Corporate Credit Rating Revisited: A Quantitative Approach Based on Neural Networks

Carlos Henrique C. Duarte[1,2*] and Ináh C. P. Garritano[2†]

Abstract— **In this paper, we present a new quantitative approach for assessing corporate credit risk based on artificial neural networks. We introduce the problem, present a credit rating model established in terms of multi-layer perceptron networks, describe a prototype system and finally justify our data selection process for training and benchmarking our model and the developed system.**

Keywords— **Credit Rating, Corporate Finances, Quantitative Analysis, Artificial Neural Networks.**

I. Introduction

Borrowing and lending are perhaps as old activities as commerce itself. With the passing of centuries, these activities have been sensibly facilitated, not only by the adoption of nationwide local currencies but also by the establishment of banks and other financial institutions where the stock of money and valuable goods could be treated in more uniform and secure ways.

The provision of credit in the form of loans or through the subscription of tradable debt instruments, such as credit bonds, is not carried out without associated risks. To understand these risks, it is important for a financial institution to assess, before granting certain amount of money to a credit customer, many aspects of the potential client, including: financial and operational risks; paying back capability; and the quality of any available collateral. Information gathered in these analyses will help establish corresponding interest rates, which are meant to cover at least operating costs and eventual losses of the institution.

The process of assessing the relative trustworthiness of a credit customer in servicing debts in a timely manner is called credit rating. This kind of risk analysis is performed by specialists, who may sometimes use incomplete or inconsistent data to carry out their work. Continuously maintaining such classifications is so important for the health of financial institutions that the Basel Committee on Banking Supervision and local supervisory authorities, such as the Brazilian Central Bank (through Resolution 2.682 [1]), require that all credit operations be internally or externally rated according to consistent and verifiable criteria. For customers, on the other hand, this process contributes to standardize and facilitate credit decisions.

In the past century, international credit rating agencies — most noticeably Standard and Poors, Moody's and Fitch — were established to supply companies and entire countries with such classifications on demand. The outcome of their work is a public grade, in the form of a sequence of letters and other symbols (such as A+ or A1), stating the perceived safety in credit contracting relative to other similar institutions. Because they were created precisely to serve as a market reference, these agencies tend to provide ratings with little distinction among themselves [2].

Credit rating methods are not widely understood. Concerning the rating of non-financial companies, institutions usually rely on quantitative data such as financial statements and projected cash flows (including some calculated financial ratios), as well as on qualitative data like auditing company remarks and adopted management procedures.

In this paper, we present a new quantitative approach for estimating the credit rating of non-financial companies based on artificial neural networks. In particular, we propose a new rating model based on multi-layer perceptron networks trained with a back-propagation algorithm and present a user-friendly prototype system which is meant to assist specialists of financial institutions and rating agencies in conducting their credit rating analyses.

In Section II, we present our credit rating model. Next, in Section III, we describe the developed prototype system. Finally, in Section IV, we justify our data selection process for training and benchmarking the system. The final section is devoted to analysing our results and presenting some prospects for future research.

II. A Corporate Credit Rating Model

Credit rating models have been proposed to attempt to capture the rationale of specialists in establishing classifications for different companies. This rationale can be represented as a function, f, that describes how the produced outputs are obtained from input data.

For the sake of defining the output of this kind of function, it is important to recall that the different classification systems adopted by the international credit rating agencies are comparable, as illustrated in Fig. 1 (extracted from [2]). The last column of the presented table contains a numerical representation of each risk category, which is represented as a different sequence of symbols by each agency. The table also shows that only grades of companies operating in the same country, even when in different economy sectors, are comparable, due to sovereign risk discrepancies between different countries.

Adopting the numerical conventions in the table and assuming that n input variables are necessary in the rating process, f can be expressed as a function with the following

[1] BNDES, Av. Chile 100, Rio de Janeiro, RJ, Brazil, CEP 20001-970, e-mail: `cduarte@bndes.gov.br`.

[2] Universidade Universidade Estácio de Sá, Rua do Bispo 83, Rio de Janeiro, RJ, Brazil, CEP 20261-902.

* E-mail: `carlos.duarte@computer.org`

† E-mail: `inahcp@hotmail.com`

signature:

$$f : \mathbf{R}^n \to \{x \in \mathbf{N} | 1 \leq x \leq 24\}$$

Each kind of credit customer may have its relative repayment capabilities approximated by different variables. If the credit is to be provided to a person, for instance, its monthly salary and expenditures would be among the most relevant proxies. Here, since we focus on the provision of credit to companies, what matters most appears to be their capability of generating results and their current debts.

We adopt as input variables some ratios, r, derived from the financial statements of each company. Since the choice of ratios covering results, debts and their relationships is a matter of personal taste [3], we use the following ones:

1. short term debt ratio = short term debts / total assets;
2. total debt ratio = total debts / total assets;
3. return on assets = net income / total assets;
4. current liquidity = current assets / current liabilities;
5. liquidity = (current assets + long term receivables) / (current + long term liabilities);

Although the ratios above can be used to provide a gross approximation of the relative repayment capability of a company, there are other variables that affect this analysis. For instance, these ratios may vary over time. In order to capture the historical evolution of a ratio r_i, $1 \leq i \leq 5$, we also use in our analyses differences, d_i, between the observations of r_i in each pair of consecutive years (as suggested in [4]). The existence of other companies in the same economic group may also affect a rating analysis. Whenever a company belongs to a group, we take into account only the consolidated financial statements of that group.

It is important to mention that all selected input variables (even the consecutive year differences) are ratios defined relative to the assets and liabilities of each company. We adopt this convention because: (i) credit rating is not dependent on company size; (ii) the manipulated numerical amounts are all directly comparable; and consequently (iii) it reduces the rounding error in the evaluation process. To confirm the first observation, just note that the size of the desired credit operations are never considered in the analysis.

A function approximating a credit rating model, f^*, can be expressed in many different ways, such as using stochastic analysis [5], fuzzy logic [6], or artificial neural networks [7]. In this paper, we focus on the last possibility. This means that we are interested in determining first the most appropriate network architecture for solving our problem and then provide a set of samples so that the network can be trained to classify corporate credit rating patterns simulating the behavior of a specialist.

In order to choose the network architecture better suited to solving our problem, we are obliged to understand first its mathematical nature. Many classification problems are of linearly separable nature and therefore can be solved using so called perceptrons by simply adjusting the weights that simulate the behavior of a neurone in a two layer network. Such linear functions can be viewed as hyperplanes that separate observation points in the n-dimensional space. It is easy to see that our problem cannot be effectively approximated using such constructions. To see why, simply note that in the absence of debts the return on assets ratio may grow arbitrarily, whereas there is an upper limit for company rating, namely that corresponding to the highest rating grade (corresponding to 24).

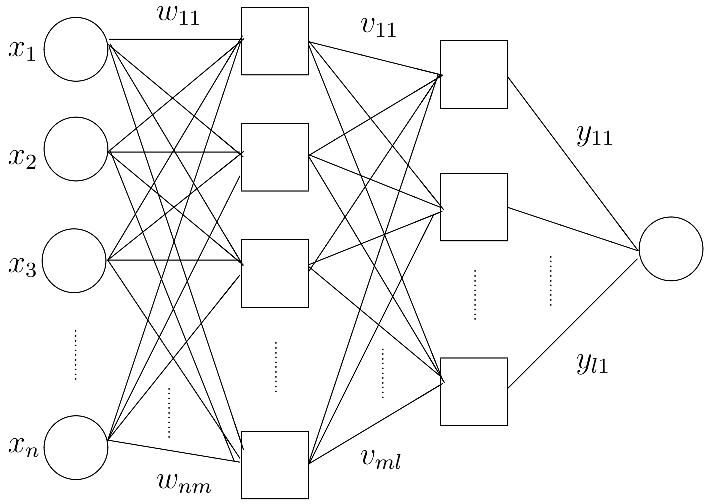

Fig. 2. Four layer perceptron network architecture

The analysis above shows that our problem is of non-linear nature and, in particular, can be represented as a discrete result folded hypersurface. We thus adopt as an architecture for solving our problem four layer perceptron networks, since the literature establishes that any such surface can be approximated in this way (see [8] referred in [9]). This kind of network is illustrated in Figure 2.

A multi-layer perceptron network, apart from the input and output layers (represented by circles in our figure), may also have hidden layers (as illustrated by the boxes in the figure). It is part of the modeling process to adjust not only the weights that relate neurones in these layers but also the number of hidden neurones. Fortunately, there is a standard back-propagation algorithm that helps fixing all these variables during the network training phase.

Using the terminology of Figure 2, f^* can be defined based on a given non-linear activation function g as follows:

$$f^*(\vec{x}) \stackrel{\text{def}}{=} \sum_{k=1}^{l} y_{k1} g \left(\sum_{j=1}^{m} v_{jk} g \left(\sum_{i=1}^{n} w_{ij} x_i \right) \right)$$

We wish to approximate f using f^* so that the resulting error can be considered negligible. This means that, after having trained our network, we expect to obtain evidence that the following formula is satisfied for a small enough ϵ:

$$\exists \epsilon \in \mathbf{N} \cdot \forall \vec{x} \in \mathbf{R}^n \cdot |f(\vec{x}) - f^*(\vec{x})| \leq \epsilon \qquad (1)$$

Here, we understand by a small enough ϵ a value that is smaller than the standard size of each credit rating category in Figure 1, that is, less or equal than two.

CREDIT RATING CATEGORIES	MOODY'S	STANDARD & POOR'S	FITCH-IBCA	GRADE
Strongest borrowing capability and smallest possibility of credit loss in relation to other local borrowers	Aaa	AAA	AAA	24
Very strong borrowing capability and low possibility of credit loss in relation to other local borrowers	Aa1 Aa2 Aa3	AA+ AA AA-	AA+ AA AA-	23 22 21
Borrowing capability above average in relation to local borrowers	A1 A2 A3	A+ A A-	A+ A A-	20 19 18
Average borrowing capability in relation to other local borrowers	Baa1 Baa2 Baa3	BBB+ BBB BBB-	BBB+ BBB BBB-	17 16 15
Borrowing capability below average in relation to other local borrowers	Ba1 Ba2 Ba3	BB+ BB BB-	BB+ BB BB-	14 13 12
Weak borrowing capability in relation to local borrowers	B1 B2 B3	B+ B B-	B+ B B-	11 10 9
Speculative and showing very weak borrowing capability in relation to other local borrowers	Caa1 Caa2 Caa3	CCC+ CCC CCC-	CCC+ CCC CCC-	8 7 6
Very speculative and showing borrowing capability extremely weak in relation to other local borrowers	Ca	CC	CC	5
Extremely speculative and showing the weakest borrowing capability in relation to other local borrowers	C	C	C	4
In default		D DD	DDD DD D	3 2 1

Fig. 1. International Credit Rating Agency Classification Standards.

III. Developing an Automated System

The requirements of an automated system to support credit rating analyses can be roughly summarized as supporting, for each company, the input of an yearly indexed sequence of financial statements (comprising their assets, liabilities and statements of results) in a standard format (say GAAP) so that, if possible, a suggested credit rating for that company is produced afterwards. This classification should be validated by specialists using their past experience and other qualitative data.

Since we decided to model our problem using artificial neural networks, the implementation of an automated system should support two operation modes, one for training the network and another for producing rating classifications. Moreover, independently of the operating mode, the treatment of company data should be subject of pre and post processing, since it is necessary not only to validate the given statements and compute the financial ratios derived from them but also to provide the end user with the final classification written in terms of one selected credit rating agency terminology.

We have implemented a prototype system to satisfy all the requirements and decisions above. The object-oriented approach was adopted during the whole development process, which took nearly one year and was divided in two phases: the production of a detailed specification using UML [10] and the implementation of the specifications using SQL, ODBC and the Java language [11]. With this architecture, it would be possible to use the system in any personal computer and network endowed with standard database management and operating systems.

Central to our implementation is the use of an offline static back-propagation algorithm for training the network. The role of the algorithm is to determine the weights that relate neurones in different network layers (namely y_{k1}, v_{jk} and w_{ij}). We adopt a standard activation function in the implementation of this algorithm:

$$g(x) \stackrel{\text{def}}{=} \frac{1}{1 + exp(-2x)}$$

Since the algorithm requires that the entire computation be continuous, we work in this way but round the network final output to obtain integer results according to Figure 1. The initial network configuration presumes that $l \leq m \leq n$ and that layers are fully connected, but some edges and even neurones may be disconnected while training by equalising some weights to zero.

The back-propagation algorithm operates in cycles, each one divided in two phases: the forward phase, when the network outputs and error are calculated for a given training sample, S, and the corresponding expected results; and the backward phase, when the weights that relate network layers are adjusted based both on a learning ratio and on the obtained quadratic training error, E. This error is defined as follows in our case:

$$E \stackrel{\text{def}}{=} \frac{1}{2} \sum_{\vec{x} \in S} (f(\vec{x}) - f^*(\vec{x}))^2$$

Since the quadratic training error is presumed to decrease during a network training process, we choose to implement the termination of such processes after reaching the allowed maximum number of cycles or a negligible error.

It happens that the algorithm does not behave well in the neighborhood of singularities in the surface being approximated. In the credit rating space, such singularity points appear not only around the mean value of the classification space, due to its sigmoidal character, but also in the entire frontier where two strong components for obtaining the result conflict. These strong components are determined, on the one hand, by the debt ratios (1 and 2), which make rating decrease as they increase, and, on the other hand, by the results and liquidity ratios (3, 4 and 5), which make rating increase as they do. In order to treat this problem, we adopt the standard approach of using a momentum term to increase the pace in the computation of adjusted weights, so that the algorithm instability in the neighborhood of singularity points is reduced [9].

IV. Training and Benchmarking

One of the worst difficulties in addressing real problems using artificial neural networks is the limited availability of training data. Fortunately, in the way we have addressed the credit rating problem, we can overcome this difficulty, since classifications are made public by the credit rating agencies and, for companies listed in open stock markets (in general, those which have a public credit rating), their financial statements are also made publicly available due to supervisory agency rules.

In order to train and benchmark our model and system, we selected a sample, P, consisting of 36 companies listed in Bovespa stock market according to CVM[1] rules. Their yearly financial statements can be found in CVM's home page (`http://www.cvm.gov.br`), whereas their credit rating is frequently made public in the classification agencies web sites (see, for instance, `http://www.fitchratings.com`). For each company, the last three available financial statements are used in the credit rating analysis.

A statistical profile of our population is presented in Figure 3. As it can be seen, ratings were selected from two distinct agencies and companies spread across many economy sectors. Although this data set is not equally distributed among the studied categories and the size of the set is not ideal, overall it does not appear to be biased, since it represents a faithful sample of the whole Brazilian economy, containing real companies among those that might require a credit rating analysis in all representative economy sectors. To confirm this, suffices it to say that our sample consists in more that one third of the whole universe of Brazilian companies rated by the three international credit agencies in the past two years.

[1]CVM is the Brazilian Securities and Exchanges Commission.

DIFFERENCES MODEL	15.15.7.1	%	15.4.1	%
$\lvert f(\vec{x}) - f^*(\vec{x})\rvert > 2$	14	38.9	18	50.0
$\lvert f(\vec{x}) - f^*(\vec{x})\rvert = 2$	4	11.1	4	11.1
$\lvert f(\vec{x}) - f^*(\vec{x})\rvert = 1$	10	27.8	7	19.4
$\lvert f(\vec{x}) - f^*(\vec{x})\rvert = 0$	8	22.2	7	19.4
Training error		6		6
Error mean value		3		4
Error standard deviation		3		4
Quadratic network error		336		566

Fig. 4. Network outputs considering ratio difference inputs.

GROWTH MODEL	15.9.2.1	%	15.11.1	%
$\lvert f(\vec{x}) - f^*(\vec{x})\rvert > 2$	8	22.2	10	27.8
$\lvert f(\vec{x}) - f^*(\vec{x})\rvert = 2$	6	16.7	1	2.8
$\lvert f(\vec{x}) - f^*(\vec{x})\rvert = 1$	11	30.6	6	16.7
$\lvert f(\vec{x}) - f^*(\vec{x})\rvert = 0$	11	30.6	19	52.8
Training error		1		1
Error mean value		2		2
Error standard deviation		3		3
Quadratic network error		228		254

Fig. 5. Network outputs considering ratio growth inputs.

We partition our population into two classes of equal size, preserving their distribution among sectors and rating categories as presented in the table. One of them is used in the training process, S, while the other one serves to test and validate our results, V, allowing us to study the generalization capability of our neural network. Initially, we adopt as network inputs the ratios r_i, together with their historical evolution over the last three years in the form of yearly differences, d_i^1 and d_i^2 respectively, for $1 \leq i \leq 5$. In this way, we have to provide 15 inputs in order to obtain a neural network credit rating evaluation.

Considering that our problem is of non-linear character, whose treatment may require two hidden layer networks, we arrange such possible network configurations as a grid and adopt an exhaustive search algorithm in order to determine the best performance configuration. The respective network is selected by yielding the smallest quadratic output error. The best performance one hidden layer network is determined using an analogous process.

The best performance networks output considering our original credit rating model are presented in Figure 4. As it can be seen, both one and two hidden layer networks perform badly in dealing with our sample data. The figure suggests that these networks are not even capable of learning data patterns during the training process. To see this, just notice that the smallest training error is equal to 6, the size of two rating categories of Figure 1.

Instead of adopting differences to take into account the historical evolution of the ratios proposed in Section II, we attempt to obtain better results by calculating their

RATING ISSUER	$\lvert P\rvert$	%	ECONOMY SECTOR	$\lvert P\rvert$	%
Agency A	18	50.0			
Agency B	18	50.0	Consumer Goods	2	5.6
TOTAL	36	100.0	Energy	8	22.2
RATING DISTRIBUTION			Forest & Paper	3	8.3
$1 \leq f(\vec{x}) \leq 6$	2	5.6	Logistics	2	5.6
$7 \leq f(\vec{x}) \leq 12$	0	0.0	Mining	2	5.6
$13 \leq f(\vec{x}) \leq 18$	18	50.0	Oil & Gas	2	5.6
$19 \leq f(\vec{x}) \leq 24$	16	44.4	Steel	4	11.1
$f(\vec{x})$ VALUES			Telecommunications	6	16.7
Mean value	18		Other	7	19.3
Standard deviation	5				

Fig. 3. Profile of the adopted data set.

SIMPLE MODEL	15.11.4.1	%	15.4.1	%
$\lvert f(\vec{x}) - f^*(\vec{x})\rvert > 2$	13	36.1	7	19.4
$\lvert f(\vec{x}) - f^*(\vec{x})\rvert = 2$	5	13.9	7	19.4
$\lvert f(\vec{x}) - f^*(\vec{x})\rvert = 1$	7	19.4	11	30.6
$\lvert f(\vec{x}) - f^*(\vec{x})\rvert = 0$	11	30.6	11	30.6
Training error	3		1	
Error mean value	2		2	
Error standard deviation	2		3	
Quadratic network error	176		166	

Fig. 6. Network outputs considering some ratio growth inputs.

growth. In this case, the networks show some learning capability, but their generalization capability is limited. More that 70% of the companies of our validation set, when analyzed by the best performing network, fall into the upper error category. These results are summarized in Figure 5.

Since there is some overlapping between the characteristics detected by the historical evolution of the liquidity and debt ratios, and between the debt ratios themselves, which could potentially cause difficulties during the network training process for reaching the smallest possible training error, we give up using the historical evolution of the first of these ratios, as well as that of short term debts. In this way, instead of having to provide 15 inputs for obtaining each classification, only 9 are required. As a result, we obtain a generalization capability far superior than in the other models, with only 33,33% of the validation companies falling into the upper error category. The best performing networks using this simplified model are presented in Figure 6, about which it is important to point out that the 15.4.1 network can already be used for practical purposes.

Nevertheless, it is still reasonable to discuss the reason for having 6 companies classified by the best performance networks (totaling the 33% validation population of the upper error category) in a rating category farther than expected. Doing a case by case analysis, we found out that these were either companies classified by agencies in transitory grades, potentially waiting for a substantial upgrade or a downgrade, or companies in the lower half of the rating scale, for which there is not considerably more real data to better train our networks.

Our experiments suggest that the conclusions above are robust, in the sense that they are not affected by the size of the studied sample. In fact, varying the sample size from 16 to 36 companies, we observed that the results of the differences model are worsened, the growth model output quality remained stable, and there was a sensible improvement in the results obtained with the simplified model.

V. Final Remarks

In this paper, we have presented a new quantitative approach for assessing corporate credit risk based on artificial neural networks. This approach is based on a credit rating model, defined using some ratios derived from the financial statements of each company, which are used as input to a multi-layer perceptron network. The model was validated through the implementation of an automated system using a back-propagation algorithm, which allowed us to train our network and benchmark its solutions. In relation to other linear regression and neural models, the literature has already shown the better performance of this kind of network architecture ([7], [5], [12]). Our experimental results suggest that this approach is relatively accurate in addressing the credit rating problem.

Our experiments can also be used as a basis for validating (or not) some conclusions reported in the literature, although obtained in different contexts. For instance, we have not confirmed that the use of historical data in the form of absolute difference values is more effective than using their growth [4]. On the other hand, we have confirmed the difficulty in discerning between adjacent rating categories [13], which seems to have motivated some supervisory authorities (such as the Brazilian Central Bank) to adopt a rating scale with less grades (only 9 in the case of Brazil). The way we organized our own model and the whole user-friendly credit rating solution, including our data selection criteria for training and benchmarking the developed system, make us believe that they constitute an original contribution to the field.

It is important to recall that the task of estimating the credit rating of a company is slightly different from assessing the possibility of its future default. The former task is more fine grain, because a company may not obtain a high credit rate (classification problem) even without apparent possibility of entering into a default state (recognition problem). Credit rating is also different from assessing the risk of buying or selling publicly tradable stocks or bonds, since the market behavior is also relevant in this case. Although these are all related problems, we do not know if the model proposed here is at least partially applicable for solving these other problems.

Our work suggests some research issues that deserve further investigation. For instance, it would be important to evaluate which additional input variables representing qualitative and other quantitative data are relevant for credit rating estimation, even by systematically validating our work against other credit rating models, as suggested in [14]. Concerning improvements in our own developments, it would be useful for the analyst to obtain a credit rating from an automated system even when some financial statements do not represent the end of year situation. In this case, the system should extrapolate the available financial data to obtain a projected end year statement. In analyses of start-up companies, it would also be interesting to rely only on projected financial data, obtained from cash flow statements, for instance. These investigations would all contribute to the development of a complete credit rating solution.

References

[1] Brazilian Central Bank, "Concerning rating criteria for credit operations and rules for the constitution of provisions (in portuguese)," Resolution 2.682, 1999.

[2] Maryse Farhi and Marcos Antônio Macedo Cintra, "Investor's information: Risk classification, accounting and conflicts of interest (in portuguese)," *Ensaios FEE*, vol. 23, no. 2, pp. 761–786, 2002.

[3] José Carlos Marion, *Corporate Accounting (in Portuguese)*, Atlas, 1998.

[4] Fernando C. Almeida, "A structured approach to neural networks in bankrupcy prediction," in *Proc. 4th Brazilian Symposium on Neural Networks (SBRN'97)*, Díbio L. Borges and Weber Martins, Eds. 1997, pp. 1–8, IEEE Computer Society.

[5] Ravipim Chaveesuk, Chat Srivaree-ratana, and Alice E. Smith, "Alternative neural network approaches to corporate bond rating," *Journal of Engineering Valuation and Cost Analysis*, vol. 2, no. 2, pp. 117–131, 1997.

[6] Antonio Morais da Silveira, Mara Lúcia Cerqueira da Silva, and Carlos Tavares da Costa, "Analysis of risk in the contracting of financial loans: An approach with diffuse logic," in *Proc. 4th Congress of Logic Applied to Technology (LAPTEC'03)*, Germano L. Torres, Jair M. Abe, Marcos L. M., and Paulo E. Crunivel, Eds. 2003, vol. II, pp. 1–8, Ed. Pleiade.

[7] Joachim Utans and John Moody, "Selecting neural network architectures via the prediction risk: Application to corporate bond rating prediction," in *Proc. 1st Int. Conf. on Artificial Intelligence Applications on Wall Street*. 1991, IEEE Computer Society Press.

[8] George Cybenko, "Continuous valued neural networks with two hidden layers are sufficient," Tech. Rep., Department of Computer Science, Tuffs University, 1988.

[9] Antônio P. Braga, André C. P. L. F. Carvalho, and Teresa B. Ludermir, *Artificial Neural Networks: Theory and Applications (in Portuguese)*, LTC, 2000.

[10] Object Management Group OMG, *Unified Modelling Language Specification*, Object Management Group — OMG, June 1999, Version 1.3.

[11] James Gosling, Bill Joy, and Guy Steele, *The Java Language Specification*, Addison-Wesley, 1996.

[12] Amir Atiya, "Bankruptcy prediction for credit risk using neural networks: A survey and new results," *IEEE Transactions on Neural Networks*, vol. 12, no. 4, pp. 929–935, July 2001.

[13] C. R. Krishnaswamy, Erika W. Gilbert, and Mary M. Pashley, "Neural network applications in finance: A practical introduction," *Financial Practice and Education*, vol. 34, no. 3, pp. 75–84, 2000.

[14] Jorge Sobehart, Sean Keenan, and Roger Stein, "Benchmarking quatitative default risk models: A validation methodology," *ALGO Reseach Quarterly*, vol. 4, no. 1-2, pp. 55–69, 2001.