr>
<tr>
<td></td>
<td>Net Worth</td>
<td>Capital + reserves</td>
</tr>
<tr>
<td></td>
<td>Debt Equity Ratio</td>
<td>It is a proportion between firms total debt and total equity</td>
</tr>
<tr>
<td>Liquidity Indicators</td>
<td>Current Liabilities</td>
<td>Creditors, Loans to be repaid within one year, provisions of taxes and expenses</td>
</tr>
<tr>
<td></td>
<td>Current Assets</td>
<td>Cash and bank balances, Inventory.</td>
</tr>
<tr>
<td></td>
<td>Current Ratio</td>
<td>It measures the proportion of a Party’s current assets to its current liabilities and thus gives a measure of the short term liquidity.</td>
</tr>
</tbody>
</table>

Table.1 The Supportive Ratios

2) Part B: Back Propagation Neural Network Model Application for Nationalized Bank

The financial data required to create the neural network model is taken from the Nationalized Bank. The entire sample consists of business firm’s loan proposals. The data is considered from the year 2004 to 2007 in the form of excel sheet. The neural network architecture works in this way.

(i) In the input layer, balance sheet data are inserted;
(ii) In the hidden layer, it calculates the activation state of each neuron;
(iii) The output layer expresses a result which is easy to interpret.

The decision systems make use of neural network algorithm. Feed-forward Back propagation neural networks are widely and successfully used models for prediction of default or non default of customers. A feed-forward fully connected network is trained in supervised manner. In supervised learning, the network is trained by providing it with input and output patterns. During this stage, the neural network is able to adjust the connection weights to match its output with the actual output in an iterative process until a desirable result is reached. The result is classified as default or non default customers.
Estimation of Credit Risk for Business Firms of Nationalized Bank by Neural Network Approach

IV. Analysis of Results and Outcomes

After the computation of the basic ratio, as suggested by Table 1, this section uses the ratios as inputs to train the network. The network after training computes the scores of the ratios from 2004 up to the year 2006 at different tolerance level. The tolerance level that provides the closest values is considered for prediction. A back propagation neural network is used for prediction.

In order to test the performance of designed model in this paper, the real world dataset from nationalized bank is used which is presented as follow.

A. Real World Dataset

The dataset from nationalized bank is considered as a input to the system. It contains 23 instances, with 16 cases of whom the loan proposals are sanctioned and 7 cases of whom the loan proposals are rejected. In these instances, each case is characterized by 13 decision attributes, 10 numerical and 3 categorical. From the given dataset 14 inputs are used as a training samples and remaining 9 as a testing samples.

B. Experimental Results

The evaluation criteria used to compare the 9 testing dataset is done by classification of Neural Network with the classification by manual calculations. The table 2 represents the comparative study of the system with traditional method.

<table>
<thead>
<tr>
<th>Classification by Manual Calculation</th>
<th>Classification by NN Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy Companies</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>30.00%</td>
</tr>
<tr>
<td>5</td>
<td>50.00%</td>
</tr>
<tr>
<td>Risky Companies</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>60.00%</td>
</tr>
<tr>
<td>4</td>
<td>40.00%</td>
</tr>
<tr>
<td>Total Sample</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>90%</td>
</tr>
<tr>
<td>9</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 2. Results Obtained by Manual Calculation of Bank and by Our Method

The neural network show more accurate results as compared to results by the manual calculation of bank. The comparative result is represented in fig 1.

Fig 1. Representation of comparative study of Manual Calculation Vs. Neural Network Calculation

The receiver operating characteristic (ROC) curve is also used for calculation for each method. The ROC method is the method of specifying the classifier performance to the correctly classified positive cases known as True positive rate and correctly classified negative cases known as false positive rate.

This confusion matrices is designed for 8 training samples

Table 3. Confusion matrices for 8 training samples
The Confusion matrices for 12 training samples

<table>
<thead>
<tr>
<th>True Positive</th>
<th>False Positive</th>
<th>False Negative</th>
<th>True Negative</th>
<th>TPRate</th>
<th>FPRate</th>
<th>No. of Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>0.2</td>
<td>0.5</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>0.66</td>
<td>0.66</td>
<td>36</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>0.6</td>
<td>0.25</td>
<td>48</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>0.6</td>
<td>0.5</td>
<td>72</td>
</tr>
<tr>
<td>0</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>0.42</td>
<td>0</td>
<td>84</td>
</tr>
</tbody>
</table>

Table 4 Confusion matrices of 12 training samples

The ROC graph is carried out on a 9 testing samples. Figure 1 shows the confusion matrices for testing dataset against 12 and 8 training samples. The cells in each table show the number of cases that were correctly classified positive cases, and the cells show the correctly classified Negative cases. The network's outputs are almost perfect.

![ROC Graph](image)

Fig 2. Roc analysis for 12 and 8 training samples
The Receiver Operating Characteristic (ROC) curve is used to inspect the classifier performance more closely as shown in the Fig.1. By definition, a ROC curve shows true positive rate versus false positive rate (equivalently, sensitivity versus 1–specificity) for different samples in which the classifier output for 12 and 8 training samples is provided. Considering the 9 tested samples the average tprate and fprate is calculated whose output is 47.8% and 65% for 12 training samples. The accuracy for 12 training samples is 55.5%. The average tprate and fprate is calculated whose output is 38.2% and 49.6% for 12 training samples. The accuracy for 12 training samples is 46.6%. From this it has been clear that the 12 number training samples gives better performance than 8 number of training samples against 9 testing samples.

![ROC Curve](image1.png)

**Fig 3.** Classified positive cases (Tprate) against no of iterations

This graph shows that the for 12 training samples, the true positive rate is high for 12 training samples as compared to the 8 training samples for the given dataset of 9 tested samples against number of iterations.

![Classification](image2.png)

**Fig 4.** Classified Negative cases (FPrate) against no of iterations

This graph shows that the for 12 training samples, the false positive rate is high for 8 training samples as compared to the 12 training samples for the given dataset of 9 tested samples against number of iterations.

V. CONCLUSION

Credit default is playing an important role in banking industry. In this paper, a neural network model is developed which applied the learning method to improve the performance of the system in the field of credit risk assessment. The research compares the structure of the artificial neural network model with the traditional method of nationalized bank conducted in the year 2004, 2005 and 2006 and shows that the designed system is having more accurate results as compared than traditional approach. The credit scoring results measured in this research support the hypothesis that ensemble NN that can be used in credit scoring applications to improve the overall accuracy of the designed system. In an overall view, the proposed ensemble model acts as a base classifiers that provides the best accuracy and performance. This research confirms that ANN is the best method for estimation of credit risks in bank.
REFERENCES


Amruta R Ghatge received the Computer Science and Engineering degree in 2007 and Pursuing the Masters Degree from Shivaji University. She published the paper in National and International Conference. She is Having 4 Years of Teaching Experience. Her area of interests are Neural Network, Data Mining.

Prof. P.P. Halkarnikar received the Electronics engineering degree and Masters Degree in Computer Science and Engineering. He is having 24 years of Teaching Experience and 15 years of Industrial Experience. His Publication Includes 20 papers at national and internal conferences and journals. He is ISTE Member of Member of Institute of Engineers (MIE). His Area of Interest are Image Processing, Data Mining, Neural Network.