December 1995

Applying Information Technology to Investigate the Global and Local Knoweldge of Loan Evluation

Ting-Peng Liang
National Sun Yat-Sen University

Chang-Er Chen
United Franchise Chain

Shin-Yuan Hung
National Sun Yat-Sen University

Applying Information Technology to Investigate the Global and Local Knowledge of Loan Evaluation

Ting-Peng Liang
School of Management
National Sun Yat-sen University
Kaohsiung, Taiwan

Chang-Er Chen
United Franchise Chain
Taichung, Taiwan

Shin-Yuan Hung
School of Management
National Sun Yat-sen University
Kaohsiung, Taiwan

Abstract

Loan evaluation is a major decision for financial institutions. This research investigates the differences in knowledge used for loan evaluation in the United States and in Taiwan. Taiwanese loan officers and LOAN PROBE, an expert system developed at KPMG Peat Marwick, were used to analyze loan cases collected from financial institutions in Taiwan. Their judgmental differences were analyzed using machine learning and statistical methods to test (1) whether knowledge differences exist in transnational settings, (2) where the differences exist, and (3) how a global knowledge structure can be constructed to offset the problem. The results indicate that significant judgmental difference exists between local experts in Taiwan and LOAN PROBE. LOAN PROBE relies more on current financial condition and future cash flow, whereas Taiwanese loan officers rely on industrial profitability. A knowledge base structure that divides knowledge of loan evaluation into global and local components is also proposed.

Keywords: *Loan Evaluation, Expert Systems, Global Information Systems.*

1. INTRODUCTION

As the world's business activities become more internationalized, it is increasingly important to understand how businesses are operated and accounting information is used in different countries. This understanding is especially important for auditors and financial institutions because they must take into account various risks involved in transnational settings. Failing to do so often results in unexpected risks. For example, a number of Asian branches of U.S. banks experienced substantial loan losses in their first few years of operation. This outcome was due primarily due inadequate knowledge of the usage difference in accounting information between the United States and Asian countries.

Bearing this problem in mind, this research investigates the transnational differences in the use of knowledge for loan evaluation and how information technologies such as machine learning and expert systems can be applied to improve the use of accounting information in transnational settings. In particular, we focus on the difference between the United States and Taiwan. Taiwan was chosen because (1) it is considered a window to Chinese culture, and (2) many US financial institutions have branches there.

The specific goals of the research project are three-fold:

1. To examine whether differences exist in using and interpreting accounting and other financial information for loan evaluation between the United States and Taiwan,
2. To locate the differences and find the global knowledge that is used across countries and local knowledge that is unique only in a certain country or area using knowledge engineering and machine learning techniques, and
3. To develop a knowledge base structure that takes into consideration the national and cultural differences in transnational loan transactions.

Loan evaluation was chosen as the domain for investigation because assessing the risks in commercial lending is a major issue for financial institutions and their auditors. Evaluation of the applicant's financial situation plays a key role in this decision. Advanced information technology, such as Expert Systems, has been used in the United States to support the decision [Baker, 1990; Blanning, 1984; Chandler and Liang, 1990; Chen and Liang, 1989; Connell, 1987; Duchessi, Shawky, and Seagle, 1988; Gardner and Mills, 1989; Ruparel and Srinivasan, 1992; Shaw and Gentry, 1988; Srinivasan and Kim, 1988; Zocco, 1985]. The application, however, becomes complicated in transnational settings because the financial information reported in different countries may have different meanings. Moreover, cultural differences may lead to different uses of the same information in different countries. As a result, knowledge or decision models that are useful in the United States may be less useful in other countries. Hence, a better understanding of the global and local factors affecting transnational loan decision is desirable.

The research discussed here comprises of three phases. First, Taiwanese loan cases were collected, and local loan officers were interviewed to acquire the knowledge they use. LOAN PROBE (an expert system developed at KPMG Peat Marwick based on expertise in the United States) was used to analyze the Taiwanese loan cases to verify the applicability of US knowledge in international settings. Second, machine learning methods were used to induce the Taiwanese and American loan evaluation models for knowledge comparison. Major factors, as well as their relative importance, were compared. Finally, LOAN

PROBE was modified to accommodate Taiwanese knowledge to see whether the expert system would perform as well as Taiwanese loan officers on Taiwanese cases. The results indicate that (1) significant knowledge differences exist between loan officers in Taiwan and LOAN PROBE, and (2) the difference can be offset by modifying the knowledge base of LOAN PROBE. In particular, a more structured knowledge base that divides its knowledge into global and local portions allows LOAN PROBE to be internationally applicable.

The remainder of the paper is organized as follows. First, literature on commercial loan evaluation is reviewed briefly. This is followed by a description of the research design. Research findings are then presented. Finally, a knowledge structure that includes global and local knowledge for loan evaluation is proposed.

2. LITERATURE REVIEW

In the AICPA [1984] Industry Audit Guide, loan evaluation is defined as "a matter of ascertaining loan collectability, that is, whether the loan will be repaid or the principal otherwise covered. The answer may depend, among other factors, on the borrowers' financial abilities as indicated in past and projected earnings and cash flows, credit history, net realizable value of loan collateral, and the financial responsibility of endorsers or guarantors." There are five primary factors (often called 5C's) that are generally considered to be important to commercial lending: credit, collateral, capital, capacity, and character (Wood, 1978). Most existing research uses either statistical or process modeling approaches to determine the relative importance of these factors.

2.1 Statistical Modeling

The statistical modeling approach uses statistical methods such as regression or discriminant analysis to identify factors that affect loan classification and to develop models that help classify loans. For instance, Altman [1970] proposed a bankruptcy discriminant model for screening commercial loans. Orgler [1970] developed a stepwise linear regression model for reviewing existing loans. Five dummy variables and the ratio of working capital to current assets were used in the study.

Abdel-Khalik and El-Sheshai [1980] applied discriminant analysis to build models from a sample of 32 firms. Six of the 18 variables were found significant for the experiment/control sample: total debt/total assets, trend of long-term debt, cash flow/total debt, net income/total assets, trend of net income/sales, and trend of net income/total assets. Current ratio, quick assets/sales, trend of long-term debt/net worth, and net income/sales were significant for the control/validation sample. The models correctly classified 90.6 percent of the experiment/control sample and 100 percent of the control/validation sample.

Dietrich and Kaplan [1982] employed a probit model to estimate bank lending officers' loan decisions. Their model provided maximum likelihood estimates of cutoff scores for classifying loans into four classes: current, especially mentioned, substandard, and doubtful. Three factors were found important in the model: the debt-equity ratio, a ratio of funds flow to fixed commitments, and a sales trend measure.

Marris, Patell, and Wolfson [1984] conducted similar research that compared the polytomous probit and recursive partitioning methods. Twenty financial statement variables and six non-financial statement variables were used. The results were evaluated in terms of expected classification loss rates based on assumed loss functions.

Gardner and Mills [1989] examined the relationship between loan attributes and default risks. Eighteen attributes were considered. The results showed that the loan-to-value ratio, payment-to-income ratio, age of loan, age of property, property location, and the appraiser's assessment of value trend had a significant impact on loan delinquency.

2.2 Process Modeling

In addition to statistical analysis, another line of work emphasizes the process by which factors are considered and evaluated by loan officers. Process modeling is different from statistical modeling in that knowledge of loan evaluation is represented in decision trees or decision rules. These trees or rules can then be used to build expert systems. The process knowledge can be constructed by interviewing experts or performing machine learning analysis on existing cases.

For example, Ribar, Willingham, and Bell [1990] interviewed an engagement partner and a senior audit manager to build a process model of more than 9000 rules for loan evaluation when they developed LOAN PROBE. The model is composed of industry knowledge, general business knowledge, lending knowledge, borrower information, and loan information. A sequence of decisions as shown in Figure 1 was identified in their study.

Instead of interviewing human experts to obtain knowledge, Messier and Hansen [1988] applied an inductive learning method called ID3 to generate rules for predicting loan defaults. The method allows a set of if-then rules to be produced automatically from a set of training cases. Data obtained from Abdel-Khalik and El-Shesai [1980] were analyzed and compared. The result indicated that ID3 performed better than the discriminant analysis in predicting the hold-out sample.

Shaw and Gentry [1988] used another inductive learning method called AQ15 to analyze the financial information of 58 companies (29 failed and 29 nonfailed). The result showed that the rules generated using AQ15 provided satisfactory predictive accuracy (86.2 percent on the whole sample and 73.3 percent on the hold-out sample).

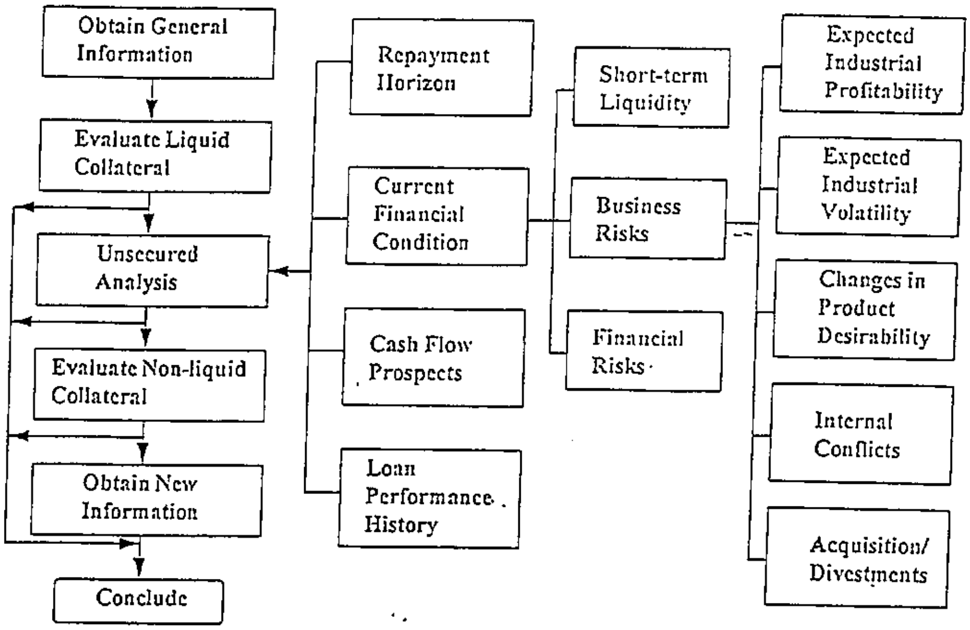


Figure 1. Knowledge of LOAN PROBE

Srinivasan and Kim [1988] presented an approach based on Analytic Hierarchy Process (AHP). Five loan attributes were considered critical to the risks of commercial loans: financial strength of the customer, customer background, payment record, business potential and frequency, and geographical location. They argued that "credit investigation proceeds hierarchically... The hierarchical order of investigation depends on the level of credit investigation and information availability. Typically, an evaluation of financial soundness is followed by an examination of pay record, customer background, business potential and frequency, and geographical location." The AHP model was evaluated by comparing its prediction with actual credit decisions of 100 customers. The hit rate of prediction was 92 percent.

Given the impressive result of the process-oriented approach, a few researchers have proposed the development of expert and decision support systems to facilitate the decision. For example, Zocco [1985] proposed a framework for building expert systems in bank loan management. Duchesai, et al. [1988] presented a knowledge-based system for commercial loan decisions. The system divides the knowledge into two levels: general analysis that examines the general trend, credit, collateral, capital, and capacity, and a detailed analysis that investigates inventory turnover and other detailed information if necessary. Ruparel and Srinivasan [1992] proposed a dedicated shell for supporting loan decisions.

Although the literature reviewed above has provided much insights into the process of commercial loan evaluation, little research has investigated the transnational applicability of the models developed in the United States. Given the rapid internationalization of financial institutions, it is interesting to study *whether knowledge difference exists in different countries, where the difference exists, and how to offset the difference* when transnational loans are involved.

3. RESEARCH DESIGN

The research is divided into three phases to examine the above three major issues in transnational applications of expert systems. To investigate *whether knowledge differences exist* for loan evaluation, we collected a set of thirty loan cases from three banks in Taiwan. These cases were then classified for comparison by representative knowledge in the United States and in Taiwan.

A key issue in this phase is how to choose representative knowledge in either country. In the U.S., KPMG Peat Marwick has invested a substantial amount of resources to develop an expert system called LOAN PROBE. The knowledge base of the system is composed of nearly nine thousand rules and has been tested and evaluated extensively (Ribar, et al., 1990, 1991). Therefore, we can reasonably assume it to be representative of the knowledge used in the United States.

Since there is no comparable system in Taiwan, three senior loan officers were asked to provide their judgments on the loan cases they offered for study. A problem in comparing LOAN PROBE with human judgments is that LOAN PROBE provides a probabilistic assessment of the loan loss reserve whereas human experts often classify loans into categories such as excellent, fair, and bad. In order for the results to be comparable, the loan loss reserve percentage was converted into the three-category scale. Those judged to have no loan loss risk are classified as

Case	Human (A)	LOAN PROBE (B)	Difference |A-B|
1	2	3	1
2	3	2	1
3	2	1	1
4	3	3	0
5	2	1	1
6	1	3	2
7	3	3	0
8	3	3	0
9	2	3	1
10	3	3	0
11	3	3	0
12	2	3	1
13	3	2	1
14	3	3	0
15	2	3	1
16	2	3	1
17	3	3	0
18	1	1	0
19	1	2	1
20	2	2	0
21	1	1	0
22	1	1	0
23	1	2	1
24	1	3	2
25	3	3	0
26	2	2	0
27	3	3	0
28	3	3	0
29	3	3	0
30	2	2	0

Note: 3 = Excellent, 2 = Fair, 1 = Poor

Table 1. Comparison of Human and LOAN PROBE

excellent. Those judged to have risks between 1 to 10 percent are classified as *fair.* The remaining cases are considered *bad.* The classification mechanism was used for convenience and was based on loan experts' subjective judgment.

With the help of HBU Bank in Taipei, we successfully obtained thirty cases with different degrees of loan risk. These cases were then analyzed by LOAN PROBE and a senior loan officer in HBU Bank. The results are then compared to test the following null hypothesis:

H_1: The classification results of LOAN PROBE and local loan officers are the same.

If differences exist, we would like to know *where they exist* in the second phase. The methods employed include both descriptive analysis and inductive learning. Local loan officers and senior auditors were asked to evaluate the knowledge base of LOAN PROBE, with an emphasis on finding how it differed from local practice. An inductive learning method was also used to create decision trees from the classifications of LOAN PROBE and local loan officers. The resulting decision trees were then compared to see the difference.

Following the discovery of knowledge differences, we proposed a knowledge base architecture for global expert systems and modified the knowledge base of LOAN PROBE to incorporate local knowledge. To know whether the modified system would operate as planned, the same set of loan cases was used to evaluate the modified LOAN PROBE. A successful modification should result in classification similar to local loan officers'. Therefore, we hypothesize the following:

H_2: The classification results of the revised LOAN PROBE and local loan officers are the same.

4. RESULTS AND DISCUSSIONS

4.1 Whether differences exist?

After analyzing all thirty loan cases, the classification results of local loan officers and LOAN PROBE is shown in Table 1. Human experts identified thirteen cases as excellent, ten as fair, and seven as poor, whereas the system identified eighteen as excellent, seven as fair, and five as poor (see Table 2). Results of a paired-t test indicate that the judgments of human experts and LOAN PROBE are significantly different ($t=4.281$, $p<0.001$). Therefore, we reject H_1 and conclude that *if we apply LOAN PROBE to Taiwanese loan cases, significant judgmental errors would occur.*

In Table 2, we reorganize the classification result and find that two cases considered *poor* by local loan officers were classified as *excellent* (i.e., no risk) by LOAN PROBE. In practice, this outcome could cause major problems. There are also five cases that were classified as *fair* by human experts but classified as *excellent* by the system. On the other hand, no case classified as *excellent* by human experts was classified as *poor* by the system. Two cases classified as *fair* by loan officers were classified as *poor* by the system. If we use human experts as a benchmark, LOAN PROBE under-estimated the risk in nine (5+2+2) cases, whereas it over-estimated the risk in four (2+2) cases. LOAN PROBE tends to be too optimistic.

4.2 Where the differences exist?

Given the judgmental difference, we further examined what caused the difference. We analyzed the knowledge in LOAN PROBE's knowledge base, traced LOAN PROBE's decision processes, showed the knowledge base and the traces to the local loan officers, and then asked them to identify the differences specifically. As a result, differences were identified in both the process and the attributes of evaluation.

Regarding the evaluation process, LOAN PROBE uses a sequential evaluation of three major components: liquid collateral, unsecured analysis, and non-liquid collateral, as shown in Figure 1. Each case is first checked for its liquid collateral. If the liquid collateral is able to cover the whole amount of loan, no further analysis needs to be done.

System / Human	Excellent	Fair	Poor	Total
Excellent	A 11	B 2 (2,13)	C 0	13
Fair	D 5 (1,9,12,15,16)	E 3	F 2 (3,5)	10
Poor	G 2 (6,24)	H 2 (19,23)	I 3	7
Total	18	7	5	30

Table 2. Contingency Table

Evaluation Attributes	Loan Probe	Taiwanese Banks
Bank Deposits	√	√
Securities	√	√
Industrial Analysis	√	√
Financial Report Analysis	√	√
Cash Flow Analysis	√	
Borrower's Record	√	√
Guarantee from Board Directors		√
Account Receivables	√	
Draft Receivables		√
Letter of Credit		√
Real Estate	√	√
Machinery	√	√
Inventory	√	√

Table 3. Comparison of attributes considered by LOAN PROBE and Taiwanese Banks

Otherwise, the unsecured analysis is activated. The unsecured analysis considers repayment horizon, current financial conditions, cash flow prospects, and loan performance history. If the unsecured analysis concludes that the residual exposure is adequately covered, no further analysis is necessary. Otherwise, the case goes through an analysis of its non-liquid collateral.

The major difference in the process is that most Taiwanese loan officers do not consider unsecured analysis as an important part of the loan evaluation. The evaluation process of local loan officers is to examine the liquid collateral first. If it is inadequate, the non-liquid collateral such as land is required. Unsecured analysis is used only if the first two are inadequate. The reason that non-liquid collateral is given a higher weight than the unsecured analysis is understandable. In the last three decades, the value of land has increased significantly more than any other asset in Taiwan. Therefore, most financial institutions have very strong interests in holding land as collateral.

In addition to the process difference, certain loan attributes evaluated in LOAN PROBE were not considered by Taiwanese loan officers, and vice versa. Table 3 shows a comparison of the major attributes examined by LOAN PROBE and by Taiwanese experts. We found that cash flow analysis and account receivable are generally not used by Taiwanese loan officers. The reason is that most loan officers do not trust cash flow predictions and the accounts receivable information. Instead, they use *checks receivable* because most account receivables are secured by long-term checks (such as checks cashable or redeemable after 3 or 6 months). In fact, information obtained from financial reports has a lower weight for Taiwanese loan officers than for LOAN PROBE. The major reason is that loan officers do not trust financial reports, because accounting systems are less reliable than in the United States.

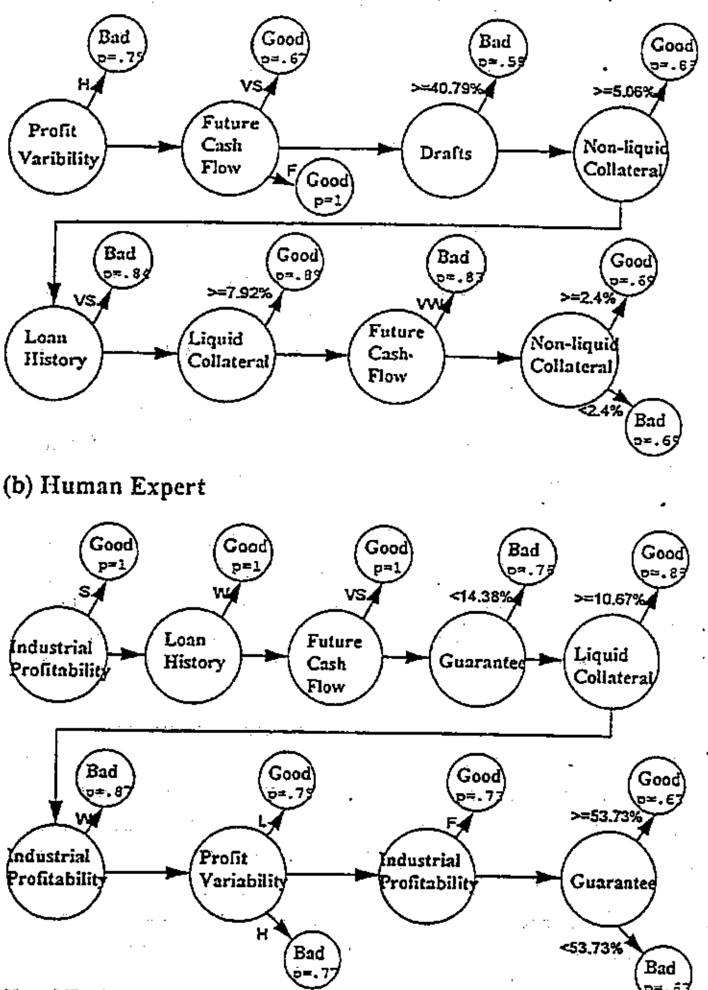
Another collateral used often in Taiwan but not included in LOAN PROBE is an LC (Letter of Credit). Because most Taiwanese companies relies heavily on international trade, the financial institutions often grant loans based on an LC issued by another financial institution. Furthermore, Taiwanese loan officers often require guarantees from the board of directors to reduce the financial institution's risks. Whether a loan is guaranteed by the Board Directors is a major factor in the evaluation.

To determine what caused the difference in the judgments, we further analyzed the types of collaterals and guarantees in all cases and found striking differences. We divided the cases into three groups: hit (LOAN PROBE and human experts gave the same classification), over (LOAN PROBE over-estimated the risks), and under (LOAN PROBE under-estimated the risks). Table 4 shows the types of collaterals and guarantees involved in the loan cases in each category.

As shown in Table 4, the nine cases whose risks were under-estimated by LOAN PROBE had no guarantees from the Board of Directors, whereas those whose risks were over-estimated all had guarantees. The under-estimated

	Under	Hit	Over
Bank Deposits	0	7	0
Stocks	5	2	0
Draft Receivable	0	1	1
LC	0	0	2
Fixed Assets	1	1	2
Real Estate	0	3	0
Board Directors' Guarantee	0	9	4
Total No. of Cases	9	17	4

Table 4. Types of Collaterals and Guarantees



Note: VS=Very Strong, S=Strong, VW=Very Weak, W=Weak, F=Fair, H=High, L=Low

Figure 2. Decision Tree

cases involved stocks and favored evaluation from unsecured analysis by LOAN PROBE. A discriminant analysis using the STEPDISC procedure in SAS resulted in six significant discriminating factors that can differentiate the classification results of LOAN PROBE and local loan officers. They are *LC* ($p<0.002$), *Board of Directors' guarantee* ($p<0.003$), *bank deposit* ($p<0.001$), *stocks* ($p<0.02$), *drafts receivable* ($p<0.05$), and *real estate* ($p<0.02$).

Finally, we applied a machine learning method, called CRIS, to examine the rules used by LOAN PROBE and human experts when analyzing the cases. CRIS stands for the Composite Rule Induction System. It integrates statistical and inductive learning concepts, and it is an accurate approach in classification problems (Liang, 1992). The resulting knowledge structures, as shown in Figure 2, also allow us to see the judgmental differences between LOAN PROBE and loan officers. The top three items differentiating the cases were profit variability, future cash flow, and draft receivables for LOAN PROBE, whereas the top three items were industrial profitability, loan history, and future cash flow for the loan officers.

As a summary, differences were found in both the process and the attributes of evaluation. In the process, unsecured analysis is given a much lower weight in Taiwan than in the United States. This is primarily due to the lack of confidence in financial reports signed by local accounting firms. Regarding the attributes, guarantees from the board of directors are used very often. Others factors, such as LC, stocks, bank deposits, and real estates are also evaluated differently.

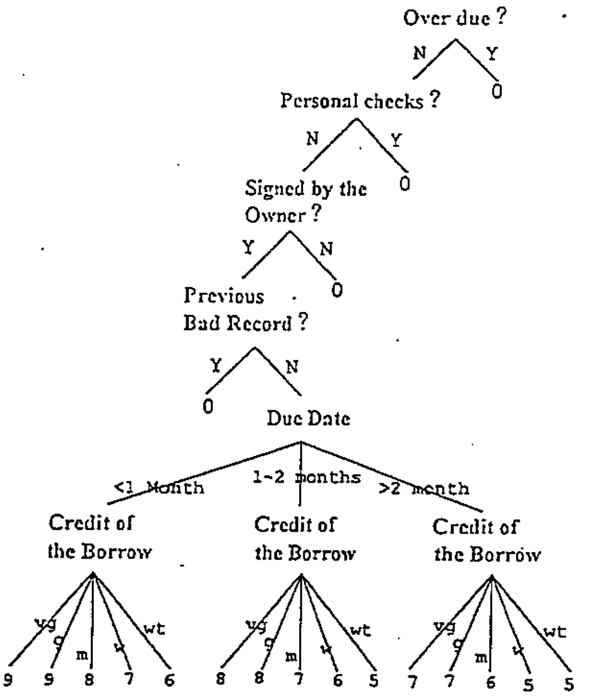
4.3 How to offset the difference?

After identifying the differences, the next step was to find ways by which LOAN PROBE could be revised to fit both countries. In this stage, we revised the knowledge base of LOAN PROBE and then classified the original loan cases using the revised system. The results were compared with local loan officers' classifications to see whether the revision would result in similar classifications (i.e., testing hypothesis H_2). The reason we used the original cases is that we would like to know whether the revision had offset the differences found in the loan cases. Therefore, it is justifiable. Since differences, as described in Section 4.3, were found in attributes as well as the evaluation process, modifications were done on both.

(1) Revisions on evaluation attributes

Based on the knowledge obtained from the Taiwanese loan officers, modifications involved liquid collateral, unsecured analysis, and non-liquid collateral. In liquid collateral, the following changes were made:

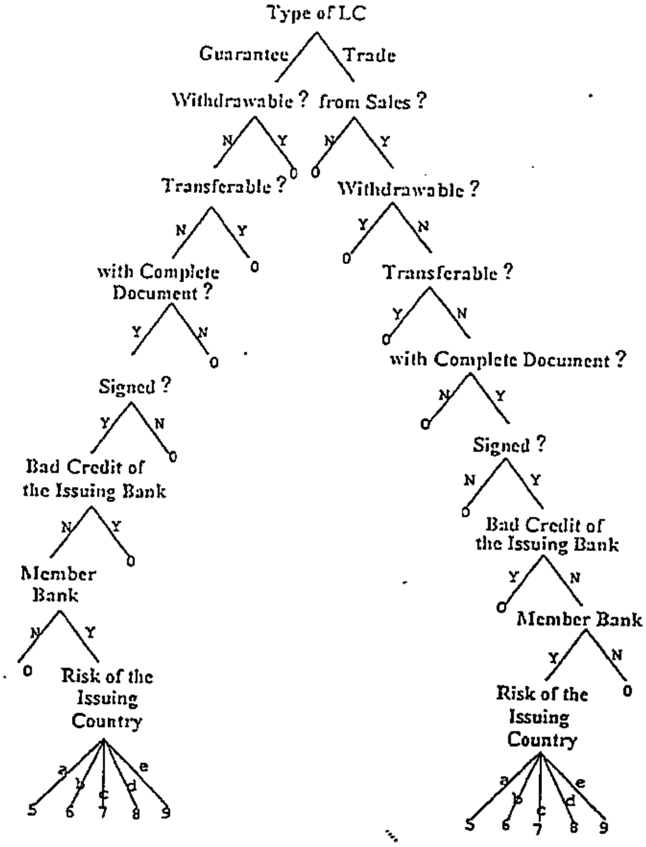
- Bank deposit: The residual value was set to zero if other loans have higher priorities.
- Account receivable: This was removed from the knowledge base.



Note: 1. vg = very good, g = good, m = moderate, w = worse, wt = worst

2. The numbers represent the risk classification with 9 to be excellent and 0 the worst.

Figure 3. Decision Tree for the Draft Receivables



Note: 1. a = very high, b = high, c = moderate, d = low, e = very low

2. The numbers indicate the degree of delingment risk with 9 to be the lowest and 0 to be the highest.

Figure 4. Decision Tree of LC

• Draft receivable: This knowledge was added into the knowledge base. The knowledge tree is shown in Figure 3.

• Letter of credits: This knowledge was added into the knowledge base. The knowledge tree is shown in Figure 4.

In unsecured analysis, the following changes were made:

• Credit ranking: This was deleted because no risk ranking was available in Taiwan at the time of this study.

• Financial statement analysis: Since financial statements are less credible in Taiwan than in the United States, the weight of the financial statement analysis was reduced.

• Guarantee: We added the guarantee from the Board of Directors and gave it a high weight.

In non-liquid collateral, we made the following changes:

• Inventory: It was removed from the knowledge base because most Taiwanese banks do not like to take inventory as collaterals.

• Real estate: Its weight was increased because the risk was pretty low in Taiwan.

(2) Revisions on the evaluation process

After discussing with loan officers in Taiwan, the evaluation process, as shown in Figure 4, was developed. It begins with an evaluation of the liquid collateral including bank deposit, other commercial papers, draft receivable, and letter of credits. The analysis of liquid collateral is followed by the evaluation of preferred non-liquid collateral such as real estate. If the above analysis does not provide adequate risk coverage, loan officers proceed with unsecured analysis of credit. In normal cases, guarantees from the Board of Directors are considered at this stage. If the unsecured analysis still cannot provide adequate coverage, inventory and other fixed assets are considered. In addition to the evaluation process, the weight assigned to each step was also modified based on the information obtained from the loan officers.

(3) Test of the new system

To know whether the above revisions would offset the differences, we used the new system again to analyze the original thirty cases. The results are shown in Tables 5 and 6. Only four cases were misclassified by the modified system, a significant improvement over the old system. By comparing results in Tables 1 and 5, we find the following:

• Case 2 moved from zone B to zone A due to the addition of draft receivables.

• Case 5 moved from zone F to zone E due to the addition of LC.

Case No.	Loan Officer (A)	Revised LOAN PROBE (B)	Difference |A-B|
1	2	3	1
2	3	3	0
3	2	1	1
4	3	3	0
5	2	2	0
6	1	2	1
7	3	3	0
8	3	3	0
9	2	2	0
10	3	3	0
11	3	3	0
12	2	2	0
13	3	2	1
14	3	3	0
15	2	2	0
16	2	2	0
17	3	3	0
18	1	1	0
19	1	1	0
20	2	2	0
21	1	1	0
22	1	1	0
23	1	1	0
24	1	1	0
25	3	3	0
26	2	2	0
27	3	3	0
28	3	3	0
29	3	3	0
30	2	2	0

Note: 3 = Excellent, 2 = Fair, 1 = Poor

Table 5. Comparison of Classifications of Human and the Revised LOAN PROBE

System / Human	Excellent	Fair	Poor	Total
Excellent	A **12**	B **1** (13)	C **0**	**13**
Fair	D **1** (1)	E **8**	F **1** (3)	**10**
Poor	G **0**	H **1** (6)	I **6**	**7**
Total	**13**	**10**	**7**	**30**

Table 6. Contingency Table

- Cases 9 and 15 in zone D moved to zone E, and case 23 in zone H and case 24 in zone G moved to zone I, due to reducing the weight assigned to stocks.
- Cases 12 and 16 in zone D moved to zone E, case 19 in zone H moved to zone I, and case 6 moved from zone G to zone H, due to reducing the weight assigned to the analysis of financial reports.

In Table 6, we see the number of cases on the diagonal cells (i.e., A, E, and I) increases from 17 to 26. The result of the paired-t test over the data in Table 5 is NOT significant. We cannot reject H_2. That is, the judgment of the new system is the same as that of the loan officers in Taiwan. We also performed a $\chi 2$ test to compare the classification of the old and the new systems. The result shows significant statistical differences ($\chi 2$ = 18.12, $p \leq .05$). Therefore, we can safely conclude that the modification was able to offset the knowledge difference between LOAN PROBE and Taiwanese loan officers.

5. CONCLUDING REMARKS

Applying expert systems across national borders is an interesting issue. There are a number of issues that must be considered. In this research, we use LOAN PROBE as an experimental tool to examine its cross-national applicability. Empirical results indicate that several problems exist in addition to language differences. The implications of the research are two-fold.

First, when multi-national firms are exporting their knowledge, they must consider its global applicability. For domains such as loan evaluation, knowledge differences between the loan officers in Taiwan and in the United States may exist in the attributes they consider and the process of evaluation. The weights given to different attributes may also differ. These differences are primarily due to differences in cultural, social, and economic practices in different countries.

Second, if we develop expert systems for global use, a new architecture that accommodates national differences may be necessary. A straight-forward approach is to design the system in module. The knowledge base may be divided into two types: global and local knowledge bases. The global knowledge base contains knowledge that is applicable world-wide. For instance, the knowledge required to analyze financial reports is a kind of global knowledge. The local knowledge base contains knowledge that is applicable only in a particular country or region. For instance, the evaluation of real estate may differ in many different countries. Hence, it should be stored in local knowledge bases. The evaluation process and weights assigned to each module is also considered local, and it may change from nation to nation. This way, expert systems developed in one country can be applied to other countries properly by revising the local knowledge base.

Due to the difficulties in otaining loan data and other research procedures, the research has a few limitations. First, we did not use new loan cases to verify the validity of the revised knowledge base. This may limit the generality of the revised knowledge. Second, we used judgments from three loan officers as benchmark for comparison. Their judgments are sometimes inconsistent or erroneous. In this case, the modification becomes useless. Finally, we do not know the actual outcomes of the loans studied in the research. As a result, we don't know the best real decision for improving the system.

Following this research, there are at least two possible extensions. First, we may extend the project to include other countries such as Japan or Hong Kong to see whether the results found in this study hold in other nations. Another extension is to study other domains such as bankruptcy analysis and audit judgment. In both areas, cultural and social differences play a key role in making decisions.

Acknowledgements

The author is indebted to many people who have helped the project. In particular, I would like to thank Gary Ribar at KPMG Research Foundation, Bearl Lin, J.J Wang at KPMG Taiwan, and Alex Lin at HBU Bank for their assistance.

References

Abdel-Khalik, A.R. and El-Sheshai (1980), "Information Choice and Utilization in an Experiment on Default Prediction," *Journal of Accounting Research*, Autumn 1980, pp. 325-342.

AICPA (1984), *Auditing Procedure Study*, Auditing the Allowance for Credit Losses of Banks, American Institute of Certified Public Accountants.

Altman, E.I. (1970), "Corporate Bankruptcy Prediction and Its Implications for Commercial Loan Evaluation," *Journal of Commercial Bank Lending*, December, pp. 8-22.

Baker, A.M. (1990), "Commercial Lending Expert Systems: A State-of-the-art Review," *Commercial Lending Review*, 5:4, pp. 34-46.

Blanning, R.W. (1984), "Management Applications of Expert Systems," *Information and Management*, pp. 311-316.

Chandler, J.S. and Liang, T.P. (1990), *Developing Expert Systems for Business Applications*, Columbus, OH: Merrill Publishing.

Chen, K.C. and Liang, T.P. (1989), *Financial Applications of Expert Systems*, A Special Focused Issue of Managerial Finance, Vol. 15, No. 5.

Chen, S. (1983), "A Study on Credit Lending Considerations of Local and Foreign Banks," *Master Thesis, National Chiao-Tung University*, Taiwan (in Chinese).

Connell, N.A.D. (1987), "Expert Systems in Accountancy: A Review of Some Recent Applications," *Accounting and Business Research*, 17:67, pp. 221-233.

Dietrich, J.R. and Kaplan, R.S. (1982), "Empirical Analysis of the Commercial Loan Classification Decision," *Accounting Review*, January, pp. 18-38.

Duchessi, P., Shawky, H., and Seagle, J.P. (1988), "A Knowledge Engineered System for Commercial Loan Decisions," *Financial Management*, Vol. 17, No. 3, pp. 57-65.

Gardner, M. and Mills, D.L. (1989), Evaluating the Likelihood of Default on Delinquent Loans," *Financial Management*, 18:4, pp. 55-63.

Ives, B. and Jarvenpaa (1991), "Applications of Global Information Technology: Key Issues for Management," *MIS Quarterly*, (March), pp. 33-49.

Jones, S.L. (1992), "How to Obtain a Small Business Loan," *Economic Development Review*, 10:2, pp. 74-77.

KPMG Peat Marwick (1990), *Banking in Taiwan*, Taipei, Taiwan: KPMG Peat Marwick.

Liang, T.P. (1992), "A Composite Approach to Inducing Knowledge for Expert Systems Design," *Management Science*, 38:1, pp. 1-17.

Lin, L.T. (1987), "A Comparative Study on Loan Evaluation Policies Between Local and Foreign Banks in Taiwan," *Taipei City Bank Monthly Journal*, 15 (in Chinese).

Marias, M.L., Patell, J.M., and Wolfson, M.A. (1984), "The Experimental Design of Classification Models: An Application of Recursive Partitioning and Bootstraping to Commercial Bank Loan Classification," *Journal of Accounting Research*, Supplement, pp. 84-114.

Messier, W.F. Jr. and Hansen, J.V. (1988), "Inducing Rules for Expert System Development: An Example Using Default and Bankruptcy Data," *Management Science*, 34:12, pp. 1403-1415.

Norris, D.M. (1992), "Real Estate Loan Underwriting Factors in the Insurance Industry," *Real Estate Finance*, 9:3, pp. 79-86.

O'Neal, Q. (1991), "International Knowledge Development in the 1990's," *Working Paper, IBM Manufacturing Expert System Project Center*.

Orgler, Y.E. (1070), "A Credit Scoring Model for Commercial Loans," *Journal of Money, Credit, and Banking*, November, pp. 435-445.

Ribar, G., Arcoles, F. and Hollo, D. (1991), "Loan Probe: Testing a Big Expert System," *AI Expert*, 6:5, pp. 43-49.

Ribar, G.S., J.J. Willingham, and T.B. Bell (1990), "The Auditor's Evaluation of Commercial Loans," *Working Paper, KPMG Peat Marwick Foundation*.

Ruparel. B. and Srinivasan, V. (1992), "A Dedicated Shell for Designing Expert Credit Support Systems," *Decision Support Systems*, 8:4, pp. 343-359.

Shaw, M.J. and Gentry, J.A. (1988), "Using an Expert System With Inductive Learning to Evaluate Business Loans," *Financial Management*, 17:3, pp. 45-56.

Srinivasan, V. and Kim, Y.H. (1988), "Designing Expert Financial Systems: A Case Study of Corporate Credit Management," *Financial Management*, 17:3, pp. 32-44.

Stonehill, A., Beekhuisen, T., Wright, R., Remmers, L., Toy, N., Peres, A., Shapiro, A., Egan, D., and Bates, T. (1975), "Financial Goals and Debt Ratio Determinants: A Survey of Practice in Five Countries," *Financial Management*, Autumn, pp. 27-41.

Wood, O.G., Jr. (1978), *Commercial Banking: Practice and Policy*, New York, NY: Van Nostrand.

Zocco, D.P. (1985), "A Framework for Expert Systems in Bank Loan Management," *Journal of Commercial Bank Lending*, 67:5, pp. 47-54.

TREATMENT OF DATA IN THE DEVELOPMENT OF INTELLIGENT INFORMATION SYSTEMS SUPPORTING FINANCIAL DECISIONS

Albert C. K. Chan and Richard F. Bonner

School of Information Systems & Management Science
Faculty of Commerce and Administration, Griffith University
Nathan-Brisbane, Australia 4111

Solving the unstructured financial problems belongs to the "knowledge-rich" task domain, collectively requiring massive current and historical data. The problem is here discussed from the standpoint of a general strategy for the deployment of intelligent information systems, concentrating on the single most critical but manageable issue: the choice and treatment of data in their development. The training in artificial neural networks is formally the construction of a "problem → solution" schema from a set of examples, and the acquired schemas depend in practice significantly on the form of the data. The suggested preparation of data consists of four basic steps: the collection, analysis, preprocessing, and the separation of data into training and testing sets. The data preparation process is illustrated by the case of a currently developed intelligent system for exchange rate risk management. It is re-emphasised in conclusions that data preparation is only one, albeit vital, aspect which makes the development of intelligent information systems quite different from that of conventional information systems.

1 Introduction

Ever since the days of Eniac, computers have been used as problem solving tools. By noting the common thread through the historical trend of computer applications from supporting structured problems to tackling unstructured problems, the knowledge of the programmer/expert becomes encoded in the program in the form of a typically fix input-output mapping, the "system transfer function" which operates on "data" transforming it into some "more useful" form. The "intelligence" of such white-box systems is thus merely a reflection of the knowledge of their creators. By contrast, the introduction of "artificial neural networks" into business information systems, as it turns out predominantly in the area of financial management (Jagielska, 1993; Trippe & Turban, 1993), in principle short-circuits the traditional role of programmers. Such systems learn by experience (Zurada, 1992): the transfer function (problem → solution) of the system is not *a priori* determined by the programmers but gradually develops as a result of repetitive exposure to data - naturally, in accordance with some in-programmed scheme. The resulting system is a black-box: its transfer function can be observed but is not easily "explained": the "rule extraction" from neural networks is a known problem.

Seen against this background, the data-rich but fairly well understood area of financial management should be ideal for neural network applications. Financial theory is conceptually advanced with all its variables readily measurable, and it should be straightforward to set up a neural network following some general financial model. Able to handle massive data, one would think, a neural network should be an ideal extension of economic and financial models, which provide structure and insight but are in themselves of limited predictive power (Meese & Rogoff, 1983). Indeed, among existing commercial applications of neural network technology, financial applications clearly dominate (Jagielska, 1993). Upon a closer examination, however, and compared to the scope and sophistication of neural network applications in technical areas such as speech (Sejnowski & Rosenberg, 1986) or image (LeCun, *et al*, 1989) recognition, the typical financial neural network applications rank rather low (Taylor, 1993), and their authors have to work hard to justify their choice of neural network technology from dozens of other, usually better understood and theoretically just as suitable, approximation schemes.

This observation, readily verifiable in isolated cases, but, of course, in need of research to give it statistical validity, is sufficiently alarming to make one consider its probable causes and potential consequences. We do this in Section 3, summarising the principles of neural network technology and discussing some of its policy issues. The overlaying general problem could be summarised as the lack of a basic policy on the development, use, and management of intelligent information systems in socio-economic context. To allow a non-technical yet meaningful discussion, we previously build up some cognitive background in Section 2, recalling simple but insightful analogies between human and machine learning. The discussion quickly focuses on a single and directly addressable problem: the treatment of data in the development of intelligent information systems. This problem is expanded in the main Section 4 outlining a scheme of data preparation for neural networks and heuristically discussing its four main steps: the collection, analysis, preprocessing, and the separation of data into training and testing sets. Finally, in Section 5, a current research project (Chan & Bonner, 1995) on intelligent systems in exchange rate risk management is used as illustration.

2 Problem schemas for knowledge-rich task domains

We start by recalling some cognitive concepts of significance for both human and machine problem solving. In particular, we hope to clarify the role of "data" in solving problems, and to prepare ground for Section 3 viewing artificial neural networks as problem schemas.

In essence, "problem solving" is the process of transforming the "data" defining a "problem" into "data" recognised as its "solution", usually by successively applying a series of simple transformations (VanLehn, 1989). According to Webster's dictionary (Mish, 1989), *datum* (pl. *data*), from Latin, literally *gift* or *present*, is "something given or admitted esp. as a basis for reasoning or inference". Thus, etymologically speaking, being "data" is not a property of any thing in isolation but a property of a (thing, transformation scheme) pair.

In one of the task domains in problem solving, the "knowledge-lean" domain, extensively studied by Newell and Simon in their classic (Newell & Simon, 1972), the problem solving transformations are fairly structured and essentially independent of the data. Thus, only little knowledge is required to solve the bulk of problems in such a domain once the transformation techniques have been mastered. In practice, however, problems are usually "contextual", each requiring its own specific transformations. Consequently, although the knowledge required to solve a single practical problem may not be significant, to keep on solving them requires copious amounts of current and historical data. Such practical problems, financial management problems included, belong to a "knowledge-rich" task domain.

It is interesting to recall some known facts explaining how "experts" solve problems in a knowledge-rich task domain (VanLehn, 1989). Roughly, experts recognise a problem as an instance of a familiar problem type, retrieve a solution template from their memory, and generate the problem's solution; they do not, in general, move about searching for a solution. Their behaviour is a product of knowledge gained through learning. As shown by the chess move experiment (Charness, 1981), for example, expertise lies not in having a more powerful overall strategy in problem solving, but rather in having a better knowledge for making elementary decisions.

The template or knowledge representation of a problem and its corresponding solution in an expert's memory is known as a problem schema (VanLehn, 1989); see also Arbib & Hesse (1986) where a theory of schemas is discussed in a broad cognitive and philosophical context. Simply and formally, a problem schema can be thought of as the "problem $\rightarrow$ solution" mapping. It consists of a class of problems the schema applies to together with the corresponding solutions. For example, in the following simple schema:

(a1) *there is a tendency for the currency to move down in the next three months,*
(a2) *the net cash outflow is high,*
(b) *so the cash flow should be hedged,*

the statements (a1) and (a2) describe the problems and the statement (b) describes a solution. Experts would have acquired many large and refined pieces of problem schemas. Each piece is highly specialised and is therefore effective in solving a particular class of problems (Chase & Ericsson, 1981).

Humans and machines alike (Poggio & Girosi, 1990) acquire problem schemas through learning, ie, through "resilient changes in the subject's knowledge about a task domain that is potentially useful in solving further problems" (Simon, 1983). Over time, problem schemas are formed by exposure to collection of problem-solution pairs. Of course, in contrast with simple supervised learning (Zurada, 1992), in more complex forms of learning, the association of a problem with its solution may not always be provided in examples, requiring instead an involved search process which may produce, along the way, new schemas or even re-define the initial problem. Whatever the particulars of the training process, after successful training, when presented with something "close" to the input of a learned pair (the problem), the learner can recall something "close" to the corresponding output (the solution).

We summarise all this on a lighter note with three "facts of life":

(i) **Data is all there is.**
In term of formal modelling of cognition, data is the basic element and all else is derived in terms of its "patterns" or "properties" (Bonner & Chan, 1995). In knowledge-lean domains, patterns are abundant giving rise to a rich transformation structure organising the "raw" data. In knowledge-rich domains, there is little pattern and more data needs to be retained in "raw", problem-specific, form.

(ii) **Learning always takes time but learning from poor data takes very long time.**
That the statement holds for human learning is a matter of direct experience: it takes a lifetime to become an "expert" even though one builds on transferred knowledge of previous generations. Machine learning is no different: the underlying physical process may be speeded up to a degree, but the task of organising experience into problem schemas is identical for man and machine.

(iii) **"Natural" and "artificial" are essentially the same.**
The distinction between "natural" and "artificial" (intelligence, learning, life (Hogeweg, 1993), ...) is simply not practical as long as both phenomena are indistinguishable in any operational (behavioural) sense. As quoted in the paper by Bonner (1993), "what does it matter whether one thinks with jelly or with wires?" The issue is no longer part of science fiction, as intelligent systems are becoming integrated with human decision making. For example, legal problems in trying to maintain a separation between the two by mixing the behavioural and the cognitive, are very real and urgently require solutions.

3 Artificial neural networks: problem schemas of a data-driven technology

Artificial neural networks are essentially as old as the von Neuman computer, and have, over the years, acquired academic standing as an acknowledged field of research. Commercial applications of neural networks are, on the other hand relatively recent but fast growing in number. In economic analysis and forecasting, there are applications to bankruptcy prediction (Koster, *et al*, 1990), stock market prediction (Wan & Chan, 1993; Baba & Kosaki, 1992), macroeconomic evaluation (Li, *et al*, 1991), and exchange rate forecasting (Refenes, 1993); and, in financial risk management, the neural network applications include bond rating (Dutta & Shekhan, 1988), mortgage underwriting (Collins, *et al*, 1988), and hedging strategy decision in foreign exchange market (Chan & Bonner, 1995). We have in this paper opted for calling information systems with neural network modules "intelligent", noting, however, that the words "adaptive" or "learning" could have been possibly more accurate but less established choices.

An artificial neural network is a "neurally inspired" (Rumelhart, 1989) computational scheme with "architecture" conceptually similar to that of its biological counterpart. Briefly, the "input signals" from the external environment are transformed into the "output signals" by the "network transfer function". Just like people, before neural networks can be entrusted with practical decisions, they must be put through a comprehensive training process, in which the parametric weights to the input signals are adjusted iteratively until a certain degree of accuracy of the output has been reached. After training, the values of the weights determine the "network transfer function" which is the problem schema to be used in life problem solving.

3.1 Promises and limitations, revisited

As pointed out in Section 2, the "intelligent" behaviour can be encoded in form of a transfer function of the "system" in question. Artificial neural networks can adapt its transfer function to the data of the environment in an "autonomous" way (Serra & Zanarini, 1990) and can thus exhibit "adaptive" or "intelligent" behaviour. Thus, so far and in principle, one could conclude that neural network technology should have no bounds in self-education, eventually perhaps surpassing all human intelligence.

In the application area here discussed, however, it is hard to find a neural system justifying the use of this technology above other, better known and often more effective computing schemes. Indeed, if restricted to the "feed-forward" networks trained by supervision, the two basic potential advantages of neural networks over classical function approximation schemes, which are the parallel structure and the speed of recall, are rarely utilised at all in existing financial applications. At the same time, the obvious disadvantage of the technology, its low explanatory value, does not seem to bother anyone.

We know now that "raw crunching power" of the computer running some universal algorithm for a giant neural network is not a realistic solution to the problem of learning from massive unorganised data. Recent hardware implementations of neural network learning algorithms, possible analogue computing, and whatever else the immediate future may bring, will of course be a continued improvement, but none of these could be a panacea solution for the problems of combinatorial explosion in dealing with unorganised data. The "contextual approach" is likely to remain for some time the only way out: problems need to be carefully analysed, data carefully selected, architectures and algorithms carefully adapted to the problems. There seem to be no universal short-cuts.

3.2 Policy issues

There is little doubt that the technical disciplines will continue their spectacular achievements in computing technologies, and that, consequently, computer-based information systems will continue to grow in their ability to take on ever more sophisticated intellectual tasks. What is required on the most general level is the establishment of a basic policy on the development, use, and management of intelligent information systems in the various socio-economic domains. Such policy will of course take time to develop as it must be grounded on a thorough understanding of basic issues. We suggest that perhaps sufficient effort is not being put into this development.

To exemplify, among basic issues which such a policy would need to address are the future intellectual needs of information systems developers, users, and managers; and the problem of adapting system development methodologies from "traditional" to intelligent systems. This is, in fact, the underlying problem we are considering in this paper, looking at the basic difference between the development of the two kinds of systems in terms of the required treatment of data.

4 Preparation of data for neural networks

It is assumed that the data preparation process here in question has been preceded by a thorough "system analysis" in the usual sense, resulting in some form of formal "conceptual schema" (Batini, *et al*, 1992), representing the "system" (organisation, firm, process, ...) considered. Such a schema is, as usual, the blueprint for the design of an information system, including its "intelligent" neural network modules. For those modules, however, a schema is only a starting point; their practical performance will essentially depend on the actual data in the schema. Such data will be used to determine the final architecture of the neural networks, and to train the networks. The data needs now to be interactively collected and tested, possibly leading to modifications in the original system schema.

The following general scheme for preparation of data for a neural network seems reasonable (Figure 1).

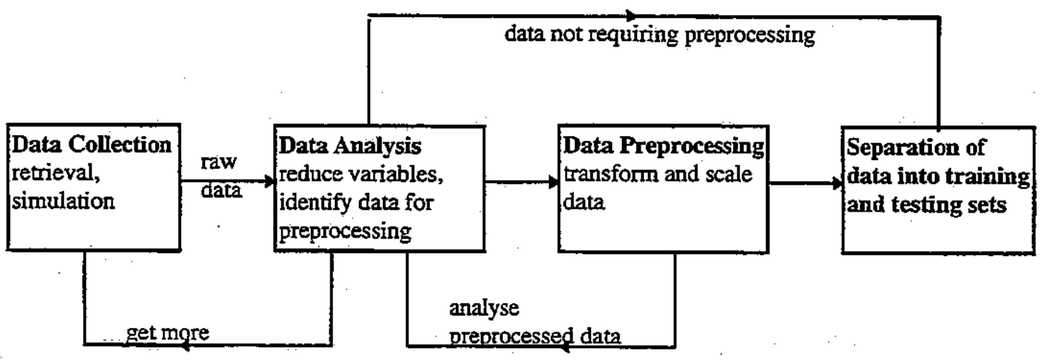


Figure 1: A data preparation scheme

Collected "raw" data is usually noisy, imprecise, and incomplete. Clearly, the training time and the resulting skills, ie the transfer mapping of the network, will depend upon the quality of the training data (Stein, 1993). The collected "raw" data should therefore be analysed before any training is attempted, and should, if possible, be preprocessed to improve its "quality".

The collected data is first analysed by, for example, statistical methods and/or neural network technology, to screen out the "irrelevant" data and to determine what preprocessing, if any, is required. The problem of analysis of the data or, more appropriately, determining the structure of the "analysis → preprocessing → analysis" loop, is a vast and virtually untouched problem, part of which is the determination of a neural network architecture suitable for the class of data to be processed. The "relevance" of data seems more basic, implying a particular need for the involvement of prospective network users and application domain experts in the system development process and especially in its data analysis stage. In this sense, the common perception that "neural networks learn by themselves and thus require little human expertise" is largely a myth.

It should be kept in mind, however, that until a comprehensive theory has been developed, any data preparation scheme will remain a trial-and-error affair to be adapted to each individual situation, guided by basic statistical principles, circumstantial evidence, general knowledge, and any applicable heuristic.

4.1 Data collection

As pointed out, data collection should be preceded by a thorough analysis of the problems so as to identify the variables and understand their relevance. The data should cover the widest range of the problem domain, including not only typical values but also exceptions and conditions at the "boundary" of the domain.

Financial analysis and decisions mostly involve historical data such as previous sales, exchange rates, stock prices, all ordinary index, and consumer price index. Such historical or "secondary" data (Davis & Cosenza, 1985) can be obtained by the retrieval method. The secondary data has been collected by others, such as governments, banks, corporations and universities, for their own purposes, and is published regularly in, for example, newsletters, government publications and bulletins, which are normally available in places such as universities, public libraries, corporations and government departments. The main advantages of the retrieval method are time and cost efficiency. However, the secondary data was not designed specifically to meet a particular problem, and we often cannot assess the accuracy of the data because we know little about the conditions under which its measurement took place, a check for accuracy, relevancy and reliability is still necessary. One method for checking accuracy is to cross-check different sources of data. If the data is obtained from a historically reliable source, it is likely to be reliable, while data from an 'expert' source is likely to be relevant.

Problems may arise when we intend to collect certain historical data such as previous sales and detailed transaction records from a firm. The required historical data may be incomplete in the firm. Collection of data from a firm needs cooperation from its management. Management may be reluctant to disclose data, particularly the data used in strategic decision making, as this data may be leaked to the firm's competitors. Computer simulation then provides a more suitable method of obtaining such kind of data.

Computers can simulate the operations of a firm and generate data to be used in substitute of the real thing. Simulation can generate data within a very short time, and the variables and operational conditions can be adjusted and controlled to produce a set of data well adapted to further processing. However, the operations of a firm are usually complex. It is, of course, not possible to simulate the firm's operations exactly, and simplifications have to be made. On the whole, simulation is a reliable and fast method to obtain data, and is particularly well suited to develop prototype systems.

4.2 Data analysis

Raw data may be noisy, poorly distributed and irregular. Though, in principle, neural networks can adapt to any kind of data, in practice, raw data has been found inappropriate for learning. High quality data for the neural networks improves its input-output mapping, and reduces learning time. However, high quality does not necessarily mean "detailed and precise". In everyday experience, we seldom need detailed data to reason about facts, events and problems. Very often, too much details are confusing and distracting, and may prevent us from organising raw data into useful patterns.

A large number of variables is usually involved in financial decisions. However, there may be correlations among some of these variables, and because of this, the actual number of variables to be considered can be less. Factor analysis is a statistical technique but does not involve sophisticated mathematics to reduce the number of observed variables, by summarising the information contained in the large number of variables into a smaller number of factors. It is done by constructing a matrix of intercorrelations between the variables, and then combining the variables into a new set of *factors* on the basis of relationships in the correlation matrix. The variables with high correlations among themselves are usually combined as one factor. However, the justification of the groupings should be checked by examining the factor loadings (ie correlation coefficients) between the variables and the factors. The grouping of the variables into a factor is justified if there is a high factor loadings, say greater than 0.55.

When human beings make decision, they normally interpret the actual values of the variables as some meaningful forms such as high or low, good or bad. If neural network has to simulate human decision making process, input to the network should better be a kind of rating scales rather than the actual values. For example, the $1,000 net cash flow is meaningless unless it is represented as a certain degree of high or low, which provides a more useful piece of information for decision making. Such variables should be identified for preprocessing.

All financial data can be presented as a collection of real-valued variables. Naturally, the variables would normally have different ranges and variabilities. During training, fluctuation in variables with large ranges will tend to 'overwhelm' the variables with small ranges. Further, variables with similar ranges need not have similar variability. Large variability of some variables would tend to 'distract' the neural network relative to a smaller but perhaps no less important variation in some other variables. In the treatment of a single variable, frequency distribution plot of data is a useful visualisation of its range and variability, and helps to determine whether the variable needs preprocessing.

4.3 Data preprocessing

Data preprocessing is essentially the implementation of the results of data analysis. It involves data transformation and scaling.

Often, the data is transformed using explicit ("elementary") mathematical functions of the original variables, eg logarithm and sine functions. Such functions may be used in combination to transform the whole collective domain of all the variables. The result of transformation can change the data range, or extract the useful information from the raw data. For example, in the analysis of time series, the statistical measures and differentiation techniques are used to extract the "trend" from the data. The variability of exchange rates can thus be measured by the coefficient of variation, and their direction of movement is determined by the first derivatives at specific points.

Scaling is a "procedure for the assignment of numbers to a property of objects in order to impart some of the characteristics of numbers to the properties in questions" (Phillips, 1971). One scaling method which is particularly useful for the continuous financial values is to convert them to "standard scores"[1], and use these as the basis of scaling. The "standard score" method makes the values independent of the units of measure, maps input to roughly equivalent ranges, in practice roughly between -3 and +3, and gives equal value to variations of equal rareness, regardless of the absolute range of the variation. This is particularly useful to minimise the influence of the absolute range and variability of one variable over the other. It is worth to re-analyse the scaled data, for example, by a frequency distribution plot, to examine its range and variability, and to revise the selected scaling method if necessary.

4.4 Separation of data into training and testing sets

It is important to assemble suitable sets of data for training and for testing the network. The data in both sets should be representative of the data on which the neural network will ultimately be used. Further, the training set should be large enough and so structured as to allow easy generalisation to data close by but not in this set.

There are no established rules to select training and testing sets. In practice, the selection seems to depend on the type of problem. In forecasting time series, such as the predictions of future stock prices (Baba & Kosaki, 1992) and foreign exchange rates (Refenes, 1993), both time-dependent, the testing set follows immediately the training set on the time-line. In credit assessment (Klimasauskas, 1993), the selection of training and testing sets has been based on the frequency of occurrence of each of the outcome categories. The automotive diagnostics classification problem (Marko, *et al*, 1990) uses a "leave-k-out" procedure, in which the network is trained and tested for multiple random (N-k, k) partitioning of the N available data sets.

5 Example: an intelligent information system in exchange rate risk management

Firms engaging in international trade are subject to risk from fluctuation of currency exchange rates. Although such risk can be actively managed with suitable hedging strategies, the problem is a complex one, involving large volume of financial data, which can be imprecise and noisy. Small firms rarely have access to required expertise to handle the problem. In a current project (Chan & Bonner, 1995), neural networks technology is used in combination with an expert system to determine a hedging strategy for the exposed net cash flow in a small import/export firm. The project is approaching the stage of a working prototype. We illustrate briefly with some of the experience gained in the treatment of data in the development of the prototype.

5.1 Data collection

The financial data, such as the interest rates, exchange rates, inflation rates, balance of payment, GNP, balance of trade and money supply, were retrieved from public sources such as the bulletins of The Reserve Bank of Australia and The Australian Financial Review. These sources contained

[1] Standard score is the number of standard deviations above or below the mean.

complete and reliable records of these financial data. The firm's data was generated in computer simulation of the firm's importing and exporting activities. As a small firm normally has an unpredicted number of sales contracts, the simulation of its importing/exporting activities thus involved an application of the Monte Carlo technique. The number of contracts obtained by the firm at a given point in the simulation was determined by a chance process described in the form of a probability distribution. The simulation model generated sales and order contracts, and calculated the revenue, expenses and balance every week over the entire period of two years of simulation.

5.2 Data analysis

Factor analysis was used to summarise the data collected by retrieval and simulation methods. A correlation matrix showing the correlation coefficients among all variables was constructed, and, using the principal component analysis technique that the correlated variables were combined as one factor, all the observed variables were then reduced to three factors, as shown in Figure 2.

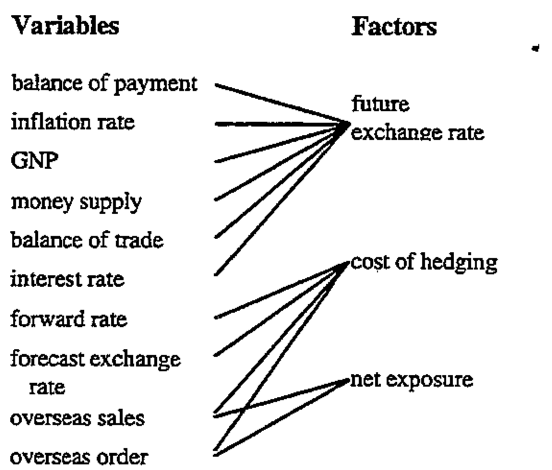


Figure 2: Factor analysis: reduce variables to a less number of factors

5.3 Data preprocessing

While it is sometimes impossible to predict future currency values with much accuracy, the firm can evaluate historical exchange rates in order to assess the potential change of the currency (Madura, 1992). The factor of *future exchange rate* was split into two variables: *currency variability* and *direction of currency movement*, the values of which were obtained by statistical and differential techniques. The standard deviation statistic was used to transform the historical exchange rates into *currency variability*, ie calculating the standard deviation of the historical exchange rates over a past period from a given point which is equivalent to the exposure period; and the *direction of currency movement* at a given point, x_0, was obtained by 'normalising' the first derivative of the line formed by past five data points, according to the formula given by Stein (1993)

$$\frac{dx_0}{dt} = \frac{25x_0 - 48x_{-1} + 36x_{-2} - 16x_{-3} + 3x_{-4}}{12|x_0|}$$

To be financially comparable, the values of *net exposure* were first converted to their present values. The values of the new variable *net exposure present value* and the *cost of hedging* were of little use in financial decisions unless they could be interpreted in a meaningful way (see Section 4.2). For example, a value of *net exposure present value*, say $1,000, could not tell whether the exposure was high or low, and a measure of exposure was related to the size of the firm. For this reason, both the *net exposure present value* and *cost of hedging* had to be scaled.

The first step of scaling the *net exposure present value* was to replace its values by their "standard scores", using the formula

$$\text{standard score} = \frac{x - \bar{x}}{\sigma}$$

where $\bar{x}$ is the mean of all x values,
σ is the standard deviation of all x values.

On basis of the "standard scores", the values were then scaled to the appropriate integers which were more useful than the actual values. The *net exposure present value* was scaled from 1 (lowest) to 4 (highest), indicating the degree of net exposure. Values less than the standard scores of -2 were scaled to 1, between -2 and 0, scaled to 2; between 0 and 2, scaled to 3; and greater than 3, scaled to 4.

The values of *cost of hedging* were scaled in a slightly different way. Firstly, the values were expressed as percentages of the net exposure. Then the percentages were replaced by the standard scores, and scaled in the same way as for the *net exposure present value*.

5.4 Separation of data into training and testing sets

After the data had been preprocessed, there would be four variables input to the neural network: *size of net exposure*, *size of cost of hedging*, *currency variability*, and *currency movement*. The output for the corresponding input was the *hedging decision*: hedged or not hedged. The input-output example pairs were divided into training and testing sets. The testing set was extracted by picking out every tenth example in the data set, in chronological order.

6 Conclusions

"A model is only as good as it is useful to people in thinking about, organising, and using data" (Tsichritzis, 1982). The development of an intelligent information system should be based on a knowledgeable assessment of its true potential and limitations. The understanding of the principles of "intelligence" is essential. An examination of typical such systems presently in use in the area of financial management suggests, however, that these requirements may not have been met in their development. In particular, neural network technology seems often to be used as substitute for thorough analysis of problems. The question "why neural networks?" remains all too often unanswered.

Though neural network technologies are, to degree, self-learning, compared to traditional information systems, both developers' and users' commitment to the development of intelligent information systems is not less vital. In particular, the technologies stand or fall with adequate data, which is not realistic to obtain without users' committed involvement. Further, like any (formal) system, a neural network system is a schema defined by its variables. The question *which variables are "relevant"?*, is thus no less important here than in the development of traditional information systems, and can only be settled with domain expertise. Another issue specific to intelligent system development is the particular need to ascertain the knowledge of the users, and the related questions of decision leverage control and user training. Summing up, a systematical approach to the development of intelligent information systems is lacking and is badly needed.

Whatever the underlying computational scheme, it is clear that the "intelligence" of computer-based information systems is continuously increasing. On the one hand, one can see new and exciting opportunities of extending analytical models with computer-based learning systems which are able to operate in a knowledge-rich environment, and are thus increasingly able to interactively solve practical and unstructured problems on the basis of partial knowledge. This, in particular, would suggest the border area of general financial theories and financial management practice as a very promising application domain. On the other hand, however, it seems rather optimistic to believe that the "intelligence" in the systems will take care of itself. There is a clear need for a basic policy on the development, use, and management of intelligent information systems, both in specific applications domains and in the society in general.

References

Arbib, M. A., and M.A. Hesse (1986) *The Construction of Reality*, Cambridge University Press, Chapter 3.

Baba, N. and M. Kosaki (1992) "An intelligent forecasting system of stock price using neural networks," *International Joint Conference on Neural Networks, (Baltimore, Maryland)*, vol 1, New York: IEEE, pp. 371-7.

Batini, C., *et al.* (1992) *Conceptual Database Design*, Benjamin/Cummings Publishing Co.

Bonner, R. F. (1993) "What can mathematics do for information systems?", *Proceedings 4th Australian Conference on Information Systems*, University of Queensland, St Lucia, Queensland, Australia, September.

Bonner, R. F, and R. Chan (1995) "A mathematical modelling perspective on information systems: basic concepts", working paper, School of Information Systems and Management Science, Faculty of Commerce and Administration, Griffith University, Australia.

Chan, A. C. K., and R. F. Bonner (1995) "A hybrid intelligent system in managing foreign exchange rate risk," to appear in *Proceedings of Symposium on Modelling and Control of National and Regional Economics* (IFAC '95), Gold Coast, Queensland, Australia, July.

Charness, N. (1981) "Search in chess: age and skill differences," *Journal of Experimental Psychology*, **7** (2), pp. 467-76.

Chase, W. G. and K.A. Ericsson (1981) "Skilled memory," in Anderson, J. R. (ed), *Cognitive Skills and Their Acquisition*, Hillsdale, NJ.

Collins, E., S. Gosh, and C.L. Scofield (1988) "An application of multiple neural network learning system to emulation of mortgage underwriting judgements," *International Conference on Neural Networks, Vol II*, IEEE, pp. 443-50.

Davis, D. and R.M. Cosenza (1985) *Business Research for Decision Making*, Boston, MA: Kent Publishing Company.

Dutta, S. and S. Shekhan (1988) "Bond rating: a non-conservative application of neural networks," *International Conference on Neural Networks, Vol II*, IEEE, pp. 443-50.

Hogeweg, P. (1993) "As large as life and twice as natural: bioinformatics and the artificial life paradigm," in Green, D. and T. Bossomaier (eds), *Complex Systems: from Biology to Computation*, Proc. 1st Australian Complex Systems Conf., IOS Press, Amsterdam.

Jagielska, I. (1993) "The application of neural networks to business information systems," *Proceedings of Australian Conference on Information Systems*, University of Queensland, Brisbane, September., pp. 565-574.

Klimasauskas, C. (1993) "Applying neural networks, Part II: A walk through the application process," in Trippe, R. R. and E. Turban (eds), *Neural Networks in Finance and Investing*, Chicago: Probus Publishing Company, pp. 54-63.

Koster, A., N. Sondak, and W. Bourbia (1990), "A business application of artificial neural network systems," *The Journal of Computer Information Systems*, 31 (2), Winter, 1990-91.

LeCun, Y., *et al* (1989) "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, 1 (4), Winter, pp. 541-51.

Li, J., J. Wu, and Y. Zheng (1991), "Neural network to support unstructured economic decision process," *International Joint Conference on Neural Networks*, 3, 18-21 November, Singapore.

Madura, J. (1992) *International Financial Management*, West Publishing Company.

Marko, K. A., L.A. Feldkamp, and G.V. Puskorius (1990) "Automotive diagnostics using trainable classifiers: statistical testing and paradigm selection," *International Joint Conference of Neural Networks (San Diego, CA), Vol. I*, New York: IEEE, pp. 33-8.

Meese, R.A. and K. Rogoff (1983) "The out-of-sample failure of empirical exchange rate models: sampling error or misspecification?" in Frankel, J. (ed), *Exchange Rates and International Macroeconomics*, Chicago: University of Chicago Press, pp. 67-112.

Mish, F. C. (ed) (1989) *Webster's Ninth New Collegiate Dictionary*, Springfield, MA: Merrian-Webster Inc. Publishers.

Newell, A. and H. Simon (1972) *Human Problem Solving*, Englewood Cliffs, NJ: Prentice-Hall.

Phillips, B.S. (1971) *Social Research Strategy and Tactics*, New York, The Macmillan Company.

Poggio, T. and F. Girosi (1990) "Networks for approximation and learning," *Proceedings of the IEEE, Vol 78 (9)*, September, pp. 1481-97.

Refenes, A. N. (1993) "Constructive learning and its application to currency exchange rate forecasting," in Trippe, R. R. and E. Turban (eds), *Neural Networks in Finance and Investing*, Chicago: Probus Publishing Company, pp. 465-93.

Rumelhart, D. E. (1989) "The architecture of mind: a connectionist approach," in Posner, M. I. (ed), *Foundations of Cognitive Science*, Cambridge, London: MIT Press, pp. 133-59.

Serra, R. and G. Zanarini (1990) *Complex Systems and Cognitive Processes*, Berlin: Springer-Verlag.

Simon, H. A. (1983) "Why should machines learn?", in Michalski, R. S., Carbonell, J. G. and Mitchell, T. M. (eds), *Machine Learning: An Artificial Intelligence Approach*, Palo Alto, CA: Tioga.

Sejnowski, T.J. and C.R. Rosenberg (1986) *NETtalk: A Parallel Network that Learns to Read Aloud*, John Hopkins University Electrical Engineering and Computer Science Technical Report (86/01).

Stein, R. (1993) "Preprocessing data for neural networks," *AI Expert*, 9, March, pp. 32-37.

Taylor, J.G. (1993) *The Promise of Neural Networks*, London: Springer-Verlag, Chapter 7.

Trippe, R. R. and E. Turban (eds) (1993) *Neural Networks in Finance and Investing*, Chicago: Probus Publishing Company.

Tsichritzis, D. C. (1982) *Data Models*, Englewood, Cliff, N. J.: Prentice-Hall.

VanLehn , K. (1989) "Problem solving and cognitive skill acquisition," in Posner, M. I. (ed), *Foundations of Cognitive Science*, Cambridge, London: MIT Press, pp. 527-79.

Wan, H.A., and A.K. Chan (1993) "A feedfordward neural network model for Hang Seng Index," *Proceedings 4th Australian Conference on Information Systems*, University of Queensland, St. Lucia, Queensland, Australia, September, pp. 575-86.

Zurada, J. M. (1992) *Introduction to Artificial Neural Systems*, St. Paul: West Publishing Company.

Enhancement of Degree of Generalization in Neural Clustering Systems and Statistical Interpretation

Jong U. Choi, Kyoung Y. Bae, & Dae W. Lee
Sangmyung Women's University

Yoon Chung
Hankuk University of Foreign Studies

ABSTRACT

Researchers in neural network community have claimed that the classification accuracy of neural network is higher or at least equal to conventional statistical methods. Based on the claims, many experimental research have been conducted and proved that the degree of generalization obtained in neural network's training scheme is better than other prediction/classification methods.

In a series of research conducted to develop an automated credit evalaution systems, various algorithms of neural networks have been experiemnted. In the experiments, back propagation algorithm and quickprop algorithms, with the increase of training data size, did not achieve classificaiton accuracy we expected. Therefore, learning vector quantization (LVQ) and generalized vector quantization (GLVQ) algorithms were employed to enhance the degree of generalization. Even though much better performance was obtained when GLVQ algorithm was employed, the prediction accuracy was still very low, around 50-60%.

The conclusion of this research is that the prediction accuracy of neural networks strongly depends on data characteristics and training environments. When the network algorithm is applied to noisy data set for training, that is, when there could not be found any meaningful statistical relation, higher accuracy might not be easily achieved. To the contrary, when there is a sound data set involved in the training, very high accuracy can be achieved.

1. Introduction

Neural-based training has been one of the hottest topics in AI research community. The technique has been successfully applied to various fields of classification: character recognition, signal processing, acoustic sound recognition, automobile licence plate recognition, business classification, and so on. For example, Kang et al. [1994a, 1994] applied neural network training to Korean character recognition in automobile licence plate and found that the accuracy rate of the system is much higher than that of an expert system. Even in business classification problem requiring experience of many years, the neural network systems have shown that the classification accuracies of the neural system are better than or at least equal to those of conventional systems or expert systems.

However, in a series of research on neural network application to credit evaluation system, it was found that the prediction(classification) accuracy strongly depends on the characteristics of the data included in training and testing. In other words, neural network approach to classification(prediction) is not a panacea to classification task, possibly applicable to every pattern classification problem. In fact, it is quite likely that neural-based training system, though many research results have claimed much better classification accuracy, can be worse than other classification techniques [Weigend and Gershenfeld, 1994]. In a series of experiments of Weigend and Gershenfeld [1994], conventional time series analysis technique showed better performance than neural-based systems in predicting future events with noise. To the contrary, in the cases of regular patterns, neural- based system showed much higher accuracy than others for predicting patterns.

In this research, the generalized learning vector quantization (GLVQ) algorithm [Pal, Bezdek, and Tsao, 1992] was applied to credit evaluation problem. GLVQ algorithm is a modified version of Kohonens' learning vector quantization (LVQ) algorithm [1991].