Proceedings of the IV Brazilian Conference on Neural Networks - IV Congresso Brasileiro de Redes Neurais
pp. 269-274, July 20-22, 1999 - ITA, São José dos Campos - SP - Brazil

A Neural Network Based Solution for the Credit Risk Assessment Problem

Germano C. Vasconcelos
Paulo J. L. Adeodato
Domingos S. M. P. Monteiro
Departamento de Informática,
Universidade Federal de Pernambuco,
Cx. Postal 7851, 50732-970, Recife-PE, Brazil
E-mails: {gcv,pjla,dsmpm}@di.ufpe.br

Abstract

The automation of decision making in financial markets is one of the major application areas of neural networks. Risk analysis is one of the problems where the technique has been efficiently applied. This paper investigates a solution to a credit analysis problem in a rather peculiar environment, characterized by a stabilized economy but subject to a high interest rate, namely the Brazilian market. A neural network based credit scoring system has been developed for the retail business in Brazil and its performance has been evaluated against that attained by a traditional discriminant analysis system. Extensive experimental results carried out with a database of 18,000 consumers of a leading Brazilian supermarket chain clearly indicate that a better solution is found with the connectionist based system.

1. Introduction

In the globalized economy, efficiency and low cost are fundamental aspects in making a product or service become competitive in the market. The huge volume of transactions turns the information processing automation into a crucial factor for cost reduction, high speed and high quality standards. The scientific and technological advance has modified the characteristics of managerial work. Automation has reached activities normally thought of as “reserved for humans” and many skeptics about this have changed their opinions as a result of the relevant successes achieved by state-of-the-art computer solutions applied nowadays. In the past, people tended to think that financial market analysis requires knowledge, experience and intuition and wondered how this activity could be automated. However, in developed countries the automation of financial market analysis has been steadily growing along with the scientific and technological advances [2]. In Brazil, the automation of this task was not possible until recently due to the high rates of inflation (at least 10% per month). This has only become feasible after the stabilization of the Brazilian economy with *Plano Real* which lowered inflation to less than 10% per year since 1994. Since then several financial institutions have already automated part of their decision support systems. But the demand for automation is still growing with new services being offered (e.g. personal leasing) and the need to expand the services to a broader market.

A major application of neural network [8,9,10] based systems to finance is risk analysis [2,3,4,7]. It involves the estimation of several aspects concerning a credit/insurance applicant: credit limit, expected profit, mean time for tardy payments, concession of credit, insurance coverage, etc.

This paper focuses on the problem of credit analysis particularly the decision on the concession of credit cards to customers [6]. A case study has been carried out on the actual database of a leading supermarket chain in Brazil which has its own credit card operator. The company had already automated the credit evaluation process and was interested in improving its performance and flexibility. This case of study is particularly interesting because of the peculiar aspects currently associated with the Brazilian market : a stabilized economy with low inflation rates (less than 1% per month) but with very high monthly interest rates (more than 5%). This has influenced the percentage of bad payers in the market. The system presented here has attained a better performance than the company’s credit score system based on linear discriminant analysis on a set of customer independent data retained at the company for evaluation.

The paper is organized as follows. Section 2 characterizes the credit concession problem. Section 3 particularizes the problem to the actual case investigated and Section 4 presents the solution adopted. Section 5 presents the results of the neural network based system and compares it to the existing credit score system. Finally, Section 6 presents final remarks on the limitations of the solution achieved, on the future work for improvement and on new desired features for such a system.

2. The Credit Concession Problem

When a person or a company asks for financial credit there are several ways to evaluate the application. Financial institutions define different cost functions to be optimized by the decision system. They may wish to maximize their profits, minimize the risks of non-payment or minimize the delays in repayments, etc. Nevertheless, the main goal is to decide whether or not to give credit to the applicant [7]. This paper will concentrate on this aspect of the problem. The block diagram in Figure 1 illustrates the process.

In the context of the credit evaluation problem described in this paper, the credit applicant supplies information through an application form which is checked by the institution for the veracity of some fields (input variables) such as social security number, annual income and permanent address (for personal loans). Additional financial information for that social security number is then obtained by the institution from credit security database bureaus (SPC, SERASA).

All the information about the applicant is fed to the credit concession system which awards a scalar score to that applicant. If the score is above a certain threshold defined by the institution, the application is approved; otherwise, it is refused [1].

After the credit concession, the institution keeps track of the applicant's financial transactions and, according to the institution's criteria, labels the customer as good or bad payer [3].

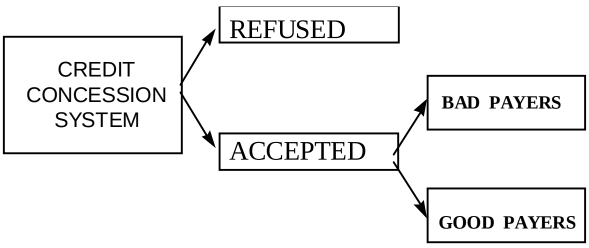

Figure 1

As a technique based on statistics, neural networks [1,8,9] requires a large database to extract knowledge from. However, usually this is not a problem once that before automating the process, the institutions always make decisions relying on human expertise. So at the time of the implementation of the automatic systems, the institutions already have a large database with the history of their customers and their classification as debt payers according to the institution's policies.

The large database provided for extraction of knowledge, however, is biased in the sense that only the applicants who were given credit are followed up as debt payers. Therefore only one type of decision error can be measured in such a system: the concession of credit to bad payers. Considering that the credit concession system maps a multidimensional space into a scalar value (score), the types of error caused by the thresholded decision policy can be illustrated by Figure 2.

Figure 2 shows the hypothetical class-conditional probability density functions for good (G) and bad (B) payers plotted against the score obtained in the credit concession system implemented at the institution. The threshold (L_B) defined by the policy of the institution determines the decision error regions. The area of region E_{B2} measures the error of type-II correspondent to bad payers who received credit whereas the area of region E_{B1} measures the error of type-I correspondent to good payers who have not received credit. Errors of type-I are not observable because of the bias in the database.

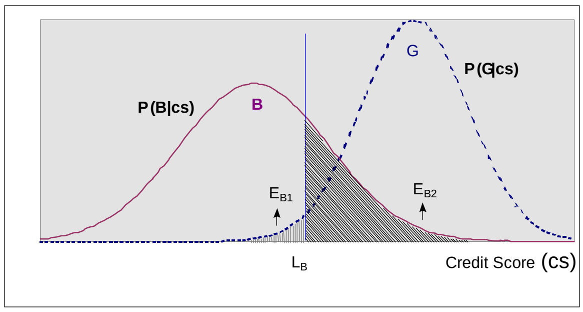

Figure 2

If another credit concession system is to be compared to the existing one for the biased database, two types of error can be measured. This can be illustrated in the previous plot with another threshold (L_R) for the new system. Figure 3 presents the types of error that occur when the new threshold (L_R) is higher than the existing one (L_B). The area of region E_{R2} measures the error of type-II correspondent to bad payers who would receive credit. The area of region E_{R1} measures the error of type-I correspondent to good payers who would not receive credit but were good payers and were followed up (observable error).

[1] The rejection score interval corresponding to cases solved by human experts has been omitted from this presentation

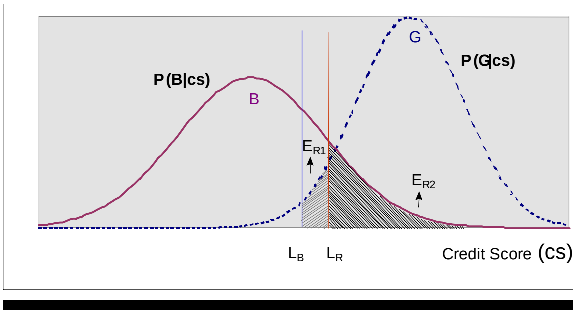

Figure 3

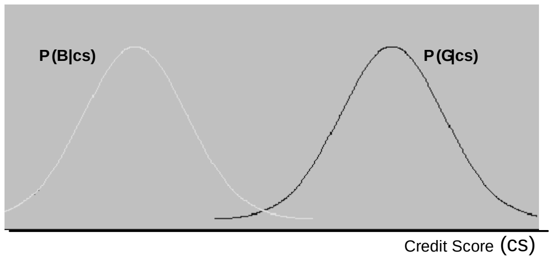

Figure 4

One way of evaluating the performances of the systems is to compare the sum of the observable areas of error regions (E_{R1} + E_{R2}) of the proposed system (Figure 3) against that (E_{B2}) of the existing system (Figure 2). The system with smaller error would have the better performance. Thus, the *advantage* of the new system can be defined as $\Delta = E_{B2} - (E_{R1} + E_{R2})$ in the comparison. This simple measure does not take into account the cost of the types of error despite being well known that the cost of the error of giving credit to a bad payer (type-II) is much higher than that of not giving credit to a good payer (type-I).

There are hypothetically two ways of achieving a performance better than that of the existing system. The first is to assume that the new system maps the input variables into a scalar producing class-conditional probability density functions similar to those produced by the existing system also assuming that the existing threshold is smaller than the optimal decision point (intersection of the curves). The new system could then have a higher threshold and reduce the overall error as illustrated in Figure 3. This situation however, may not be the case; the threshold may be beyond the optimal decision point, not giving credit to much more good payers than it should.

The second and more plausible solution is that the non-linear mapping of the neural network based system (as opposed to the linear mapping of the linear discriminant analysis based system) produces class-conditional probability density functions more separated than those produced by the existing system (with less area of intersection under the curves), thus reducing the decision errors for thresholds near the optimal value. Figure 4 illustrates this situation.

3. The Actual Case Study

The actual problem to be solved was to evaluate the performance of the neural network-based system on the retailer's database for applicants who had received credit in recent years (after *Plano Real*). All of them had been approved by the system based on linear discriminant analysis. This financial institution is a credit card company with 500.000 customers.

The database consisted of a set of 18,000 customers from which 6,000 had their labels withheld by the institution for performance evaluation. This apparently large database turned out to be small since from the 12,000 labeled examples available for development, there were only 450 customer labeled as bad payers by the institution criteria. This strong unbalance in the classes (3.75% of bad payers and 96.25% of good payers) required a maximum performance error of 3.75% from the proposed system.

The unbalance in the classes also suggested that the institution's system was operating with a threshold beyond the optimum level; that is, to the right of the intersection of the distributions in Figure 1. But that could not be proved because the was no follow up on the applicants with credit refused.

The system had also to cope with other features which make the problem slightly more complex. There are missing data. Some fields in the application form need not to be filled (non-obligatory). There are also several non-numerical fields in the form the applicants fill. The truth of the information on the form is only verified for particular fields.

4. The Solution Adopted

Given the complexity of the problem and the high accuracy needed, the system could not be trivial to achieve the expected goal.

4.1. Data preparation

The data of this case of study contains the features generally found in data mining problems. Therefore it required several levels of data preparation: data cleansing, statistical data analysis, number scaling and attribute encoding [5,7].

The first problem encountered on the data was the lack of standard filling of attribute fields such as city, borough, profession etc. requiring a long non-automated data cleansing.

Another difficulty faced was missing data which is inherent to this problem once that some fields in the application form need not to be filled (non-obligatory). Statistical analysis helped on completing these blanks.

The reliability of the data is assured only for a few particular fields of the application form which have their veracity confirmed through documents.

There are also several non-numerical fields in the form the applicants fill. Random variables converted the attributes into numerical values in a way to optimize the data representation for the classifiers. The encoding was typically *1*-out-of-*n* for fields with few options (e. g. marital status). For fields with many different ways of filling (e. g. town, profession etc.), statistical analysis helped in grouping all rare events in a single value. In this case, the *k*-out-of-*n* encoding was used; a compromise between the bad distance mapping of the dense binary code and the large dimension of the *1*-out-of-*n* code.

Some fields (e. g. time in the present job: years and months) needed to be combined for expressing their information in a simpler way (months).

Numerical data were all scaled to the fall in the interval between zero (*0*) and one (*1*).

4.2. The architecture

Considering that the data made available by the company had only a small fraction of bad customers (only 3.75% of the total) and that no information about the refused applications was supplied, the system has been developed as a 2-stage classifier described below and shown in Figure 5.

The first stage consisted of a k-nearest neighbors-like (k-NN) classifier to separate credit applicants in two groups: applicants "close" to the bad payers go for further evaluation whereas those "far" from bad payers are classified as good applicants to receive credit. The output threshold of the this classifier can be adjusted according to the institutions policies to accept or refuse credit applications.

The credit applicants clustered around the bad payers then pass through a finely tuned Multi-Layer Perceptron (MLP) dedicated to the separation of a more restricted set with a much more balanced expected ratio of good/bad payers.

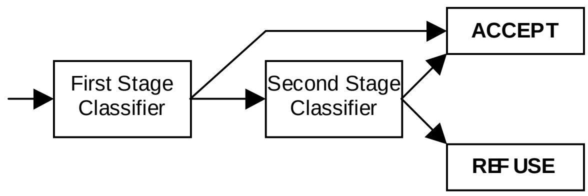

Figure 5

4.3. Training and testing

The system was trained on 6,000 examples and tested on the remaining 6,000 from the 12,000 made available to the developers.

5. Results

The results were presented to the institution in the worksheet supplied to the developers with the 6,000 unlabeled customers. For each customer, a classification label was given indicating whether or not that customer should have received credit. To reach this binary decision, the test set consisting of 6,000 examples was used for the purpose of setting the threshold of the first stage classifier. The goal was to maximize *advantage* of the system in terms of number of applications (as defined in section 2).

Figure 6 shows the process of selecting the threshold based on the test data. As the threshold increases, the systems starts detecting bad payers in the set but soon the system starts refusing credit to good payers. The aim is to optimize the difference between the number of bad payers detected and the number of good payers with credit refused. The threshold for this optimal net number of customers in the test set was then used for the deciding about the 6,000 customers on the set retained by the institution. In fact 3 different thresholds were chosen to account for statistical variations on the database once that the number of bad payers in the sample was too small.

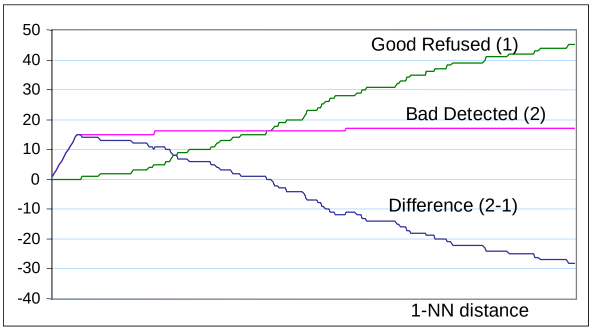

Figure 6

The behavior observed in Figure 6 suggests that the neural network mapping into a scalar separates the class-conditional probability density functions of bad and good payers from the biased database and that it allows the threshold to be placed either to the left or to the right of its optimal value.

Table 1 below shows the performance (Det = bad detected, Ref = good refused, Dif = difference between the two) of the system on both test set and the data set retained by the institution for the 3 thresholds chosen. The proposed system performed well both on the 6,000 examples retained by the developers and on the 6,000 retained by the institution. Should the cost of the errors be taken into account (as mentioned in section 3), the performance of the neural network based system would have been far better than has already been.

Table 1: System Performance

	Test Set			RetainedSet		
	Det	Ref	Dif	Det	Ref	Dif
Threshold-1	10	0	10	17	0	17
Threshold-2	15	0	15	23	2	21
Threshold-3	17	6	11	25	6	19

After the institution's evaluation of the neural network based system, the developers had access to the actual status of the payers in the retained data set. Figure 7 shows the results obtained on that data set in the same way as shown in Figure 6 for the test set. The similar behavior observed confirms the power of the solution offered and the statistical representativeness of the sample.

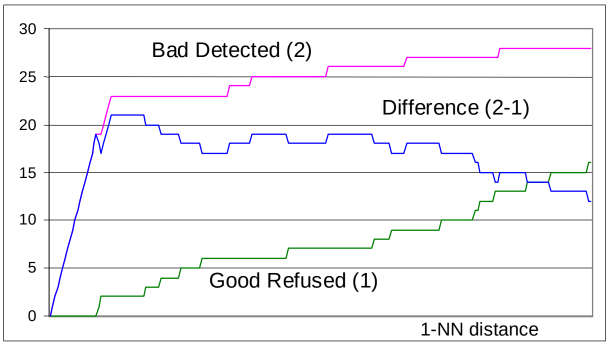

Figure 7

An interesting feature of the new system was observed some months later. Several customers who had been mistakenly classified as bad payers by the new system (classification errors) became bad payers after the initial evaluation of the system. This power could be better analyzed to predict when a customer is going to turn into a bad payer if more information is made available.

6. Final Remarks

This article has presented a neural network based system for automatic support to credit analysis in a real world problem. The proposed system reached a higher performance than the existing one based on linear discriminant analysis despite the huge unbalance among the classes which required a maximum of less than 4% classification error from the proposed system. That was not an unexpected result considering that k-NN and neural networks implement non-linear mappings whereas the existing system is constrained to linear ones.

The proposed system is also better in terms of flexibility for adapting to gradual changes in the applicants' behavior during the usage of the system. The existing system, however, is still important for explaining the reasons for the decisions made and for adapting to drastic changes in the economy (economic law changes, stock market crashes, etc.).

Since the proposed system achieved a performance better than the credit score system, the institution is going to operate both systems in parallel giving credit to the applicant whenever one of the systems decides for it and the other does not refuse strongly. Thus, a much better performance evaluation will be carried out. Also, an automatic combination of both systems through a neural network can be used for a better decision making where the best of each system will be taken into account.

It should be emphasized that the second condition "*...and the other does not refuse strongly.*" for parallel operation is crucial for the neural network based system because it has not been trained on any data from applicants with credit score below the threshold. This guarantees that the company's credit score system will prevail in those cases. The proposed solution has only been tested in a cascade combination of the two systems. The developers have asked the institution to create a database for storing data about the refused applications for them to develop a more general and powerful neural network based system under the unsupervised learning paradigm.

References

[1] D. Rumelhart, B. Widrow, and M. Lehr, The basic ideas in neural networks, *Communications of The ACM*, Vol. 37, (**3**) 1994.

[2] *The Economist*, pp.75-77, 15 April 1995.

[3] E. Gately, *Neural Networks for Financial Forecasting*, John Wiley & Sons, Inc., 1996.

[4] A. Ainslie and X. Drèze, *Data Mining: Using Neural Networks as a Benchmark for Model Building*, 1996.

[5] K. M. Decker and S. Focardi, *Technology Overview: A Report on Data Mining,*. CSCS-ETH, 1995.

[6] W. K. Schrickel *Credit Analysis,*. Atlas, 1995 (in Portuguese).

[7] J. P. Bigus, *Data Mining with Neural Networks*, McGraw-Hill, 1996.

[8] S. Haykin, *Neural Networks: A Comprehensive Foundation*. McMillan College Publishing Company Inc, 1994.

[9] R. Beale and T. Jackson, *Neural Computing: An Introduction*, Adam Hilger, IOP Publishing Ltd, 1991.

[10] D. Rumelhart, B. Widrow, and M. Lehr, Neural networks: applications in industry, business and science, *Communications of The ACM*, Vol. 37, (**3**) 1994.