1 cm

Decision Support System for Bank Loan Classification Using Neural Networks

EJKP Nandani and JR Wedagedara

University of Ruhuna, Matara, Sri Lanka.

nandani@maths.ruh.ac.lk, janak@maths.ruh.ac.lk

Introduction

We report the designing and application of an Artificial Neural Network (ANN) to classify the consumer loan application in banking sector. The system can be used as a "Second Level" filter, upon which shall be supplied with data of loans that already had been approved by the bank officer, to serve as a decision support system in the loan approval process. In particular, this shall be useful in identifying potential risks associated with these loan applications before giving the final approval by the management so can be integrated into the bank's information management system.

Approach: The loan application data set of size 1000 was generated with the computer algebraic and numeric package - Mathematica [1] via Monte-Carlo simulation with 14 covariates (Age,Salary,Spend,Income,etc.) applying the exact criteria adopted by a government bank in Sri Lanka in its actual housing loan application process. Half of the data set which has not been used in the training process has been used to test the best model.

The loan classification process

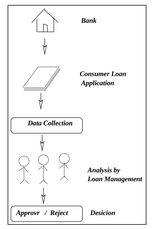


Figure 1: Historically, banking was a local and personal operation. If a customer wants a loan, then he applied for a loan by filling the loan application. The bank manager would then approved or deny. The loan based on combination of the information in the application and their personal knowledge of the customer.

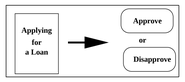


Figure 2: Classifying loan applications for acceptance or rejection.

Method ○ Back-Propagation Neural Network is employed. ***The goal is to minimize the error between the actual output of neural network and desired output.*** ○ The Back-propagation algorithm is an iterative process. In each iteration weights coefficients of nodes are modified. We try to minimize error by using *Steepest Descent Method*

Building A Model: There is no well defined criteria for model selection in neural network. However, *"trial and error"* is usually be the best guide in most cases. The initial weight sets between $(-1, 1)$, for Input and Output layers were generated randomly. Architecture: A 14-X-2 network has been chosen as the optimal architecture. ***The goal of modeling is to find the most appropriate model (in the present case to determine X in the above), fitting the data via mean square error minimization, without over-fitting.***

Results

To get a prediction of classes use a decision rule which approves an input if the probability is greater than 0.5 and rejects with probability less than 0.5.

The dependence of the input variable - Age: We demonstrate the effect of a single variable on the final decision. Let us take the input variable **Age:**

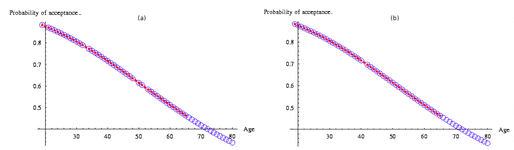


Figure 3: Each point shows the decision taken by the loan officer (i.e. approved or rejected) with the associated probability computed by the Neural Network. "$*$"-approve ,"o"-reject

Clearly, we can see that, there are cases where a loan was rejected for higher values of probability of acceptance and those approved for lower probabilities.

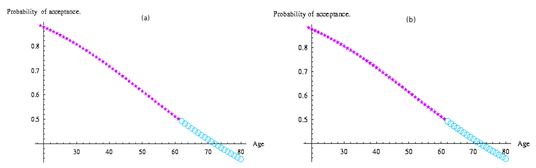


Figure 4: Network decision and Actual output for training and testing data set."$*$"-approve , "o"-reject.

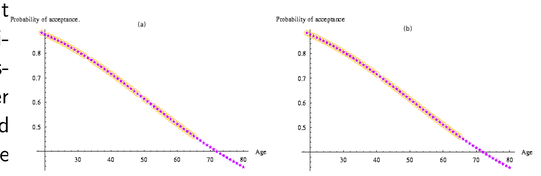


Figure 5: Shows that real and network decision,there are no misclassified decisions over 65 years of age and below 65 years of age both decisions are possible.

Using all covariates as input variables

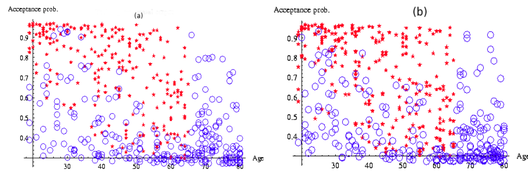
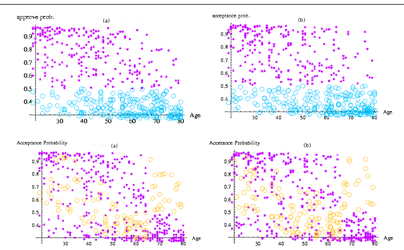


Figure 6: Each point shows the decision taken by the loan officer (i.e. approved or rejected) with the associated probability computed by the Neural Network. "$*$"-approve ,"o"-reject

By comparing the target decision with the probability of acceptance which is the output value in the network, one can see that a loan could have been approved even it has a minimum probability of acceptance and a loan could have been rejected even it has a maximum probability of acceptance. For this case have an accuracy 70% and 72.6% on the training and testing data set respectively.



Network decision and the corresponding probability for training and testing data set respectively."$*$"-approve ,"o"-reject.

Shows how the neural network has performed in classifying the input. **Left:** training set and **Right:** test data set ."$*$"-correct classification , "o"-miss-classification.

Conclusions: (1) The officers would gain their own *experiential knowledge or intuition* (other than those guidelines given from their institution) in deciding whether an application is loan-worthy or not. (2) As the co-relation among the covariates are not considered in this study, we cannot say whether some of the explanatory variables absolutely necessary or not.

References: [1] http://www.wolfram.com

[2] Herbert Lee K.H , *Bayesian Non parametrics Via Neural Network.* , SIAM; Philladelphia [2004]

[3] James A Freeman and David M.Skapure *Neural Networks. (Algorithms, Applications, and Programming Techniques.* , Addison Wesley longman (Singapore) pvt Ltd , [1999]

[4] Sheng-Tun Li , Weissor Shiue and Meng-Huah Huang , *The evaluation of Cnsumer loans using Support Vector Machines.* , Expert System with Applications [2006]