

An Analysis of Credit Scoring for Agricultural Loans in Thailand

¹Visit Limsombunchai, ¹Christopher Gan and ²Minsoo Lee

¹Commerce Division, P.O. Box 84, Lincoln University, Canterbury, New Zealand 8150

²School of Business and Management, P.O. Box 26666,
American University of Sharjah, Sharjah United Arab Emirates

Abstract: Loan contract performance determines the profitability and stability of the financial institutions and screening the loan applications is a key process in minimizing credit risk. Before making any credit decisions, credit analysis (the assessment of the financial history and financial backgrounds of the borrowers) should be completed as part of the screening process. A good credit risk assessment assists financial institutions on loan pricing, determining amount of credit, credit risk management, reduction of default risk and increase in debt repayment. The purpose of this study is to estimate a credit scoring model for the agricultural loans in Thailand. The logistic regression and Artificial Neural Networks (ANN) are used to construct the credit scoring models and to predict the borrower's creditworthiness and default risk. The results of the logistic regression confirm the importance of total asset value, capital turnover ratio (efficiency) and the duration of a bank - borrower relationship as important factors in determining the creditworthiness of the borrowers. The results also show that a higher value of assets implies a higher credit worthiness and a higher probability of a good loan. However, the negative signs found on both capital turnover ratio and the duration of bank-borrower relationship, which contradict with the hypothesized signs, suggest that the borrower who has a long relationship with the bank and who has a higher gross income to total assets has a higher probability to default on debt repayment.

Key words: Agricultural loan, credit scoring, neural networks, logistic regression

INTRODUCTION

Loan contracts performance determines the profitability and stability of the financial institutions and screening the loan applications is a key process in minimizing credit risk. Before making any credit decisions, credit analysis (the assessment of the financial history and financial backgrounds of the borrowers) should be completed as part of the screening process. Good borrowers with low credit risk would be granted a loan, while a high risk borrower would be denied. A good credit risk assessment assists financial institutions on loan pricing, determining amount of credit, credit risk management, reduction of default risk and increase in debt repayment.

Credit analysis is the primary method in reducing the credit risk on a loan request. This includes determining the financial strength of the borrowers, estimating the probability of default and reducing the risk of nonpayment to an acceptable level^[1]. In general, credit evaluations are based on the loan officer's subjective assessment (or judgmental assessment technique). However, this technique seems to be inefficient, inconsistent and non-uniform^[2,3].

A major evolution in the credit evaluation practices has been the risk assessment (or credit scoring) of borrowers based on sophisticated statistical analysis of the borrower's financial data and other information related to creditworthiness. Credit scoring models have the potential in reducing the variability of credit decisions and adding efficiencies to the credit risk assessment process. Furthermore, the models not only assist financial institutions in loan approval, but also on loan pricing, loan monitoring, determining amount of credit, credit risk management and assessment of loan portfolio risks^[4].

Credit scoring has been broadly applied in consumer lending, especially in credit cards and it is becoming more commonly used in mortgage lending. Credit scoring has not been widely applied in business lending because business loans substantially differ across conventional borrowers and make it more difficult to construct an accurate scoring method. However, the flexibility of the statistical models and the computing technology has made such scoring method possible. Several financial institutions are currently using the credit scoring model to assess loan applications, which is a cost effective credit management tool^[5].

Corresponding Author: Visit Limsombunchai, Commerce Division, P.O Box 84, Lincoln University, Canterbury, New Zealand 8150, Tel: 64-3-325-2811, Fax: 64-3-325-3847

The overall idea of credit scoring model is quite straightforward. A large historical loan sample of the similar loan type is divided into two categories, good loans and bad loans. Based on statistical probabilities, the combination of borrowers' characteristics differentiating "good" from "bad" loans generate a score (or probability) serving as an estimate of the risk level of each new loan^[2] when the lenders decide whether to make the loans or not.

Several statistical methods have been used to estimate credit scoring models, including discriminant analysis^[6-8], linear probability models^[7,9], logit models^[7,8,10] and probit models^[7,11]. The logit model has dominated the literature and has been widely used because of its simplicity. Recently, there has been an increase in the use of the artificial neural networks (ANN) in the lending decision process^[8,12,13].

The purpose of this study is to estimate a credit scoring model (lending decision) for the agricultural loans in Thailand. A special class of artificial neural networks called "probabilistic neural network (PNN)" is employed to estimate the credit scoring model together with the logit model and a widely used artificial neural network called "multi-layer feed-forward neural network (MLFN)". The study also empirically compares the predictive power among the three different estimation methods.

Factors used in lending decision models: The major factors used in lending decision models include borrowers' liquidity (i.e. current ratio, quick ratio and net working capital), profitability (i.e. return on assets and return on equity), solvency (i.e. leverage ratio and debt-to-equity ratio), efficiency (i.e. gross ratio and capital turnover ratio) and repayment capacity (i.e. Interest expense ratio, interest coverage ratio and debt repayment ratio).

The variables can be easily calculated from a borrower's financial statement. Thus, lenders always use these financial criteria in combination with other factors, such as the borrower's personal attributes, enterprise type, region and etc., In the credit decision model. Since it has been found that the relationship between bank (lender) and the borrower has an influence on the availability of credit and the cost of credit^[14-16], the lender-borrower relationship should have an influence on the lending decision. Therefore, the relationship indicators will be included in the credit scoring model to further enhance the analysis.

DATA AND METHODOLOGY

The data in this study are obtained from the Bank of Agriculture and Agricultural Cooperative (BAAC), Thailand. BAAC is considered a major lender in the Thailand agricultural sector with a high significant share in the agricultural financing market (more than 55 percent of the total loan in 2003). The credit files were

retrieved from the "Credit BPR" (Credit Business Process Reengineering) database on June 2004. During the period of 2001 to 2003, a total of 16,560 agricultural loans was made available. The data set comprises of 14,383 good loans (GL) and 2,177 bad (or default) loans (BL). All loans are under the normal loan scheme (excluding the government loans for specific projects). Unfortunately, there is no information about borrowers' current assets, current liabilities and debt repayment available on the database. As a result, the borrower's liquidity and repayment capacity cannot be estimated.

Logistic model: We assume that the probability of a good loan follows the logistic distribution and is a function of the borrower characteristics, credit risk proxies, relationship indicators and dummy variables. The credit scoring model for the agricultural loans in Thailand can be written as follows:

Lending decision = f (Borrower characteristics, Credit risk proxies, Relationship indicators, Dummy variables) (1)

Where, Lending decision = 1 if loan is paid (good loan); 0 if loan is default (bad loan)

Borrower characteristics include:

Assets (+) = total asset value (in Thai Baht),
Age (+) = age of the borrower (in years),
Education (+) = 0 if the qualification of the borrower is primary school or lower; 1 otherwise;

Credit risk proxies include:

Collateral (+) = value of collateral (in Thai Baht),
Return on assets (+) = net return / total assets,
Leverage ratio (-) = total liabilities / total assets,
Capital turnover ratio (+) = gross income / total assets;

Relationship indicators include:

Borrowing from others (-) = 1 if the borrower has an outstanding debt with other financial sources; 0 if the borrower has a loan from BAAC only,
Duration (+) = the duration of bank-borrower relationship prior to the credit decision (in years);

Dummy variables include: Province (Province 1 in 17), Farm type (Horticulture, Orchard/Vegetable, Livestock/Aquaculture and others), Loan type (Cash credit loan, Short-term loan, Medium-term loan and Long-term loan), Loan size (Small loan, Medium loan and Large loan) and Lending year (2001 to 2003)

Priori hypotheses are indicated by (+) or (-) in the above specification. For example, assets, age, education, collateral, return on assets, capital turnover ratio and duration are positively related to the probability of a good loan. On the other hand, leverage ratio and borrowing from others are negatively related to the probability of a good loan.

Dummy variables such as province, farm type, loan type, loan size and lending year are included to describe the systematic effects relating to the type of borrowers and the type of contracts and are hypothesized to influence the borrower's credit risk and the probability of a good loan. For example, borrowers who have a cash crop (horticulture) as the major production would require a smaller amount of credit than the other farm types, and the contract term for the cash crop production is a short-term contract. Thus, this group of borrowers would have a higher probability to obtain a loan than the others. This is because the short-term loan is less risky than medium-term or long-term loans and the lending risk is relatively low. In contrast, if the major production of the borrowers is either orchards or livestock, which requires a larger long-term loan, they would be expected to have higher credit risks and a higher probability to default.

MLFN model: The ANN model, inspired by the structure of the nerve cells in the brain, can be represented as a massive parallel interconnection of many simple computational units interacting across weighted connections^[17]. Each computational unit (or neuron or node) consists of a set of input connections that receive signals from other computational units, a set of weights for input connection and a transfer function (Fig. 1). The output of the computation unit (node j), U_j , is the result of applying a transfer function F_j to the summation of all signals from each connection (X_i) times the value of the connection weight between node j and connection I (W_{ij}) (Eq. 2):

$$U_j = F_j \left(\sum W_{ij} X_i \right) \quad (2)$$

where, U_j is output for node j and F_j is a transfer function with different functional forms: linear functions, linear threshold functions, step functions, sigmoid functions or Gaussian function^[18].

The multi-layer feed-forward neural network (MLFN) computational units are grouped into 3 main layers-the first layer is the input layer, the last layer is the output layer and the layer(s) in between is called the hidden layer(s). Figure 2 shows the structure of the multi-layer feed-forward neural network with one hidden layer. Since the output of one layer is an input to the following layer, the output of the network (Z) can be algebraically exhibited as follows:

$$Z = F \left(\sum_{j=1}^j W_j^{(2)} \cdot F_j \left(\sum_{i=1}^i W_{ij}^{(1)} X_i \right) \right) \quad (3)$$

where, Z is the output of the network, F is the transfer function in the output node, $W_{ij}^{(1)}$ and $W_j^{(2)}$ are connection weights from input layer (node i) to hidden layer (node j) and from hidden layer (node j) to output layer, respectively.

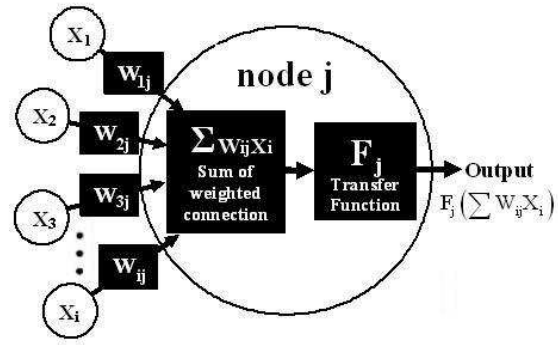


Fig. 1: Structure of a computational unit (node j), (source: modified from James and Carol^[18])

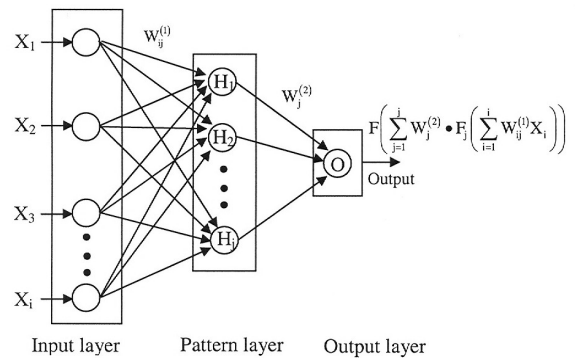


Fig. 2: MLFN structure with one hidden layer (source: modified from West *et al.*^[19] and Gradojevic and Yang^[20])

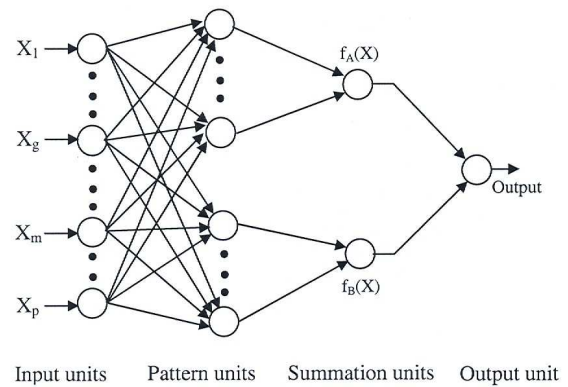


Fig. 3: The PNN architecture (source: modified from Specht^[21])

The calculation of the neural network weights is known as training process. The process starts by randomly initializing connection weights and introduces a set of data inputs and actual outputs to the network. The network then calculates the network output and compares it to the actual output and calculates the error. In an attempt to improve the overall predictive accuracy and to minimize the network total

mean squared error, the network adjusts the connection weights by propagating the error backward through the network to determine how to best update the interconnection weights between individual neurons.

PNN model: The PNN model proposed by Specht^[21] is basically a classification network. Its general structure consists of 4 layers - an input layer, a pattern layer, a summation layer and an output layer (Fig. 3). The PNN model is conceptually based on the Bayesian classifier statistical principle. According to the Bayesian classification theorem, X will be classified into class A, if the inequality in equation 4 holds:

$$h_A c_A f_A(X) > h_B c_B f_B(X) \quad (4)$$

where, X is the input vector to be classified, h_A and h_B are prior probabilities for class A and B, c_A and c_B are costs of misclassification for class A and B, $f_A(X)$ and $f_B(X)$ are probabilities of X given the density function of class A and B, respectively^[22].

To determine the class, the probability density function is estimated by a non-parametric estimation method developed by Parzen^[23] and further extended by Cacoulos^[24]. The joint probability density function of a set of the p variables can be expressed as follows:

$$f_A(X) = \frac{1}{(2\pi)^{p/2} \sigma^p n_A} \sum_{j=1}^{n_A} e^{-\frac{(X-Y_{Aj})^2}{2\sigma^2}} \quad (5)$$

where, p is the number of variables in the input vector X, n_A is the number of training samples which belongs to class A, Y_{Aj} is the j^{th} training sample in class A and σ is a smoothing parameter^[25].

The working principle of the PNN model starts from the input layer, where the inputs are distributed to the pattern units. Then the pattern unit, which is required for every training pattern, is used to memorize each training sample and estimate the contribution of a particular pattern to the probability density function. The summation layer comprises of a group of computational units with the number equal to the total number of classes. Each summation unit that delicate to a single class sums the pattern layer units corresponding to that summation unit's class. Finally, the output neuron(s), which is a threshold discriminator, chooses the class with the largest response to the inputs^[26,27].

NeuroShell2 package is used to construct the artificial neural network models, while the logistic model is estimated by the maximum likelihood method used in the LIMDEP software. To examine the predictive power of the models, the out-of-sample forecasting technique is applied. The sample is randomly divided into two sub-samples: an estimation sample and a forecasting sample. The estimation sample and the forecast sample contain 80 and 20

percent of the total observations, consecutively. All models are re-estimated by using only the estimation samples and the out-of-sample forecasting are conducted over the forecasting samples. To evaluate the forecast accuracy of the model, the classification rates and the expected misclassification loss of each model are computed and compared.

RESULTS AND DISCUSSION

The results of the logistic credit scoring models are shown in Table 1. In general, the model I and model II (without and with duration, respectively) fit the data quite well. The chi-square statistics fail to accept the null hypothesis that the parameter estimates for the models are equal to zero. Both models correctly predict the lending decision at 87.19 and 85.30 percent, respectively. However, the model I and II exhibit 93.98 and 90.70 percent of Type I error (wrongly reject H_0 or accepting a bad loan (BL) as a good loan (GL)), respectively. Although model I have a higher overall percentage correct, model II can predict the bad loan (BL) better than the model I.

In model I, the estimated coefficients of assets, education, leverage ratio and capital turnover ratio are found to be significant at the 5 percent level (Table 1). The probability of a good loan would increase if the borrower has larger assets and more than a primary education. On the other hand, the probability of a good loan deteriorates as the borrower improves his or her leverage (solvency) and capital turnover ratios (efficiency). This contradicts the hypothesis on capital turnover ratio, which shows that the borrower who has a higher gross income to total assets has a higher probability to default on debt repayment. It seems when borrowers earn more they prefer to spend the extra earned income on other activities rather than repaying their debt.

When the duration is included into the model (model II), there are only 3,965 observations that can be used to estimate the model. This is because there is no available information to estimate the duration for all the samples, due to recent changes in the BAAC's database policy. The estimated results show that assets and capital turnover ratio are significant at the 5 percent level, while education and leverage ratio are insignificant. Furthermore, the estimated coefficient on a capital turnover ratio is negative, which is consistent with the estimated result in the model I. However, the relationship between duration and lending decision contradicts the postulated hypothesis. The estimated coefficient is negative and significant at the 5 percent level. Thus, it suggests that the borrower who has a longer relationship with the bank has a higher probability to default on debt repayment and bank should cautiously deal with this group of borrowers.

Table 1: Logistic models

| Variable ^{1/, 2/} | Model I | | Model II | | | |
|---|-------------|-------------------------------|-------------|-------------------------------|-------|---------|
| | Coefficient | Marginal Effect ^{3/} | Coefficient | Marginal Effect ^{3/} | | |
| Log(Assets) | 0.3197* | 0.0289 | 0.3719* | 0.0387 | | |
| Age | -0.0009 | -0.0001 | -0.0016 | -0.0002 | | |
| Education | 0.1686* | 0.0161 | 0.1769 | 0.0190 | | |
| Log(Collateral) | -0.0339 | -0.0030 | -0.0689 | -0.0072 | | |
| Return on assets | 0.0383 | 0.0030 | 0.005 | 0.0005 | | |
| Leverage ratio | -0.9629* | -0.0874 | -0.8326 | -0.0868 | | |
| Capital turnover ratio | -0.0634* | -0.0060 | -0.0596* | -0.0062 | | |
| Borrowing from others | 0.1081 | 0.0095 | 0.0329 | 0.0034 | | |
| Duration | | | -0.1915* | -0.0199 | | |
| Province, Farm type, Loan type, Loan size, Lending year dummies ^{4/} | yes | yes | yes | yes | | |
| Constant | yes | yes | yes | yes | | |
| No. Of observation | | 16,560 | | 3,965 | | |
| LR statistic (χ^2) | | 1,446.85* | | 398.97* | | |
| Degree of freedom | | 34 | | 35 | | |
| Log likelihood | | -5,720.45 | | -1,489.09 | | |
| McFadden R ² | | 0.1123 | | 0.1182 | | |
| Prediction classification ^{5/} | BL | GL | Overall | BL | GL | Overall |
| % Correct | 6.02 | 99.48 | 87.19 | 9.30 | 98.90 | 85.30 |
| % Incorrect | 93.98 | 0.52 | 12.81 | 90.70 | 1.10 | 14.70 |

Note: 1/ Dependent variable is lending decision.
 2/ Maximize using logistic likelihood function and quasi-maximum likelihood (QML) standard errors and covariance.
 3/ Marginal effect is at the mean value. For dummy variable, marginal effect is P|1 - P|0.
 4/ To avoid the singularity problem, a dummy variable is dropped from each group.
 5/ Cut-off point = 0.50
 *Represent 5% significant level.

Table 2: Prediction classification of ANN models (in-sample forecast)

| | Model I | | | Model II | | |
|-------------|---------|-------|---------|----------|-------|---------|
| | BL | GL | Overall | BL | GL | Overall |
| MLFN | | | | | | |
| % Correct | 14.47 | 98.89 | 87.80 | 8.47 | 99.29 | 85.50 |
| % Incorrect | 85.53 | 1.11 | 12.20 | 91.53 | 0.71 | 14.50 |
| PNN | | | | | | |
| % Correct | 87.51 | 98.92 | 97.42 | 88.37 | 97.98 | 96.52 |
| % Incorrect | 12.49 | 1.08 | 2.58 | 11.63 | 2.02 | 3.48 |

Note: Cut-off point = 0.50

Table 3: Out-of-sample prediction classification

| | Model I | | | Model II | | |
|-------------|---------|-------|---------|----------|-------|---------|
| | BL | GL | Overall | BL | GL | Overall |
| LOGIT | | | | | | |
| % Correct | 4.05 | 99.41 | 86.62 | 5.13 | 99.11 | 85.25 |
| % Incorrect | 95.95 | 0.59 | 13.38 | 94.87 | 0.89 | 14.75 |
| MLFN | | | | | | |
| % Correct | 10.59 | 99.13 | 87.26 | 4.27 | 99.26 | 85.25 |
| % Incorrect | 89.41 | 0.87 | 12.74 | 95.73 | 0.74 | 14.75 |
| PNN | | | | | | |
| % Correct | 11.04 | 99.23 | 87.41 | 40.17 | 91.57 | 83.98 |
| % Incorrect | 88.96 | 0.77 | 12.59 | 59.83 | 8.43 | 16.02 |

Note: Cut-off point = 0.50

The marginal effects in Table 1 represent a quantitative change in the conditional probability that results from the change in the independent variable. For example, a one unit increment in the borrower's total assets value and capital turnover ratio would increase and decrease the probability of a good loan by 0.0387 and 0.0062, respectively (see model II). Furthermore, the marginal effect on duration shows when the

relationship between the bank and the borrower increases by 1 year, the probability of a good loan would decrease by 0.0199 on the average.

The estimated coefficients of the province, farm type, loan type, loan size and lending year dummy variables are not presented here, but the estimated results show that horticultural production, short-term loan and small borrowing are less risky and have a

lower credit risk than others. Therefore, the probability to default on loan repayment of the borrower (farmer) in this group is relatively lower than the others. Furthermore, the estimated coefficients of provinces show that the borrower's credit risk differs according to the residential province.

Since the ANN model is usually nonlinear and their training process is always regarded as a black-box, it is very difficult to write out the algebraic relationship between a dependent variable and an independent variable compared to the logistic models. Furthermore, the learned outputs, connection weights or coefficients, cannot be interpreted and tested. Therefore, only the classification results of the models are presented in Table 2.

The classification results in Table 2 show that the PNN models (both models I and II) exhibit a superior ability to learn and memorize the patterns corresponding to the borrower's default risk. The overall percentage correct of the PNN for both models I and II are 97.42 and 96.52 percent, respectively. Thus, the PNN models offer a better classification results than the logistic models, whereas the MLFN models yield almost the same level of correctness as the logistic models. However, the results do not provide strong and conclusive evidence of superiority in term of prediction capability among the models, as shown by the in-sample results.

The classification rates on the out-of-sample prediction for the logistic, MLFN and PNN models are shown in Table 3. The results show that the prediction accuracy of the three models is similar to each others in the model I, but in model II, the logistic and MLFN models are slightly better than the PNN model. However, a closer examination indicates that the logistic model can predict well only on the good loan (GL). The Type I error rate shows that the logistic model is unable to predict the bad loan (BL), as it has more than 90 percent of Type I error. In contrast, the Type I error of the PNN model is smaller than the logistic and MLFN models, especially when duration is introduced into the lending decision model (model II).

It is generally accepted that the misclassification cost of Type I error is more costly than Type II error. For Type I error, the lender may lose not only the principal but also the interest on the principal. On the other hand, for Type II error, the lender loses only the interest and expected profit from the loan. Therefore, the overall percentage correct may be misleading in this case, as it is calculated under the assumption that the misclassification costs of both types of errors are identical. Thus, to interpret the model performance in a meaningful way, the misclassification costs of both types of errors must be differentiated and taken into the account. The expected loss of misclassification on the

out-of-sample forecasting must be estimated. The lending decision model that offers the smallest expected loss is considered as the most preferable model.

According to Koh^[28], the expected misclassification loss of the model can be calculated as follows:

$$EL = (PB)(PI)(CI) + (PG)(PII)(CII) \tag{6}$$

Where:

- EL = Expected loss of misclassification
- PB = Prior probability of being bad loan
- PI = Conditional probability of Type I error
- CI = Misclassification cost of Type I error
- PG = Prior probability of being good loan
- PII = Conditional probability of Type II error
- CII = Misclassification cost of Type II error

Since the prior probabilities of PB and PG are unobserved, they are estimated by dividing the total number of bad and good loans by the total number of samples, respectively. Since the consequences of incorrect classification are intangible and immeasurable (such as loss of existing and potential clients, loss of depositor's trustworthy, etc.), It is not easy to quantify the misclassification costs of Type I and Type II errors (CI and CII). Therefore, the relative misclassification costs of Type I and Type II errors are used. The relative cost ratios are assumed to vary accordingly from 1:1, 2:1, 3:1, 4:1 and 5:1, with the relatively higher misclassification cost on Type I error where a bad loan is classified as a good loan.

Table 4 summarizes the models expected misclassification loss on out-of-sample forecasting at the different relative cost ratios. The PNN model without duration (model I) has the lowest expected loss when the relative cost ratio is 1:1. Although the PNN model with a duration (model II) has lower overall percentage correct than the logistic and MLFN models (model II) on out-of-sample forecasting, when the cost ratio is 2:1 or higher, the PNN model becomes the top performer since it has the lowest expected loss. Therefore, the PNN model can be considered as the superior model in predicting the lending decision.

Table 4: Expected loss (out-of-sample prediction)

| CI : CII | 1:1 | 2:1 | 3:1 | 4:1 | 5:1 |
|----------|---------------|---------------|---------------|---------------|---------------|
| LOGIT | | | | | |
| Model I | 0.1313 | 0.2574 | 0.3835 | 0.5097 | 0.6358 |
| Model II | 0.1324 | 0.2571 | 0.3819 | 0.5066 | 0.6313 |
| MLFN | | | | | |
| Model I | 0.1251 | 0.2427 | 0.3602 | 0.4778 | 0.5953 |
| Model II | 0.1323 | 0.2581 | 0.3840 | 0.5098 | 0.6356 |
| PNN | | | | | |
| Model I | 0.1236 | 0.2406 | 0.3575 | 0.4745 | 0.5914 |
| Model II | 0.1519 | 0.2305 | 0.3092 | 0.3878 | 0.4665 |

Note: Bold and italic indicate the minimum expected loss

CONCLUSION

The estimated results of the logistic credit scoring model show the significance of total asset value, capital turnover ratio and duration in determining the probability of a good loan. The results show that a higher value of assets implies a higher creditworthiness and a higher probability of a good loan. However, the negative signs in both capital turnover ratio and duration, which contradict the hypothesized signs, suggest that the borrower who has a longer relationship with the bank and who has a higher gross income to total assets has a higher probability to default on debt repayment.

The overall prediction accuracy of the logistic credit scoring models is above 85% on both in-sample and out-of-sample forecast and is higher than the PNN model II on out-of-sample forecast. In most cases, the logistic models' performances are quite similar to the MLFN model. Therefore, in terms of precision, the ANN model might not necessarily predict the borrower's creditworthiness and default risk better than the logistic regression model. However, most of the ANN models can detect Type I error much better than the logistic regression models. Since it is generally accepted that the costs of classifying a bad loan as a good loan (Type I error) are significantly greater than the costs of misclassifying a good loan as a bad loan (Type II error), the overall prediction accuracy is not completely reliable, since it ignores the relative cost difference between Type I and Type II errors. Thus, when the expected loss of misclassification is computed and compared, the results indicate that the misclassification cost of the PNN model is the best credit scoring model with the lowest misclassification costs. In summary, the empirical results of this study support the use of the PNN model in classifying and screening agricultural loan applications in Thailand.

REFERENCES

1. Plata, V. and G.N. Nartea, 1998. Credit analysis procedures of rural lenders in Canterbury. Paper Presented at the 5th Ann. Conf. New Zealand Agril. and Resource Econ. Soc. (Inc.).
2. Crook, J.N., 1996. Credit scoring: An overview. Working Paper no. 96/13, Department of Business Studies. The University of Edinburgh.
3. Glassman, C.A. and H.M. Wilkins, 1997. Credit scoring: probabilities and pitfalls. *J. Retail Banking Service*, 19: 53-56.
4. Turvey, C.G. and R. Brown, 1990. Credit scores for a federal lending institution: The case of Canada's farm credit corporation. *Agril. Finance Rev.*, 50: 47-57.
5. Mester, L.J., 1997. What's the point of credit scoring? *Business Review*, Federal Reserve Bank of Philadelphia, Sep. /Oct., pp: 3-16.
6. Dunn, D.J. and T.L. Frey, 1976. Discriminant analysis of loans for cash grain farms. *Agril. Finance Rev.*, 36: 60-66.
7. Turvey, C.G., 1991. Credit scores for agricultural loans: A review with application. *Agril. Finance Rev.*, 51: 43-54.
8. Altman, E.I., M. Glancario and F. Varetto, 1994. Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (The Italian experience). *J. Banking and Finance*, 18: 505-529.
9. Barney, D.K., O.F. Graves and J.D. Johnson, 1999. The farmers home administration and farm debt failure prediction. *J. Accounting and Public Policy*, 18: 99-139.
10. Turvey, C.G. and A. Weersink, 1997. Credit risk and the demand for agricultural loans. *Can. J. Agril. Econ.*, 4: 201-217.
11. Lufburrow, J., P.J. Barry and B.L. Dixon, 1984. Credit scores for farm loan pricing. *Agril. Finance Rev.*, 44: 8-4.
12. Lee, T.H. and S.C. Jung, 1999. Forecasting credit worthiness: Logistic vs. Artificial neural net. *J. Business Forecasting*, Winter: 28-30.
13. Wu, C. and X.M. Wang, 2000. A neural network approach for analyzing small business lending decisions. *Rev. Quantitative Finance and Accounting*, 15: 259-276.
14. Petersen, M.A. and R.G. Rajan, 1994. The benefit of lending relationships: Evidence from small business data. *J. Finance*, 49: 3-38.
15. Berger, A.N. and G.F. Udell, 1995. Relationship lending and lines of credit in small firm finance. *J. Business*, 68: 351-381.
16. Harhoff, D. and T. Korting, 1998. Lending relationships in Germany- empirical evidence from survey data. *J. Banking and Finance*, 22: 1317-1353.
17. Venugopal, V. and W. Baets, 1994. Neural networks and statistical techniques in marketing research: A conceptual comparison. *Marketing Intelligence & Planning*, 12: 30-38.
18. James, R.C. and E.B. Carol, 2000. Artificial neural networks in accounting and finance: Modeling issues. *Intl. J. Intelligent Systems in Accounting, Finance and Management*, 9: 119-144.
19. West, P.M., P.L. Brockett and L.L. Golden, 1997. A comparative analysis of neural networks and statistical methods for predicting consumer choice. *Market. Sci.*, 16: 370-391.

20. Gradojevic, N. and J. Yang, 2000. The application of artificial neural networks to exchange rate forecasting: The role of market microstructure variables. Bank of Canada, Working Paper, 2000–23.
21. Specht, D.F., 1990. Probabilistic neural networks. *Neural Networks*, 3: 109-118.
22. Albanis, G.T. and R.A. Batchelor, 1999. Using probabilistic neural networks and rule induction techniques to predict long-term bond ratings, in M. Torres (Ed.), *Proc. 5th Ann. Conf. Information Systems, Analysis and Synthesis*, Orlando: IIS.
23. Parzen, E., 1962. An estimation of a probability density function and mode. *Ann. Math. Stat.*, 33: 1065-1076.
24. Cacoullos, T., 1966. Estimation of a multivariate density. *Ann. Institute of Stat. Math. (Tokyo)*, 18: 179-189.
25. Chen, A., M.T. Leung and H. Daouk, 2003. Application of neural networks for an emerging financial market: Forecasting and trading the Taiwan Stock Index. *Comp. Operations Res.*, 30: 901-923.
26. Etheridge, H.L. and R.S. Sriram, 1997. A comparison of the relative costs of financial distress models: Artificial neural networks, logit and multivariate discriminant analysis. *Intl. J. Intelligent Systems in Accounting, Finance and Management*, 6: 235-248.
27. Yang, Z.R., M.B. Platt and H.D. Platt, 1999. Probabilistic neural networks in bankruptcy prediction. *J. Business Res.*, 44: 67-74.
28. Koh, H.C., 1992. The sensitivity of optimal cutoff points to misclassification costs of Type I and Type II errors in the going-concern prediction contest. *J. Business Finance and Accounting*, 19: 187-197.