International Journal of
Operations and Quantitative Management

AIMS AND SCOPE

The objective of the International Journal of Operations and Quantitative Management (IJOQM) is to provide an international forum for discussion of advancements in operations management, operations research, quantitative management and management science. The Journal publishes high-quality research articles in all these areas. The Journal particularly encourages articles dealing with international issues, practical experiences, pedagogical aspects including case research, developments in the emerging nations, service management, and other nontraditional and innovative areas. Purely theoretical articles should make significant contribution to the existing knowledge.

Editor-in-Chief

Omprakash K. Gupta
Prairie View A&M University
College of Business
Prairie View, TX 77446, USA

Phone: 1-936-261-9206
Fax: 1-936-261-9220
e-mail: editor@@ijoqm.org
website: www.ijqom.org

The International Journal of Operations and Quantitative Management (ISSN 1082-1910) is published quarterly (March, June, September and December) by International Forum of Management Scholars. Subscriptions, changes of address, and all correspondence should be sent to: AIMS International, P.O. Box 207, 11659 Jones Road, Houston, TX 77070.

International Journal of
Operations and Quantitative Management

Volume 12, Number 3, September 2006

Contents

66. Sanders, S.E. (1998), "Scarlet Letters, bilboes and cable TV: Are shame punishments cruel and outdated or are they a viable option for American jurisprudence?", *Washburn Law Journal*, Vol. 37, pp 359 – 382.
67. Schroeder, C.H., (2002), "Lost in translation: What environmental regulation does that tort cannot duplicate", *Washburn Law Journal*, Vol. 41, pp. 583 – 605.
68. Sheth, J.N. (1975), "Buyer-seller interaction: A conceptual framework", in Anderson, B.B. (Ed.), *Advances in Consumer Research*, Association of Consumer Research, Cincinnati.
69. Sunstein, C. (1996), "On the expressive function of law", *University of Pennsylvania Law Review*, Vol. 144, pp 2021 – 2053.
70. Svensson, G. & Wood, G. (2003), "The dynamics of business ethics: a function of time and culture – cases and models", *Management Decision*, Vol. 41, No. 4, pp. 350-361.
71. Takala, T & Uusitalo O., (1995), "Retailers Professional and Profession-Ethical Dilemmas: The case of the Finnish Retailing Business", *Journal of Business Ethics* 14(11), 893-907.
72. Thompson, J.A. and Hart, D.W.: 2006, "Psychological Contracts: A Nano-Level Perspective on Social Contract Theory", *Journal of Business Ethics* 68(3), 229-241.
73. Tsahuridu, E.E.: 2006, "Anomie and Ethics at Work", Journal of Business Ethics 69(2), 163-174.
74. Tyler, T. R., (1990), *Why people obey the law*, New Haven, Yale University Press.
75. Vavra, T.G., (1992), *After marketing: Hope to keep customers for life through relationship marketing*, Business One Irwin, Homewood, IL.
76. Velasquez, M.G., (1998) *Business Ethics: Concepts and Cases*, 4th ed. Prentice Hall, Englewood Cliffs, New Jersey.
77. Whitman, J.Q. (1998), "What is wrong with inflicting shame sanctions?", *Yale Law Journal*, Vol. 107, pp 1055 – 1092.
78. Williams, K.R. and Hawkins, R. (1986), "Perceptual research on general deterrence: A critical review", *Law & Society Review*, Vol. 20, pp. 545-66.
79. Williams, S.F. (1993), "The soft underbelly of deterrence theory in tort", *Harvard Law Review*, Vol. 106, pp 932 – 944.
80. Wilson, D.T. (1995), "An integrated model of buyer-seller relationships", *Journal of the Academy of Marketing Science*, Vol. 23, No. 4, pp. 335-345.

Econometric Study of Home Loan Approvals Using Statistical Methods

IJQM
Volume 12, *Number* 3
September 2006, *pp.* 231-256
Received: June 2006
Accepted: August 2006

Varsha Bane, Palaniappa Krishnan
Operations Research Program
Department of Food and Resource Economics
University of Delaware, Newark, DE, 19716, USA
(varshakb@yahoo.com)
(baba@udel.edu)

In this paper we discuss the effectiveness of statistical methods used for credit scoring. Financial institutions carry out credit management by adopting statistical methods that develop credit scoring models. Credit or application scoring model is an automated credit approval model for loans or credit applications. Statistical methods such as logistic regression and probabilistic neural network are efficiently used to construct and validate these models based on the internally and externally available credit information of the applicants. Results from both techniques on comparison show that out of all the credit and demographic variables that are available from internal database and/or external credit bureaus only few are the most effective predictors of loan approvals.

1. Objective

The objective of this study was to generate automated credit scoring models using logistic regression and neural network technology. The results obtained from both the models would be compared for accuracy and efficiency of automated loan approvals. This study focuses on the importance and effectiveness of using cross-validation methods to judge model validity.

2. Introduction and Literature Review

According to published literature, statistically based credit decision-making systems were first introduced during the late 1950s. However, they only saw mainstream use during the 1990s with the widespread increase in use of electronic credit information[2]. These statistical techniques are commonly referred to as "credit scoring" models. Initially, scoring models (credit or application) were employed in the consumer credit portfolios of most major banks and credit card issuers. Primary goal was to increase the speed of the credit decision making, enhance the uniformity of the decision process and reduce the overall cost and time duration of decision making. The relative homogeneity of the widely available data that shows credit and the performance or behavioral attributes of applicants, made the initial implementation of scoring models by the lenders successful and profitable[3]. Increased use of credit scoring have materially altered the optimal tradeoffs among information quality, customer service, loan processing costs and bank scale. The output generated by credit scoring models improves the ability of lenders to assess and price default risk. Additionally, it also reduces banks' loan production costs by eliminating expensive on-site visits, reducing loan processing time and generating scale economies associated with automated lending processes[4]; thereby increasing the profitability of the marginal loan applications. However, these methods involve an increased information costs.

These developments will likely have strategic ramifications going forward, such as whether banks do both relationship lending and transactional lending or engage exclusively in just one or the other[5].

What are credit scores? Credit scores are statistically derived measures of creditworthiness of the borrowers that rank order credit applicants according to their degree of credit or default risk[6]. A score (credit or application) is typically associated with an odds ratio, addressing the question: How many applicants are likely to go delinquent (or be rejected) at the corresponding score? Although the models may not predict the absolute level of risk or which borrowers within a score range are likely to perform poorly, the published literature[7] has shown them to be effective tools for ranking the risk of applicants.

"Credit scoring" uses quantitative measures of the performance and characteristics of past loans to predict the future performance of loans with similar characteristics. It is a scientific method of assessing the credit risk associated with new credit applications. Statistical models derive predictive relationships between application information and the likelihood of satisfactory repayment. Models are designed empirically; which means that they are developed entirely from information gained through prior experience. Therefore, credit scoring is an objective risk assessment tool, as opposed to subjective methods that rely on a loan officer's opinion.

Statistical Models (Objective risk assessment method): If a bank has a large pool of loan application data and repayment history for loans over a multi-year period, it is efficient to use the data to derive a statistical model that predicts the risk of a good or bad outcome as defined by the bank's credit policy. Statistical models are the most powerful scoring models. Building a statistical model generally requires an initial data loan pool with at least 1,000 "good" and "bad" outcomes, but ideally considerably more. The actual factors in any custom statistical loan model are determined by testing and generating some key indicators that are likely to be found predictive in many models[8].

Judgmental Models (Subjective risk assessment method): For banks that lack an adequate pool of historic loan data required for deriving a statistical model, a customized rules-based model can be set up. This model can consistently weight the key factors the bank has described as the credit risks of borrowers. The model can then rank order loans from low to high risk by assigning a risk rating.

As mentioned earlier, credit scoring cannot predict individual loan loss; rather it predicts the likelihood or odds of a "bad" outcome, as defined by each bank's credit policy. Nor should a credit scoring system alone approve or reject a loan application; rather the underwriter must decide how he or she will incorporate the credit score into the loan review without any bias[9]. Banks have been using credit scoring successfully for nearly 20 years to make decisions on consumer loans for autos, personal lines of credit and credit cards. The credit quality of a business often

mirrors its principal's credit behavior "if the principal of a business does not pay personal creditors, chances are good that he or she will not pay a business loan". These methods are helpful to make more efficient and reliable loan application decisions, better target marketing plans, to devise more effective collection strategies and increase client retention. They form a solid foundation for pricing loans based on individual client risks and more accurately provisioning against loan losses.

As application credit scoring is the process of predicting the probability that an applicant for a credit product will fail to repay the loan in an agreed manner, to assess this process we require a model which represents the behavior of all applicants for credit. However, typically we have only information about the repayment behavior of those who have been accepted for credit in the past. The behavior of those who had been rejected, if they had been accepted, is unknown. If one estimates a model using accepted applicants only, it may generate biased parameters, if those parameters are applied to a model representing the behavior of all applicants. In addition, if cut-offs are chosen to equalize the actual and predicted number of "bads", then a sample of accept-only is likely to yield inappropriate cut-offs for the population of all applicants.

Several techniques for reducing the magnitude of this bias have been proposed either in the literature. These include extrapolation, augmentation[10], iterative reclassification[11], bivariate probit[12], "parcelling", use of the EM algorithm[13], using a multinomial logistic model[14] and collecting repayment performance data for rejects[15]. The necessary assumptions for the use of these techniques along with data typically used in credit scoring models have been reviewed by a number of authors[16]. Stein[17] provides an analysis of how such cut-off values might be set in various lending environments.

Since lenders must comply with fair lending and fair pricing laws and regulations, the data shortcomings create a critical risk management weakness. That is, gauging regulatory compliance requires a full set of information on credit applicants, including those who did not receive credit. In fact, it is commonly observed that the denied credit applications are the primary focus of fair lending rules and yet, these are the credit applicants for whom institutions have the least information under legacy systems.

In keeping with regulatory guidelines, risk managers create a database of critical information required to comply with the Home Mortgage Disclosure Act (HMDA) and the Community Reinvestment Act. These loan approval processes are expensive and time-consuming. The data items provide the regulator with a sense of the loan application denial disparities among protected classes of persons, including non minority versus minority borrowers and male versus female borrowers.

As stated earlier, application scoring is a statistical means of assessing risk at the point of application for credit. It is primarily used for:

- Determination of Credit risk
- Approval of loan
- Setting limit

Basic characteristics used in scorecards are similar to those used in traditional judgemental lending, for example any loan Application has following three attributes:[18]

1 Purpose of loan: Home loan applications may have varied purposes such as residential, rental, mortgage or tax benefits, equity building, investment, etc.
2 Deposit: Each loan application is supported with documents showing monetary deposits, credit line, credit scoring of the applicant.
3 Security: Application also needs guarantee documents such as sufficient lien over personal or institutional assets for the purpose of security

Financial attributes of loan application include: Assets, Liabilities, Monthly repayment, Total Monthly income of the loan applicant.

Character of an individual applicant is assessed base on: Time at current employment, Residential status, Time at current address, etc. The difference being that attributes within these characteristics are given formal weights (scores) and added to produce a resulting score for an applicant.

Credit scoring (or application scoring) is designed to separate "Goods" from "Bads" using[19]:

Link Function

- The discrete outcome variable (i.e. reject/approve or default/no default) is linked to a set of predictors combined with the help of a set of weights. This is called the *link function*.
- It is calculated on the basis of *earlier applicants*.
- It turns the discrete binary outcome into a continuous probability distribution. The scores predict the place of the *new applicant* in this probability distribution.

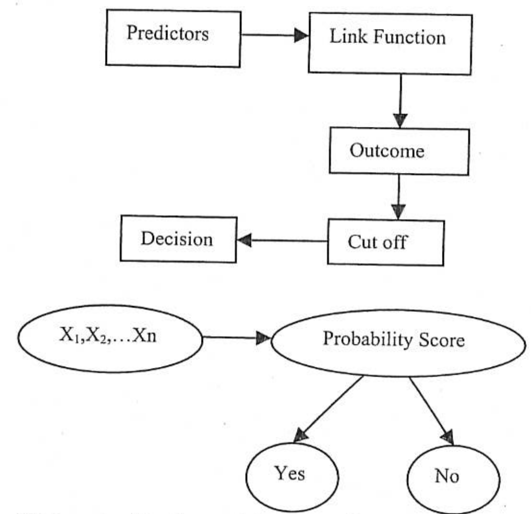
Cut Off Rule

- The scores must be translated into decisions, into a discrete variable once the continuous scores are created.

Good/Bad Odds: A scoring system does not individually identify a good performer from a bad performer, but it classifies an applicant in a particular "Good/Bad odds" group. For example an applicant belonging to a 200 to 1 group, appears pretty safe and profitable while if the applicant belongs to a 4 to 1 risk group, we would no doubt find the risk unacceptable.

There is a "cut-off" point where it is not profitable for the bank to accept a certain Good to Bad ratio[20]. Based on this, it is accepted that there will be some "bads" above the cut-off level set, and some "goods" below the cut-off level set.

'Good/Bad' Discrimination: The objective of a scorecard is to have characteristics which discriminate between Good and Bad accounts with a sufficiently high probability. As discussed earlier, some characteristics are legally or ethically not used (such as gender, race, etc.)



Various methods used for assessing creditworthiness:

- Expert Judgment
- Rules of Thumb
- Point System
- Statistical Scoring

The choices of the link functions[21]: $Yi=f(Xki)$

1. Discriminant function
2. Regression based functions
   - Linear regression (OLS)
   - Logistic regression (logit)
   - Probit
3. Functions based on biological models
   - Neural networks model
   - Genetic algorithm (GA)
4. Other functions
   - Linear programming (LP)
   - Classification/decision trees or recursive partitioning algorithms (RPA)
   - Nearest neighbors.

As Yegorova et al.[22] point out "common variable selection techniques used by practitioners include stepwise logistic regression[23], selection based on the neural network weights[24] and simple plotting of input variables against the dependent variable to identify potentially highly discriminating variables[25]."

Logistic Regression: Risk managers have been using binary logistic regression analysis to determine credit worthiness of an applicant. Use of regression analysis allows for the contemporaneous control of loan characteristics of all applicants within a particular loan program and provides a probability of approval score for every loan application. This approach enables the risk manager to focus on only those marginal or potentially misclassified applications i.e. the situations in which the logistic regression outcome differed from the actual decision. A review of these situations aids in preventing problems when the internal banking data quality and the model are periodically refined and deployed.

Logistic regression has the following form:

$$\ln\left(\frac{p}{1-p}\right) = \sum_{j=0}^{k} \beta_j x_j$$

$$p = \frac{\exp\left(\sum_{j=0}^{k} \beta_j x_j\right)}{1+\exp\left(\sum_{j=0}^{k} \beta_j x_j\right)}$$

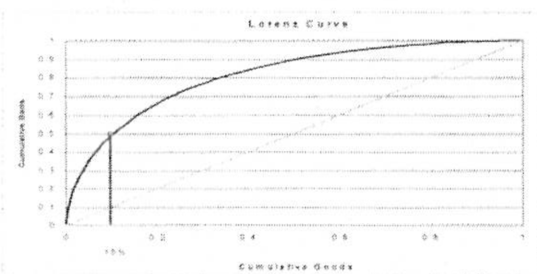
Measures of Discrimination (I):

Receiver Operating Curve (ROC): The Receiver Operating Curve is the area under the curve generated when the cumulative Bads are plotted against the cumulative goods (Lorenz Curve).

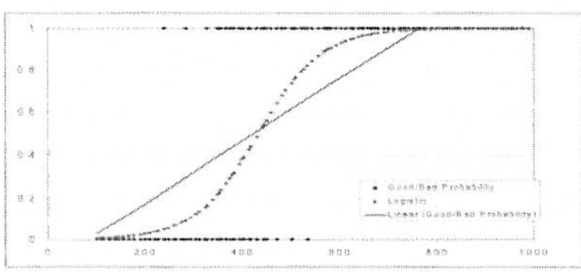
Gini coefficient (G): This discrimination measure is geometrically defined as the ratio of the area of the shaded semi-circular area to the area of the triangle in the Lorenz diagram.

$$ROC = \frac{1}{2}(G+1)$$

PH (percentage Good for 50% Bad): This is defined as the cumulative proportion of Goods up to the median value of the Bads.



As seen in following figure, logistic regression fits the probability better than linear regression.



Measures of Discrimination (II): Discrimination measures that need to be determined for discrete attributes:

- Chi-Squared $\sum \frac{(Obs - Exp)^2}{Exp}$

- Fico (Kullback Divergence)[18revisited] $100\sum(G_i - B_i)\ln\left(\frac{G_i}{B_i}\right)$

Probabilistic Neural Network (PNN): There has been much written about the use of Artificial Neural Networks in the loan approval process. In part this analysis is carried out to build managerial confidence in the use of PNN's in general for predicting loan approvals. PNNs combine some of the structures of traditional statistical pattern recognition and feed-forward NNs. These NN implementations use discriminant analysis process structures that were introduced by Donald Specht[26]. PNNs feature very fast training times and produce outputs using Bayes posterior probabilities. They have been found to be much more efficient and effective than general purpose back propagation NNs. They are ideal for the purpose of this study. More indepth discussion of PNNs can be found in literature[27]. One of the advantages of PNNs is their ability to fit models to data with missing values.

In the present study, it is found that the importance of weights of different variables is very useful in trimming variables from the pool of 50 potential input variables. This is because the genetic algorithm (GA) of neural network software provides a relative measure of the importance of each variable in predicting the dependent variable i.e. loan approvals.

3. Data and Methodology

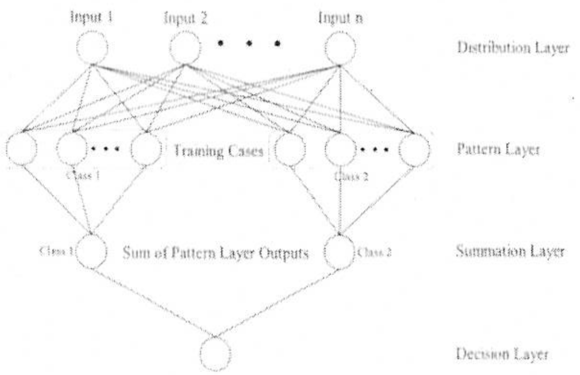
A commonly available dataset that was originally used in a famous study by researchers at the Boston Federal Reserve Bank[1] was used. In order to comply with the rules for fair lending, five variables were eliminated from the study which was legally and ethically not allowed. The input variables, mainly exhibiting traditional financial characteristics, were then used to build a logistic regression model and a probabilistic neural network model to correctly predict the ultimate disposition of majority of the loan approvals.

The data set used for this study was 1996 data on 1,989 home loan applications drawn from a large regional lender's activities in a single Metropolitan Statistical Area (MSA). We have chosen to build statistical models that adhere to the rules and regulations of fair lending thereby generating automated application scoring models for improving the overall quality of loan approval. PNN and MLR techniques were used to build scoring models. Data is available upon request. Following figure (Source Yegorova et al. 2001) illustrates a simple PNN.

Data–Predictors: 50 input variables which include loan Amount, term, interest rate, applicant income, outstanding liability, etc. (for details refer Table 1 of Appendix)

Data –Response: Single binary variable Approve or Reject

Total Number of Loan Applications: 1,989



Data set used for this study was already normalized for outliers, invalid values, duplicates. We have chosen a 60% random sample of 1,195 of these 1,989 customers to create a binary logistic regression model; then use the model to validate or score the remaining 40% of the data set or 794 customers as good or bad credit risks.

4. Results and Discussion

(I) *Using Multinomial Logistic Regression (MLR)-Model Build:* A multivariate statistical method was used:

Steps followed:

- Logistic Regression (Binary logistic regression is an appropriate technique for the present data because the response variable is dichotomous).
- Stepwise selection methods.
- Residual analysis.

Not all predictive characteristics were used in the model because an inter-correlation effect may exist between variables. e.g. mortno and mortperf are excluded from model built due to correlation effect.

A random sample of 1,071 (124 missing of 1195) applications namely "training" was used for building and validating the model, holding the 793 (1 missing of 794) prospects namely "test" aside for later scoring using the best predictors from the model built. Categorical variables were defined using class statement and we have used the default effects coding for the same.

Out of 50 predictor variables, 12 variables were found to be the most statistically predictive variables (refer Table 3 of Appendix). From the pool of 1,071 customers modeled 939 are approvals and 132 rejects. The resulting applicant-level probabilities of approval were used to "score" the applications based on the *prima facie* probability of denial of 0.12 percent.

The response variable was dichotomous and the event modeled was 'approval' of loans. The model used an automatic backward stepwise selection procedure. This selection begins by selecting the strongest candidate predictor, ten testing additional

candidate predictors, one at a time, for inclusion in the model. At each step, it checks to see whether a new candidate predictor will improve the model significantly at 0.1 level. It also checks to see if a new predictor is included in the model, any other predictors already in the model should stay or be removed. If a newly entered predictor does a better job of explaining loan approval inference, then it is possible for a predictor already in the model to be removed from the model because it no longer uniquely explains enough. This stepwise procedure continues until all the candidate predictors have been thoroughly tested for inclusion and removal at 0.1 level of significance. Summary of this stepwise selection is shown in Table 2 of Appendix which displays all 12 predictors in order of their importance or significance in the model.

However, due to the stepwise variable selection method, the significance levels associated with the model predictors may be somewhat inaccurate because they are assuming a single step process rather than a multi-step process. Therefore, additional diagnostics were used to give us more confidence in our model.

As the discrete binary outcome was converted into a continuous probability distribution, the Hosmer and Lemeshow goodness-of-fit test was performed using lackfit option. This test divides subjects into approximately ten groups of roughly the same size based on the percentiles of the estimated probabilities. The discrepancies between the observed and expected number of observations in these groups are summarized by the Pearson chi-square statistic. Resulting large p-value of 0.9966 suggests that the fitted model is an adequate model[28]. The results of Hosmer and Lemeshow test are displayed in Table 3 of Appendix.

Classification and Validation: Cross tabulating observed response categories with predicted categories helped to determine how well the model identified approvals. Classification table output is shown in Table 4 of Appendix. "Ctable" option was used to classify the input binary response observations according to whether the predicted event probabilities are above or below the a cut point value in the range (0,1). An observation was predicted as an event if the predicted event probability exceeds the cut point.

If growing the customer base or approval is the primary concern, then one would want to lower Type I error and maximize "Sensitivity"(sensitivity is the probability that an approval is correctly classified). However, if decline or reject a customer is the priority, then one should lower Type II error and maximize "Specificity" (specificity is the probability that a reject is correctly classified).

Usually both are major concerns, we decided to choose a decision rule for classifying customers that gives the best mix of sensitivity and specificity. A *prima facie* probability cut off point of 0.9 (90%) was chosen because it seems to give a balance of all the five measures of classification accuracy. But in practice, depending on the specific objectives and judgmental underwriting model, one may want to experiment with various cut points to see how they affect the model's sensitivity and specificity by examining the rates of correct classification for each model.

The ROC curve shown in Figure 1 of Appendix, is the sensitivity plotted against one minus the specificity for various value of probabilities shows a c-value of 0.957

which is the area under the ROC curve. It is a common measure of the ability of a model to correctly classify the observations. The high value of 'c' shows that the model built had a very good predictive power.

Since there was no interaction between covariates, such as gdlin*obrat interaction, the odds ratio for these predictors were computed. The estimated odds ratios for all predictors and 90 percent confidence interval for the same are shown in Table 5 of Appendix. Note that these odds ratio estimates were not the same as the corresponding values in the Exp(Estimates) in the parameter estimates table because effect coding was used.

Reschi option was used to specify Pearson (Chi) residual for identifying the observations that were poorly accounted for by the model. A very high Pearson Chi-Square residual indicates that very few observations were poorly accounted for by the model which is confirmed in Table 6 of Appendix.

R-Squared statistics which measures proportion of variance of the response that is predictable from (that can be explained by) the regressor variables that is explained by a linear regression model, can not be computed for a logistic regression model because of dichotomous response than continuous. Instead a pseudo R-Squared statistics with a value of 0.6672 was used. It seems that approximately 67% of the variation in the dependent variable was explained by the 12 predictors in our final model. Table 7 of Appendix displays the R-Squared statistics.

Note[29]: In logistic regression, there is no true R^2 value as there is in OLS regression. However, because deviance can be thought of as a measure of how poorly the model fits (i.e., lack of fit between observed and predicted values), an analogy can be made to sum of squares residual in ordinary least squares. The proportion of *unaccounted* for variance that is reduced by adding variables to the model is the same as the proportion of variance accounted for, or R^2. Where the null model is the logistic model with just the constant and the *k* model contains all the predictors in the model.

$$R^2_{\text{logistic}} = \frac{-2LL_{null} - 2LL_k}{-2LL_{null}}$$

$$R^2_{OLS} = \frac{SS_{total} - SS_{residual}}{SS_{total}} = \frac{SS_{regression}}{SS_{total}}$$

Model Fit Statistics: Akaike information Criterion (AIC) = -2LogL + 2 p where p= k+s=(total number of response levels-1) + number of explanatory variables. Table 8 of Appendix exhibits model fit statistics. AIC is a penalized maximum likelihood statistic that gives a very measure for assessing how well a particular model fits. Stepwise logistic regressions were carried out on various randomized samples of the data and the sample which gave the lowest AIC value for intercept and covariates was selected. Comparing AIC values of different models is a useful tool of model fit statistics.

Also, there was a quasi-complete separation of data points that separated the events from non-events almost completely. If the iterative process of maximizing the likelihood function is allowed to continue, the dispersion matrix becomes unbounded and the log likelihood diminishes to a nonzero constant. The coefficient estimates may be more suspect in this case and it would be difficult to know the structure i.e. which covariate influences the event and by how much. The 'score' procedure in SAS was used to score remaining 794 randomly selected control group (hold out sample) or proxy portfolio namely 'test' to score the model based on the model derived from "training" sample. The 692 approvals and 101 rejects from this new data sample show mean of -0.25 and 1.68 respectively for Factor1 as shown in Figure 2 of Appendix. In a common factor analysis the true factor scores have mean zero and variance one, but the computed factor scores were only estimates of the true factor scores[30].

This valid predictive model containing 12 most predictive variables was used to score a prospect sample namely "test". The following logistic regression equation for estimated logit was used to score and cross validate "test" data set.

$\widehat{\textbf{logit}}$ = .8012 + gdlin *3.4621 + typur *4.6739 + inson *-4.3785 + prop *-0.4207 + unver *-2.5493 + fixadj *0.8607 + pubrec *-1.3538 + vr *-0.8776 + obrat *-0.0296 + old *-0.745 + married *0.8341 + gift *-0.7603

To estimate the probability of approval in "test" data set, an inverse link function was applied as follows: Probability (loan approval) = exp ($\widehat{\text{logit}}$)/(1+exp($\widehat{\text{logit}}$))

These 794 customers were then divided into ten segments or deciles based on their scores, 1 being most likely to be approved and 10 being most likely declined. Model predictions were then compared (Refer Table 9 of Appendix) with actual approval records. Approval decisions were predicted accurately for 83% customers and failed totally for 0.00433. While reject decisions were predicted accurately for 37% customers and failed totally for 11%. The remaining marginal 17% approval and 52% reject decisions can be scored using the model based on specific objectives. This difference in prediction accuracy of approvals and rejects is already explained earlier in discussing cut off point. Although an accurate prediction of both approvals and rejects is a major concern, one need to choose a decision rule for classifying customers giving the best mix of sensitivity and specificity Table 9 of Appendix compares the outcomes from the MLR with the judgmental underwriting model for the test sample. Using the weights derived from the logistic regression, the probability of approval derived for each applicant and the *prima facie* cutoff of 0.9 gave a well fitted MLR model for predicting loan approvals. Figures 3.a and 3.b of Appendix display dichotomous response of MLR for "training" and "test" data sets while 3.c and 3.d display approval probabilities for the same.

Marginal Applicant Focus: As discussed earlier, the benefit and/or disadvantage of credit scoring tends to be illustrated best when reviewing marginal applicants. Marginal applicants are those with credit scores that are at or near the cutoff for denial.

However, based on specific objectives, one may want to experiment with various cut points based on our business needs (either reject accurately or approve accurately). For a good classification rule the sensitivity, specificity and percent correct should be almost 100 and false positive and false negative to be zero. The sensitivity and specificity are antagonistic in character. One can get high sensitivity at the expense of a low specificity and vice versa. Therefore, increasing accuracy of prediction is a business rule than statistical based on primary goal of the business.

(II) *Using Probabilistic Neural Networks (PNNs)*:

Network Training and Testing: In this study, a PNN was used where the structure of the network was selected via a genetic algorithm (PNN-GA). The PNN-GA was fitted using the software package NeuroShell®Predictor by Ward Systems on all 1,989 applications with 50 input variables. As shown in Figure 4 of Appendix, the 9 input variables with relative importance greater than 30% were chosen to be the most effective predictors which were further used for training and testing the PNN using NeuroSolutions5.

Since the evaluation version of NeuroSolutions5 allowed to fit PNN for only 500 applications, a random sample of 500 applications from the total 1,989 applications was used as follows: 300 for training, 75 for cross-validation and remaining 125 for testing the trained network. Also, note that the dependent variable was dichotomous with a 1 for approvals of loan applications and a zero for rejection.

The sensitivity of these 9 predictors about mean was tested using the training data set which is shown in Figure 5 of Appendix. For training the network, 1,000 epochs were carried out for 'training' and 'cross-validation' samples. Figure 6 of Appendix shows that the minimum squared errors (MSE) versus numbers of epoch is almost constant over 1,000 epochs.

Subsequently trained network was used to measure the accuracy of predictions by using 'test' sample. A test data base of 125 loan applications was used to assess trained network performances to 1,000 fold cross-validation methods. As shown in Table 10 of Appendix, PNN was found to be an effective predictor of loan approvals with 88% correct predictions. Table 11 of Appendix shows a detailed performance of testing network on test data set. A graphical display of actual and desired (or predicted) output is shown in Figure 7 of Appendix.

Thus, the 1,000-fold cross-validation was found to be useful in predicting future performance and confirms model validity.

5. Comparison of logistic regression and Neural Network Techniques

As discussed above, the original PNN based on 300 training observation predicted new loan approvals more accurately than that obtained using multinomial logistic regression (MLR) model. The PNN model predicted 110 of 125 loan decisions(88%) accurately while MLR model is accurate predicting 578 of 693 approvals (83%). However, the PNN model was fitted on a sample of total 500 applications and MLR was fitted for all 1,989 applications. It is therefore not recommended to confirm which model is a better suit in predicting loan approvals.

As displayed in Table 12 of Appendix, the 12 most important predictor candidates obtained using MLR were compared with the 9 obtained using genetic algorithm of Neural Network model using NeuroShell predictor-2. With exception of few predictors, both techniques came up with a different set of key predictors.

Although, PNN is found to be more effective than traditional statistical procedures including logit and discriminant analyses; neural network software used has a disadvantage of taking longer time to process PNN-GA when working with a large pool of data. It also runs the problem of over training the network used. On the other hand MLR is a very simple method to analyze large data set in short span of time using stepwise selection.

From the practical standpoint, the outcome of some loans are impossible to predict because of the uncertainty of future events such as sector-specific or macroeconomic downturns. Thus, those studying these results must remember that some results are completely unpredictable.

6. Summary

Methods considered for data analysis and their brief outcomes are as follows:

Logistic Regression (pursued):

- Carried out discrimination between the values of final 12 predictors.
- Classified applicants into approved/rejected based on the built model.
- Predicted future loan approvals automatically using our model.
- Reduced misclassifications using the best cut off point.

Discriminant Analysis (pursued):

- Actual proportion of approved loans was known and compared with predictions.
- Cost of misclassification unknown.
- Predictors do not appear normally distributed.

Probabilistic Neural Network (pursued)

- Genetic algorithm trimmed 50 input variables to best 9 predictors.
- Trained network is used to accurately classify applicants into approved or rejected.

The most critical piece of the fair lending compliance puzzle is the capture and inclusion of information on credit history and related financial characteristics. In present study, both models have been successful in using binary logistic regression and PNN techniques to identify behavioral and demographic characteristics associated with likelihood to approve or reject a home loan application. MLR does a good job of predicting these approvals and rejects depending upon what we want to improve the most approval prediction or reject prediction. In other words, the accurate use of the model built using statistical models depends on the ultimate business decision.

Further research needs to address the following:

- Re-estimation of PNN model using the full data set of 50 input variables for all 1,989 loan applications using NeuroSolutions5 full version. The research issue of interest is network model stability in independent variables, network structure, and network weights.
- A simulated usage of models updating and validation using the chronological sequence of loan applications for the Post-One-Year loans. The addressed research issues will be network model stability in independent variables, network structure, and network weights.

7. Conclusion

Clearly, credit scoring is a risk management tool. Scoring systems can help a bank ensure more consistent underwriting and can provide management with a more insightful measure of credit risk. Consequently, credit scoring is not meant to increase approval rates; rather, it promotes consistency and efficiency while maintaining or reducing historic delinquency rates. It also allows the users to focus their attention and time on applications that are not obvious approvals or obvious declines also referred to as 'marginal applicants'. These scoring models are of great benefit to senior risk managers, financial-services professionals, and regulators to manage large number of loan applications thereby enhancing the fair lending as a result. These are also useful for consistent risk-based pricing, credit policy implementation and prioritization of collections. Finally, the results have important implications for bank supervision. Currently, bank supervisors, financial regulators and researchers focus their fair lending concern on dealing with disparate treatment.

Acknowledgments: The authors wish to thank the following people for their valuable support and guidance:

1. Jeffrey M. Wooldridge, for the dataset of 1,989 individual observations included in Wooldridge, Jeffrey M. Introductory Econometrics: A Modern Approach, 2nd Edition, South Western College Publishing, (2003). **Source:** W.C. Hunter and M.B. Walker (1996), "The Cultural Affinity Hypothesis and Mortgage Lending Decisions," *Journal of Real Estate Finance and Economics*13,57-70. **Notes:** These data were originally used in a famous study by researchers at the Boston Federal Reserve Bank. See A. Munnell, G.M.B. Tootell, L.E. Browne, and J. McEneaney (1996), "Mortgage Lending in Boston: Interpreting HMDA Data," *American Economic Review* 86, 25-53.

2. Dr. John Pesek, Assistant Professor, University of Delaware—for help with logistic regression.

3. Dr. John Bernard, Assistant Professor, University of Delaware – for help with econometric analysis.

8. References

1. Hunter, W.C. and Walker, M.B. (1996), "The Cultural Affinity Hypothesis and Mortgage Lending Decisions," *Journal of Real Estate Finance and Economics*13,57-70.

2. Lewis, E. M. (1994), *An Introduction to Credit Scoring* for a history of credit scoring models.

3. Collins, M. C., *University of Tennessee,* Keith D. Harvey, *Boise State University,* Peter J. Nigro, "The influence of bureau scores, customized scores and judgmental review on the bank underwriting decision-making process", *The Office of the Comptroller of the Currency*

4. Mester, L.J. (1997), "What's the Point of Credit Scoring?" Federal Reserve Bank of Philadelphia, *Business, Review* September/October: 3-16. AND Rossi, C.V. (1998), "Mortgage Banking Cost Structure: Resolving an Enigma." *Journal of Economics and Business* 50(2): 219-234.

5. Boot, A.W.A. and Thakor, A.V. (2000), "Can relationship banking survive competition?" *Journal of Finance* 55, 679-713. AND DeYoung, R., Hunter, W.C. and Udell, G.F. (2004), "The Past, Present, and Probable Future for Community Banks." *Journal of Financial Services Research* 25(2/3): 85-133.

6. Avery, R., Bostic, R. W., Calem, P. and Canner, G. (1996), "Credit Risk, Credit Scoring, and the Performance of Home Mortgages," *Federal Reserve Bulletin*, 82 (7), pp. 621-648 AND McLean, F. M. (1995), "The Predictive Power of Selected Credit Scores," *Freddie Mac Letter* AND Pierzchalski, L. (1996), "Guarding Against Risk," *Mortgage Banker* (June), pp. 38-45.

7. Avery, R., Bostic, R. W. , Calem, P. and Canner, G. (1996), "Credit Risk, Credit Scoring, and the Performance of Home Mortgages," *Federal Reserve Bulletin*, 82 (7), pp. 621-648 AND McLean, F. M. Virginia (1995), "The Predictive Power of Selected Credit Scores," *Freddie Mac Letter* AND Pierzchalski, L. (1996), "Guarding Against Risk," *Mortgage Banker* (June), pp. 38-45.

8. Caire, D. and Kossmann, R. (2003), for Bannock Consulting on a Technical Assistance funded by the *European Bank for Reconstruction and Development and the European Union*, "Credit Scoring: Is It Right for Your Bank?"

9. Caire, D. and Kossmann, R. (2003), for Bannock Consulting on a Technical Assistance funded by the *European Bank for Reconstruction and Development and the European Union*, "Credit Scoring: Is It Right for Your Bank?"

10. Hsai, D. C. (1978), Credit scoring and the Equal Credit Opportunity Act. *The Hastings Law Journal*, 30, November: 371-448.

11. Joanes, D. N. (1993/4), Reject inference applied to logistic regression for credit scoring. *IMA Journal of Mathematics Applied in Business and Industry*, 5: 35-43.

12. Boyes, W. J., Hoffman, D. L. & Low, S. A. (1989), An econometric analysis of the bank credit scoring problem. *Journal of Econometrics* 40: 3-14.

13. Demster, A.P., Laird, N.M. & Rubin, D.B. (1977), Maximum Likelihood from incomplete data. *Journal of the Royal Statistical Society* B 39: 1-38.

14. Reichert, A.K., Cho, C. C. & Wagner, G.M. (1983) An examination of the conceptual issues involved in developing credit scoring models. *Journal of Business and Economic Statistics,* 1: 101-114.

15. Hand, D.J. and Henley, W.E. (1993/4), Can reject inference ever work? *IMA Journal of Mathematics Applied in Business and Industry,* 5: 45-55. AND Ash, D. and Meester, S. (2002), *Best Practices in Reject Inferencing.* Presentation at Credit Risk Modelling and Decisioning Conference, Wharton Financial Institutions Center, Philadelphia, May 2002.

16. Ash, D. and Meester, S. (2002), *Best Practices in Reject Inferencing.* Presentation at Credit Risk Modelling and Decisioning Conference, Wharton Financial Institutions Center, Philadelphia, May 2002. AND Banasik, J. B., Crook, J. N. & Thomas, L. C. (2001), Sample selection bias in creditvscoring models. Working Paper 01/5, Credit Research Centre, University of Edinburgh. AND Hand, D.J. and Henley, W.E. (1993/4), Can reject inference ever work? *IMA Journal of Mathematics Applied in Business and Industry,* 5: 45-55. AND Joanes, D. N. (1993/4), Reject inference applied to logistic regression for credit scoring. *IMA Journal of Mathematics Applied in Business and Industry,* 5: 35-43. AND Thomas, L. C., Edelman, D. E. & Crook, J. N. (2002) *Credit Scoring and its Applications.* Monographs on Mathematical Modelling and Computation, Philadelphia: *Society for Industrial and Applied Mathematics.*

17. Stein, R.M. (2005), "The Relationship Between Default Prediction and Lending Profits: Integrating ROC Analysis and Loan Pricing." *Journal of Banking and Finance* 29(5): 1213-1236.

18. Marinopoulos, J., Head of Retail Decision Model, (2002), "Credit Scoring Development and Methods" AND Based on a book by Solomon Kullback *"Information Theory and Statistics"*

19. Rona-Tas, A., University of California, San Diego, "Rational Calculation and Trust: A Comparative Institutional Analysis of Emerging Credit Card Markets in Transition Economies".

20. Marinopoulos, J., Head of Retail Decision Model, (2002), "Credit Scoring Development and Methods".

21. Rona-Tas, A, University of California, San Diego, "Rational Calculation and Trust: A Comparative Institutional Analysis of Emerging Credit Card Markets in Transition Economies".

22. Yegorova, I, Andrews, B. H., Jensen, J.B., and Smoluk, B. J., Walczak, S., (2001), "A Successful Neural Network-Based Model for Predicting Small Business Loan Default," *The Credit and Financial Management Review,* Vol. 7, Fourth Quarter.

23. Klein, B.D., Rossin, D.F., (1999), "Data Errors in Neural Network and Linear Regression Models: An Experimental Comparison," *Data Quality,* Vol. 5 (1) AND Walczak, S., and Cerpa, N., (1999), "Heuristic Principles for the Design of Artificial Neural Networks," *Information and Software Technology,* Vol. 41(2), 109-119.

24. Tyree, E., and Long, J.,(1996), "Bankruptcy Prediction Models: Probabilistic Neural Networks versus Discriminant Analysis and Backpropagation Neural

Networks," Department of Business Computing, School of Informatics, City University, London, UK.

25. Tucker, J., (1998), "Neural Networks versus Logistic Regression in Financial Modeling:a Methodological Comparison," Plymouth Business School, University of Plymouth, UK.

26. Specht, D. (1988), Probabilistic Neural Networks for Classification, Mapping, or Associative Memory. Proceedings of the IEEE International Conference on Neural Networks, 1, 525-532.

27. Specht, D. (1990), Probabilistic Neural Networks. Neural Networks, 3, 109-118 ANDhttp://chemdivwww.nrl.navy.mil/6110/sensors/chemometrics/pnn.htmlANDhttp://www.mathtools.net/MATLAB/Neural_Networks/

28. Hosmer and Lemeshow, (2000), Applied Logistic Regression, Second Edition

29.www.ioa.pdx.edu/newsom/da2/ho_logistic3.doc

30.http://support.sas.com/91doc/docMainpage.jsp

Appendix

Table 1: *Basic statistics of the population of 1,989 customers with 56 explanatory variables and single binary or dichotomous response variable namely 'action'. We have eliminated the five variables represented in bold letters because they are not used legally and ethically*

Variable #	Variable Code	Variable Description	#obs	# Misses	Mean	Std Dev	Min	Max
1	occ	Occupancy	1989	0	1.03	0.19	1	3
2	Loanamt	loan amt in thousands	1989	0	143.25	80.52	2	980
3	Suffolk	=1 if property in suffolk co.	1989	0	0.15	0.36	0	1
4	Appinc	applicant income, $1000s	1989	0	84.68	87.06	0	972
5	Typur	type of purchaser of loan	1989	0	1.53	2.61	0	9
6	Unit	number of units in property	1989	4	1.12	0.44	1	4
7	Married	=1 if applicant married	1989	3	0.66	0.47	0	1
8	Dep	number of dependents	1989	3	0.77	1.10	0	8
9	Emp	years employed in line of work	1989	0	0.21	1.00	0	9
10	Yjob	years at this job	1989	0	0.45	1.12	0	9
11	Self	=1 if self employed	1989	0	0.13	0.34	0	1
12	atotinc	total monthly income	1989	0	5195.55	5269.06	0	81000
13	cototinc	coapp total monthly income	1989	0	1547.18	2361.81	0	41667
14	Hexp	propose housing expense	1989	0	1504.90	833.98	154	10798
15	Price	purchase price	1989	0	196.26	128.12	25	1535
16	other	other financing, $1000s	1989	0	2.37	28.23	0	1020
17	liq	liquid assets	1989	0	4618.03	67126.13	0	1000000
18	rep	no. of credit reports	1989	9	1.50	0.99	0	9
19	gdlin	credit history meets guidelines	1989	0	1.58	21.09	0	666
20	lines	no. of credit lines on reports	1989	0	516.36	22422.11	0	999999
21	mortg	credit history on mortgage paym	1989	0	1.71	0.56	1	4
22	cons	credit history on consumer stuf	1989	0	2.11	1.66	1	6
23	pubrec	=1 if filed bankruptcy	1989	0	0.07	0.25	0	1
24	hrat	housing exp, % total inc	1989	0	24.79	7.12	1	72
25	obrat	other oblgs, % total inc	1989	0	32.39	8.26	0	95
26	fixadj	fixed or adjustable rate?	1989	0	0.31	0.46	0	1
27	term	term of loan in months	1989	0	2351.46	44795.74	6	999999
28	apr	appraised value	1989	0	205.09	156.13	25	4316
29	prop	type of property	1989	0	1.86	0.54	1	3
30	inss	PMI sought	1989	0	0.20	0.40	0	1
31	inson	PMI approved	1989	0	0.02	0.12	0	1
32	gift	gift as down payment	1989	0	0.16	0.37	0	1
33	cosign	is there a cosigner	1989	0	0.03	0.17	0	1
34	unver	Unverifiable info	1989	0	0.04	0.20	0	1
35	review	number of times reviewed	1989	0	113.73	314.66	0	999
36	netw	net worth	1989	0	266.57	1110.18	-7919	28023
37	unem	unemployment rate by industry	1989	0	3.88	2.16	1.8	10.6
38	**min30**	**=1 if minority pop. > 30%**	**1989**	**183**	**0.06**	**0.23**	**0**	**1**
39	bd	=1 if boarded-up val > MSA med	1989	0	0.42	0.49	0	1
40	mi	=1 if tract inc > MSA median	1989	0	0.87	0.33	0	1
41	old	=1 if applic age > MSA median	1989	0	0.47	0.50	0	1
42	vr	=1 if tract vac rte > MSA med	1989	0	0.41	0.49	0	1
43	sch	=1 if > 12 years schooling	1989	0	0.77	0.42	0	1
44	**black**	**=1 if applicant black**	**1989**	**0**	**0.10**	**0.30**	**0**	**1**
45	**hispan**	**=1 if applicant Hispanic**	**1989**	**0**	**0.06**	**0.23**	**0**	**1**
46	**male**	**=1 if applicant male**	**1989**	**15**	**0.81**	**0.39**	**0**	**1**
47	mortno	no mortgage history	1989	0	0.12	0.33	0	1
48	mortperf	no late mort. payments	1989	0	0.88	0.33	0	1
49	mortlat1	one or two late payments	1989	0	0.33	0.47	0	1
50	mortlat2	> 2 late payments	1989	0	0.64	0.48	0	1
51	chist	=0 if accnts deliq. >= 60 days	1989	0	0.02	0.14	0	1
52	multi	=1 if two or more units	1989	0	0.01	0.10	0	1
53	loanprc	amt/price	1989	0	0.84	0.37	0	1
54	thick	=1 if rep > 2	1989	4	0.09	0.28	0	1
55	**white**	**=1 if applicant white**	**1989**	**0**	**0.77**	**0.19**	**0.02**	**2.57**
56	Action(Response)	type of action taken, 1=if Approve	1989	0	0.88	0.33	0	1

Table 2: *Summary of Stepwise Selection shows the final 12 most predictive variables in the order of their importance.*

Step	Effect	Description of Predictor	DF	Number In	Score Chi-Square	Pr > ChiSq
1	gdlin	credit history meets guidelines	2	1	389.6207	<.0001
2	typur	type of purchaser of loan	7	2	79.8814	<.0001
3	inson	PMI approved	1	3	37.499	<.0001
4	prop	type of property	2	4	34.2513	<.0001
5	unver	unverifiable info	1	5	18.4631	<.0001
6	fixadj	fixed or adjustable rate?	1	6	8.4448	0.0037
7	pubrec	=1 if filed bankruptcy	1	7	9.1295	0.0025
8	vr	=1 if tract vac rte > MSA med	1	8	5.6248	0.0177
9	obrat	other oblgs, % total inc	1	9	4.8863	0.0271
10	old	=1 if applic age > MSA median	1	10	5.4273	0.0198
11	married	=1 if applicant married	1	11	5.2876	0.0215
12	gift	Gift as down payment	1	12	3.5043	0.0612

Table 3: *Hosmer and Lemeshow test for Goodness of fit shows that observed and predicted values are almost the same for different probability percentiles.*

Partition for the Hosmer and Lemeshow Test

Group	Total	action = Approve		action = Reject	
		Observed	Expected	Observed	Expected
1	107	18	19.36	89	87.64
2	107	81	81.24	26	25.76
3	107	99	97.33	8	9.67
4	107	101	101.61	6	5.39
5	107	104	104.15	3	2.85
6	107	107	106.39	0	0.61
7	107	107	106.94	0	0.06
8	108	108	107.98	0	0.02
9	87	87	87	0	0
10	127	127	127	0	0

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
1.2092	8	0.9966

Table 4: *Classification Table*

Probability	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensitivity	Specificity	False+ve	False-ve
0	939	0	132	0	87.7	100	0	12.3	
0.02	936	13	119	3	88.6	99.7	9.8	11.3	18.8
0.04	935	23	109	4	89.4	99.6	17.4	10.4	14.8
0.06	935	29	103	4	90	99.6	22	9.9	12.1
0.08	935	34	98	4	90.5	99.6	25.8	9.5	10.5
0.1	933	42	90	6	91	99.4	31.8	8.8	12.5
0.12	931	48	84	8	91.4	99.1	36.4	8.3	14.3
0.14	931	52	80	8	91.8	99.1	39.4	7.9	13.3
0.16	931	56	76	8	92.2	99.1	42.4	7.5	12.5
0.18	931	58	74	8	92.3	99.1	43.9	7.4	12.1
0.2	931	59	73	8	92.4	99.1	44.7	7.3	11.9
0.22	931	62	70	8	92.7	99.1	47	7	11.4
0.24	931	63	69	8	92.8	99.1	47.7	6.9	11.3
0.26	930	64	68	9	92.8	99	48.5	6.8	12.3
0.28	929	66	66	10	92.9	98.9	50	6.6	13.2
0.3	929	68	64	10	93.1	98.9	51.5	6.4	12.8
0.32	929	71	61	10	93.4	98.9	53.8	6.2	12.3
0.34	928	73	59	11	93.5	98.8	55.3	6	13.1
0.36	927	75	57	12	93.6	98.7	56.8	5.8	13.8
0.38	925	77	55	14	93.6	98.5	58.3	5.6	15.4
0.4	925	77	55	14	93.6	98.5	58.3	5.6	15.4
0.42	924	80	52	15	93.7	98.4	60.6	5.3	15.8
0.44	924	81	51	15	93.8	98.4	61.4	5.2	15.6
0.46	922	82	50	17	93.7	98.2	62.1	5.1	17.2
0.48	921	82	50	18	93.7	98.1	62.1	5.1	18
0.5	918	84	48	21	93.6	97.8	63.6	5	20
0.52	917	85	47	22	93.6	97.7	64.4	4.9	20.6
0.54	915	87	45	24	93.6	97.4	65.9	4.7	21.6
0.56	915	89	43	24	93.7	97.4	67.4	4.5	21.2
0.58	913	89	43	26	93.6	97.2	67.4	4.5	22.6
0.6	912	91	41	27	93.7	97.1	68.9	4.3	22.9
0.62	910	92	40	29	93.6	96.9	69.7	4.2	24
0.64	905	92	40	34	93.1	96.4	69.7	4.2	27
0.66	904	93	39	35	93.1	96.3	70.5	4.1	27.3
0.68	903	94	38	36	93.1	96.2	71.2	4	27.7
0.7	899	96	36	40	92.9	95.7	72.7	3.9	29.4
0.72	897	96	36	42	92.7	95.5	72.7	3.9	30.4
0.74	891	96	36	48	92.2	94.9	72.7	3.9	33.3
0.76	889	98	34	50	92.2	94.7	74.2	3.7	33.8
0.78	881	100	32	58	91.6	93.8	75.8	3.5	36.7
0.8	874	102	30	65	91.1	93.1	77.3	3.3	38.9
0.82	867	102	30	72	90.5	92.3	77.3	3.3	41.4
0.84	855	106	26	84	89.7	91.1	80.3	3	44.2
0.86	844	109	23	95	89	89.9	82.6	2.7	46.6
0.88	829	111	21	110	87.8	88.3	84.1	2.5	49.8
0.9	804	116	16	135	85.9	85.6	87.9	2	53.8
0.92	777	121	11	162	83.8	82.7	91.7	1.4	57.2
0.94	725	123	9	214	79.2	77.2	93.2	1.2	63.5
0.96	644	126	6	295	71.9	68.6	95.5	0.9	70.1
0.98	550	132	0	389	63.7	58.6	100	0	74.7
1	0	132	0	939	12.3	0	100	.	87.7

Table 5: *Odds Ratio Estimates*

Effect	Point Estimate	90% Wald Confidence Limits	
typur 0 vs 9	<0.001	<0.001	>999.999
typur 1 vs 9	0.004	<0.001	>999.999
typur 3 vs 9	3.128	<0.001	>999.999
typur 5 vs 9	0.873	<0.001	>999.999
typur 6 vs 9	0.511	<0.001	>999.999
typur 7 vs 9	1.114	<0.001	>999.999
typur 8 vs 9	8.665	<0.001	>999.999
married 0 vs 1	0.483	0.29	0.805
gdlin 0 vs 666	<0.001	<0.001	>999.999
gdlin 1 vs 666	<0.001	<0.001	>999.999
pubrec 0 vs 1	4.051	1.902	8.626
obrat	0.971	0.949	0.994
fixadj 0 vs 1	0.394	0.238	0.65
prop 1 vs 3	3.575	1.686	7.579
prop 2 vs 3	6.376	3.225	12.604
inson 0 vs 1	119.516	18.724	762.871
gift 0 vs 1	2.055	1.086	3.887
unver 0 vs 1	11.609	4.704	28.653
old 0 vs 1	2.386	1.446	3.935
vr 0 vs 1	1.95	1.185	3.211

Table 6: *Residual Chi-Square Test*

Chi-Square	DF	Pr > ChiSq
28.1359	36	0.8223

Table 7: *R-Squared Statistics*

R-Square	0.4
Max-rescaled R-Square	0.7

Table 8: *Model Fit Statistics*

Criterion	*Intercept Only*	*Intercept & Covariates*
AIC	801.714	380.074
SC	806.691	479.601
-2 Log L	799.714	340.074

Table 9: *Outcome of MLR*

Decile or Segment	*Prediction for segment*	*Approvals*	*Prediction Vs Actual (Approvals)*	*Rejects*	*Prediction Vs Actual (Rejects)*
1	**Most Likely to be approved Prob>=0.9**	**578**	**83% Correct**	**11**	**11% False Negative**
2		66		9	
3		18		2	
4		8		4	
5		4		5	
6		4		8	
7		2		4	
8		3		7	
9		6		13	
10	**Most Likely to be rejected Prob <0.1**	**4**	**0.1 % False Positive**	**38**	**38% Correct**
		693		**101**	

Table 10: *PNN Predictions-Basic Statistics*

Test data set	*# Applications*	*%*
Incorrect Predictions	**15**	**12%**
Correct Predictions	**110**	**88%**
Total # applications	**125**	

Table 11: *Performance of Trained Network on Test Data Set.*

Performance	*action*	*action Output*
MSE	0.0933149	0.01750095
NMSE	0.8360351	0.64698811
MAE	0.1656041	0.09036844
Min Abs Error	0.0010625	0.00106248
Max Abs Error	0.9945641	0.47659806
r	0.4456898	0.70783952
Percent Correct	12.5	99.0825688

Table 12: *Most Effective Predictors using both Techniques*

Logistic Regression		*Neural Network*	
gdlin	credit history meets guidelines	typur	type of purchaser of loan
typur	type of purchaser of loan	other	other financing, $1000s
inson	PMI approved	chist	=0 if accnts deliq. >= 60 days
prop	type of property	cons	credit history on consumer stuf
unver	unverifiable info	obrat	other oblgs, % total inc
fixadj	fixed or adjustable rate?	mortperf	no late mort. payments
pubrec	=1 if filed bankruptcy	cototinc	coapp total monthly income
vr	=1 if tract vac rte > MSA med	apr	appraised value
obrat	other oblgs, % total inc	prop	type of property
old	=1 if applic age > MSA median		
married	=1 if applicant married		
gift	gift as down payment		

Figure 1: *Lorenz Curve-Measures of Discrimination. Scorecard performance can be judged on the level of discrimination. Gini (or ROC) with c- value of 0.957 shows that the model is good predictor model.*

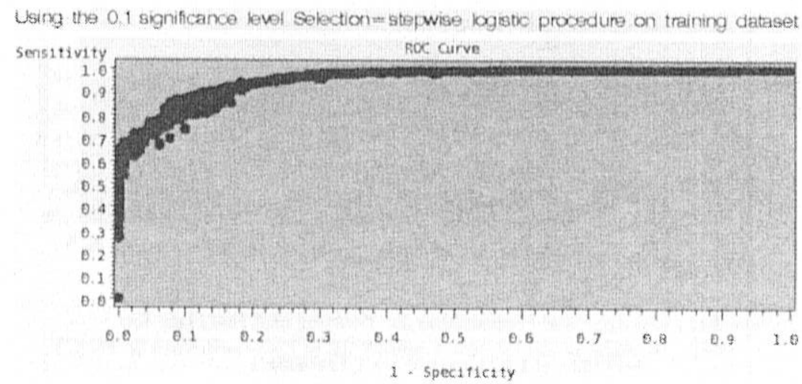


Figure 2: *Common Factor Analysis*

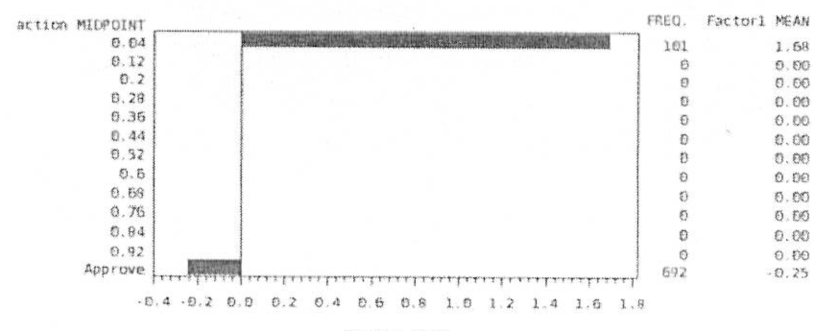


Figure 3.a: *Probabilities of Action (Approval or Reject) For 'Training' Data Set*

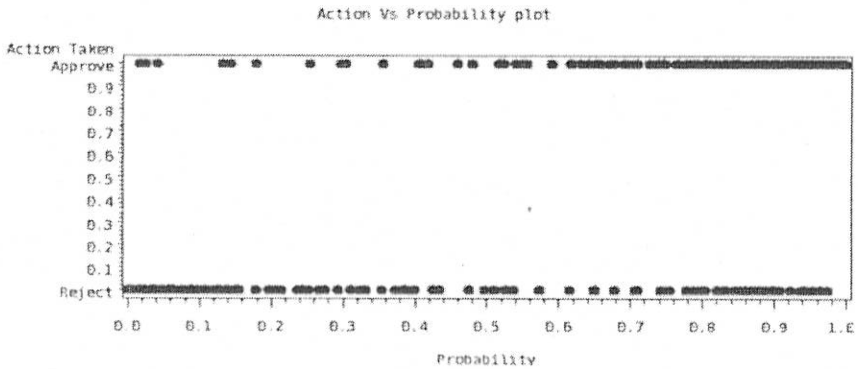


Figure 3.b: *Probabilities of Action (Approval or Reject) For 'Training' Data Set*

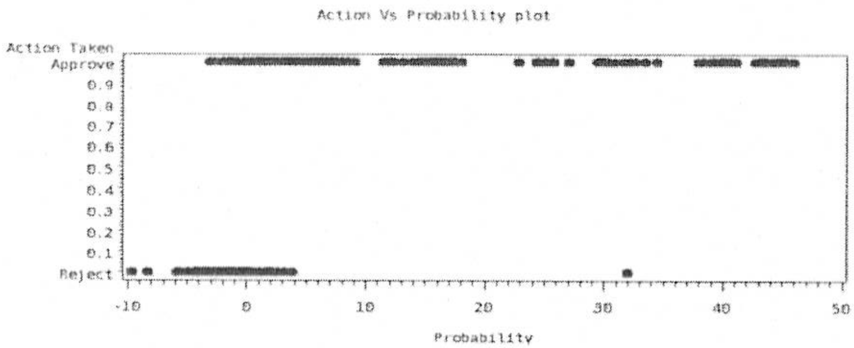


Figure 3.c: *Loan Approval Probabilities for Training and Test Data Sets.*

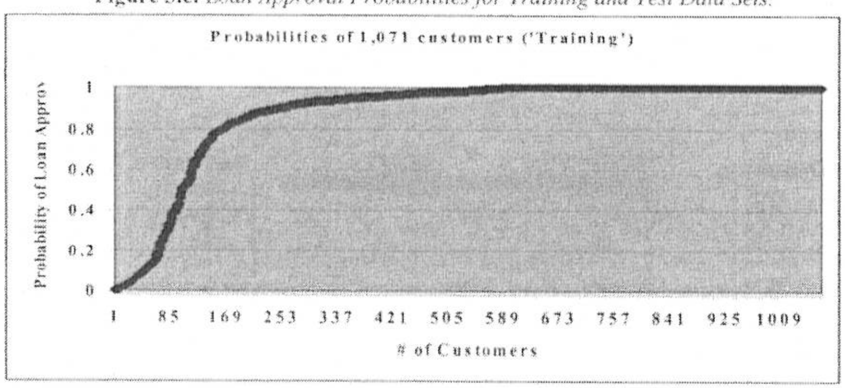


Figure 3.d: *Loan Approval Probabilities for Training and Test Data Sets*

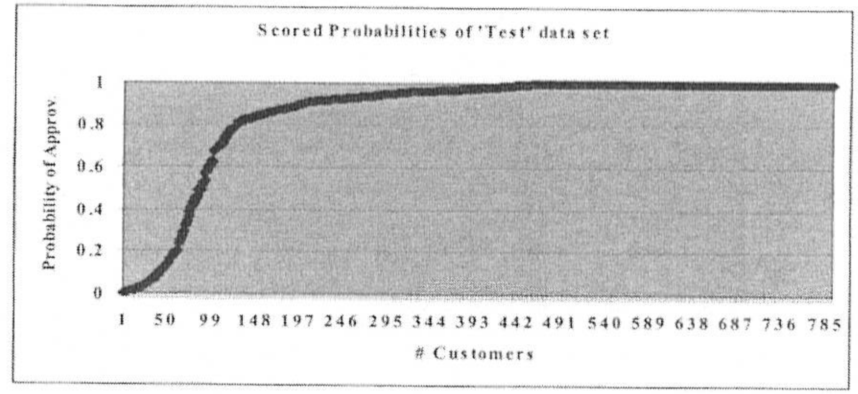


Figure 4: *Source Ward Systems Group, Inc. NeuroShell®Predictor.*

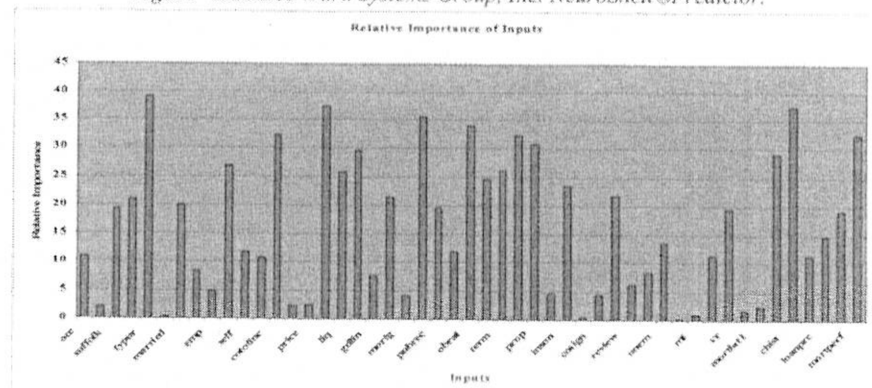


Figure 5: *PNN-Sensitivity of All 9 Input Variables about Mean of Training Data Set.*

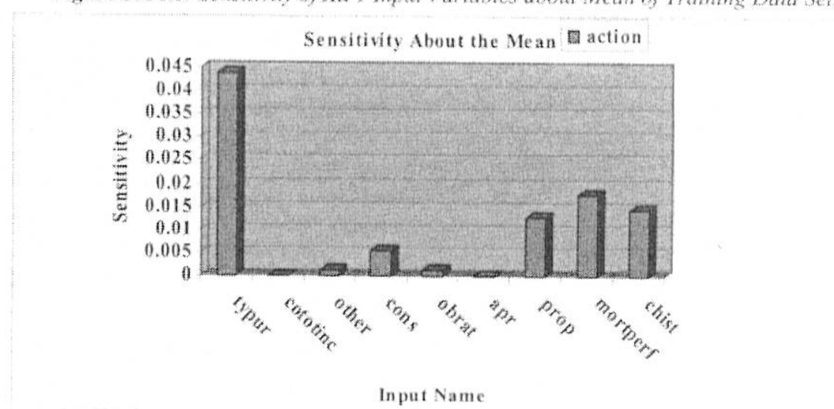


Figure 6: *MSE Over 1,000 Epochs.*

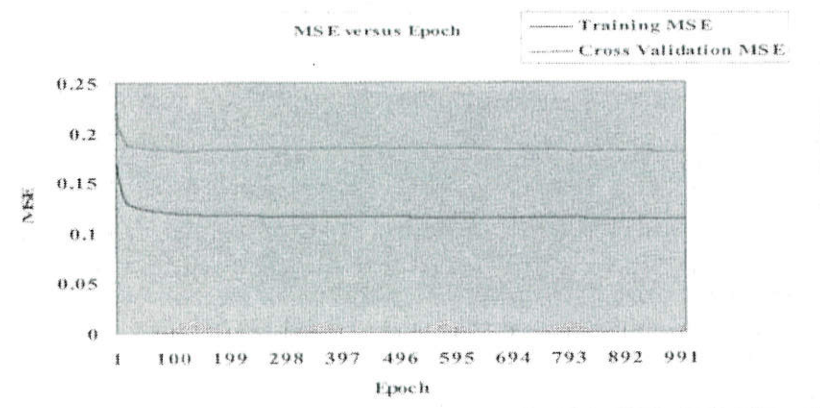


Figure 7: *Graphical Comparison of Actual and Predicted Outputs.*