ELSEVIER

Expert Systems with Applications 29 (2005) 75–81

Expert Systems with Applications

www.elsevier.com/locate/eswa

Predicting bond ratings using publicly available information

Kee S. Kim*

Finance Department, Southwest Missouri State University, 901 S. National Ave, Springfield, MO 65804, USA

Abstract

This paper developed a model to predict bond ratings using financial and non-financial variables although the rating agencies believe that bond ratings could not be replicated quantitatively by using a computer model. The model used an artificial intelligence (AI) technique that is non-parametric and designed to capture a dynamic relationship between input and output variables. The results showed that bond rating could be assessed quite accurately and critical variables were successfully identified. In addition the investment grade bonds were successfully distinguished from the speculative bonds.

1. Introduction

All of the publicly traded debt issues in the United States are rated by one or more of the three ratings agencies (Moody's, Standard & Poor's, and Fitch) and these ratings are intended to measure the long-term default risk of the bond. Bond rating agencies claim that their ratings reflect each agency's opinion about an issue's potential default risk and rely heavily on a committee's analysis of the issuer's ability and willingness to repay its debt and therefore researchers would not be able to replicate their ratings quantitatively

On the other hand, many researchers believe that the companies in each individual category of the ratings should possess some commonality in their patterns of financial data on which their opinions are based although they can vary depending on the size of the firm, accounting practices, and the industry as indicated by Gibson and Frishkoff (1986). Several studies used statistical methods, including regression, multi-variate discriminant analysis, probit and logit models to predict bond rating. (Ang & Patel, 1975; Altman & Katz, 1976; Bhandari, Soldosky, & Boe, 1979; Martin, Henderson, Perry, & Cronan, 1984). McAdams (1980) employs the use of multiple discriminant analysis to design a statistical credit analysis model to assist portfolio managers to predict agency downgrades of electric utility bonds. Horrigan (1966) and Pogue and Soldofsky (1966) use multiple regression model to predict Moody's findings. Pinches and Mingo (1973) use factor analysis to screen variables for predicting bond ratings and then apply multiple discriminant analysis. Kamstra, Kennedy, and Suan (2001) improve the statistical predictive model by combining several forecasting methods to predict bond ratings in the transportation and industrial sectors. They use ordered logit method to combine forecasts and they find that combined forecasts outperform their input forecasts. Recently neural networks have been applied to bond rating (Dutta & Shekhar, 1988; Surkan & Singleton, 1990; Kim, 1992; Kwon, Han, & Lee, 1997). Although neural networks approaches have several advantages over statistical methods (Salchenberger, Cinar, & Las, 1992; Tam & Kiang, 1992), the results of these studies were less than expected because the real data in application is usually unevenly distributed among classes and these approaches are limited in dealing with the ordinal nature of bond rating.

This paper intends to build a model that predicts bond ratings using publicly available data, using an artificial intelligence technique, adaptive learning networks (ALN). Our main concern is to examine whether an ALN technique can be used to overcome the difficulties associated with assessing the credit standing (ratings) of the firms using publicly available data. This paper also intends to identify the critical variables that determine the ratings. An ALN model is known to be better than other statistical models because the ALN model is able to learn and generalize the knowledge obtained from the correlations between the input

* Tel.: +1 417 836 6345; fax: +1 417 836 6224.
E-mail address: ksk882f@smsu.edu.

0957-4174/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2005.01.007

and output, which are frequently nonlinear, incomplete and/or unclear. When the data are discontinuous or nonlinearly separable, the innate representation become inconsistent, and therefore the mapping cannot be learned by multi-variate statistical techniques. Since the ALN model is non-parametric by nature and it can also ignore undesirable noise in the input data, it should be able to identify more accurately the credit rating of the firm whose data are frequently inconsistent, non-parametric, and have multicollinearity.

In the following the ALN is briefly discussed in Section 2. Next, the research design and data collection were discussed in Section 3. The study results and their analyses were reported in Section 4. A brief concluding remarks were shown in Section 5.

2. The adaptive learning network (ALN)

The adaptive learning networks (ALN) technique (Barron, Mucciardi, Cook, Craig, & Barron, 1984) was developed from almost three decades of statistical modeling, neural network, and artificial intelligence research. The ALN technique automatically generates the trained network from the database by fitting model coefficients to bases of observational data. It uses a network structure that resembles neurons and synapses of a human brain and also uses mathematical functions that represent numeric knowledge on each processing unit. The power of the network lies in its ability to decompose complex problems into much smaller and simpler ones and to solve them. The network structure makes decision-making much easier because the numbers of factors to consider and the alternatives to evaluate become smaller.

Montgomery (1989) developed an effective computer-based algorithm called the abductory induction mechanism (AIM), from which the final solution is synthesized in the form of a network. It uses a refinement of the group method of data handling (GMDH) algorithm of the Ivakhnenko and Ivakhnenko (1974) and polynomial network training routine (PNETTER) algorithm to create the polynomial network. The final model is a layered network of feed-forward functional elements in which the coefficients, number and types of network elements, and the connectivity are learned inductively and automatically.

Each processing unit (node) has a unique equation of multi-variable configurations: singles, doubles, triples, normalizers, white elements, unitizers, and wire elements (Montgomery, 1989). Normalizers transform the original input variables into standardized normal variables with a mean of zero and a variance of one. The white element is a linear combination of all inputs to the current layer. Unitizers convert the normalized data back into the original data to assess the output values. The algebraic form of singles, doubles, and triples is shown in the following equations:

$$\text{Single} = W_0 + W_1X_1 + W_2X_1^2 + W_3X_1^3$$

$$\text{Double} = W_0 + W_1X_1 + W_2X_2 + W_3X_1^2 + W_4X_2^2 + W_5X_1X_2 + W_6X_1^3 + W_7X_2^3$$

$$\text{Triple} = W_0 + W_1X_1 + W_2X_2 + W_3X_3 + W_4X_1^2 + W_5X_2^2 + W_6X_3^2 + W_7X_1X_2 + W_8X_1X_3 + W_9X_2X_3 + W_{10}X_1X_2X_3 + W_{11}X_1^3 + W_{12}X_2^3 + W_{13}X_3^3$$

where X_i and W_i denote input variables and coefficients, respectively. These elements are homogeneous multinomials of degree 3 in one, two, three variables and allow interaction among input variables. It is well known that a suitably high degree multinomial—a polynomial of n variables in which all cross products appear and combinations of the variables to a different degree are included—can approximate arbitrary functions of many variables very accurately (Barron et al., 1984). All these terms in the equation may not always appear in a node since these processes will throw out the terms that do not contribute significantly to output. The output of elements in one layer will then feed into subsequent layers, together with the original input variables. Networks are synthesized from layer to layer until the adaptive network model ceases to improve based on predicted squared error (PSE) criterion.

The objective of the ALN algorithm is to train and identify the model that minimizes the predicted squared error (PSE) that is a sum of the trained squared error (TSE) and the errors on as yet unforeseen data, without over fitting the data (A.R. Barron, 1984).

$$\text{PSE} = \text{TSE} + 2\sigma_\mathrm{p}^2\frac{K}{N}$$

where TSE is the average squared error of the model on the training sample observations, K is the number of coefficients that are estimated to minimize TSE, σ_p^2 is the prior estimate of true error variance, and N is the size of the training sample observation. The TSE alone would likely provide a poor estimate of future performance because TSE will create an overly complicated model structure that could be detrimental to the future error on the application data set.

A.R. Barron (1984) presented a statistical analysis of PSE that includes the expected squared error on future data set, and explained why PSE is a good estimate of future performances. The minimum PSE is always attainable because as each coefficient is added to the model, TSE decreases at a decreasing rate while the overfit penalty increases linearly. If the adaptive model is obtained by minimizing TSE alone, the model will perform well on the training data set, but it can perform poorly on evaluation

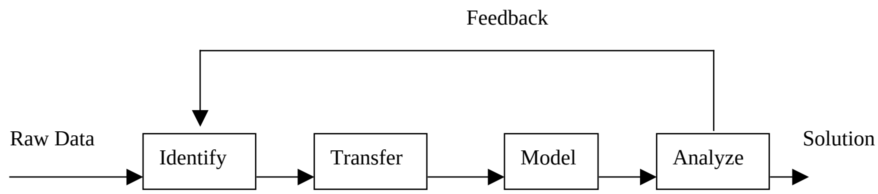

Fig. 1. ModelQuest™ modeling procedure.

samples. When the model has an overly complex structure with many coefficients, it will give a poor estimate of error on the test data set. By adding a term for the over-fit penalty, the minimum expected squared difference between the estimated model and the true model on the future data set can be obtained (Barron, 1984).

ModelQuest™ (2000) integrates advanced data modeling algorithms such as StarNet™ with more traditional data analysis technologies in a very easy-to-use and powerful data mining technology. An application of ModelQuest™ entails the following four steps as shown in Fig. 1. The first step involves identifying and characterizing data for a better solution. The second step involves representing and transforming the data through mapping, sampling, and feature extraction routines to provide additional inputs and to compensate for outliers and sparse regions. The third step is to split data into training and evaluation subsets, and train a model from a training data set. The ModelQuest™ automatically synthesizes the optimal ALN network including model size, connectivity, and parameter values. The last step involves applying the model to an evaluation data set and predicting output values. The accuracy and consistency of prediction determines how well the model works.

3. Model development and data

Asset size, subordination status, financial leverage and profitability ratios are directly related to the default risk of the firm and also included in most of the previous bond rating studies as explanatory variables (Ederington, 1985; Kaplan & Urwitz, 1979). Asset management, liquidity and cash flow ratios are also directly associated with the default risk. Professional rating agencies listed qualitative factors such as management ability, value of intangible assets, financial flexibility, operating efficiency, industry risk, accounting quality and market position. However, most of these qualitative factors are likely reflected in the quantifiable data such as financial and other non-financial variables, and could be assessed indirectly from analyzing these quantifiable data..

This model used 27 financial and non-financial variables as explanatory variables as shown in Appendix A. The input variables include both stock- and flow-variables that represent cash reserves, leverage, liquidity, profitability, and market value ratios. In addition, expected values and standard deviation of key earnings and financial leverage variables are also included initially as input variables. The output variable is *S* and *P* bond rating which represents 23 different bond ratings, including AAA$^+$, AAA, AAA$^-$, AA$^+$, AA, AA$^-$,…,CC$^+$. In this research, investment grade bonds whose rating is higher than BB+ were categorized into two classes and all speculative bonds also into two classes, since the sample is unevenly distributed among classes (see Appendix B). Security analysts and investors can use these ratings as the primary source of obtaining information about the quality and marketability of various bond issues and assess also the market risk premium attached to securities while investment bankers use the bond ratings for determining commission rates on underwritings. The reclassification of the ratings in four classes could enhance model performances because mathematical and statistical approaches have generally limits in dealing with the ordinal nature of the bond rating. It was known that, as the number of bond classification increases, the predictive power could likely decrease (Kwon et al, 1997).

A sample of 1080 observations (companies) in 2001 was collected primarily from the COMPUTSTAT database, Dun and Bradstreet database, and *S* and *P* bond manuals. The data set excluded utilities, transportation, and financial companies because their financial structures are quite different from the rest of companies. Next, the sample was divided randomly into one (75%) for training and another (25%) for testing the model. A larger sample size is used for training because an accurate training normally requires a larger sample by nature.

4. Empirical results and implications

A final adaptive learning network (ALN) is synthesized in Fig. 2 from training, using the ModelQuest™. It is a layered network of feed-forward functional elements, which contain the best network structure, node types, coefficients, and connectivity to minimize the predicted squared error (PSE). The model used nine different input variables with a repeat of ROA and LNTA to synthesize the final ALN model. None of the statistical variables such as standard deviation and expected variables was selected for predicting bond ratings.

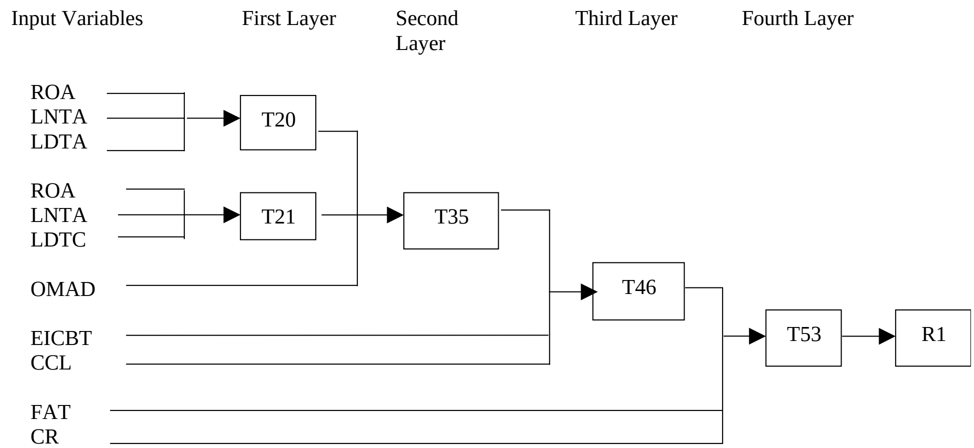

Fig. 2. Adaptive learning network model.

The equations in Appendix C show the final ALN model in polynomial equation forms, and each equation number represents the node number of the ALN network as shown in Fig. 2. The nine input variables were first transformed into standardized normal variables with the mean of zero and a variance of one using normalizers in the Appendix C. These standardized variables were next fed into the first layer to generate a series of intermediate output values. For example, the node triplet 20 (T20) was synthesized using normalized values of return on total assets (ROA), natural log of total assets (LNTA), long-term debt to total assets (LDTA) ratio, and then fed into the node T35 with two other inputs: T21 and operating margin after depreciation (OMAD). The value of the node T35 with two other variables, EICBT and CCL are fed again into the subsequent node T46 in the third layer, and finally T53 in the fourth layer is synthesized with two other variables, fixed asset turnover (FAT) and current ratio (CR). These values are again converted back to bond ratings with the mean and variance of the original output variables. When this model is validated using a holdout sample, this final ALN model becomes a knowledge base from which bond ratings could be assessed, using the nine different input variables.

The prediction results of the ALN on the evaluation sample are shown in Fig. 3. The overall performance of the model shows that the trained ALN model was successful in predicting two hundred twenty eight observations correctly (84%) out of 272 cases. The ALN model categorized eight cases (57%) correctly out of 14 companies in the best rating group while six cases (43%) was categorized correctly in the second group. The prediction accuracy on the level 2 and 3 are 88 and 91%, respectively, and shows convincingly that both the investment grade and speculative bonds were correctly identified using financial ratios alone without referring to the opinion of the two rating companies. The prediction result of companies in the lowest rating group seems to indicate that the companies in the rating level 4 might not be distinguishable from those in the level 3 since 10 out of 12 were categorized into the third group while none of the companies identified correctly.

This result is much better than most of the studies as shown in Fig. 4. For example, Ederington (1985) reported that the ordered probit regression approach produced 78% of correct prediction, 73% by the unordered logit and 64% by the multinomial discriminant analysis approaches.

Actual \ Predicted	1	2	3	4	Total
1	8	6	0	0	14
2	1	113	14	0	128
3	1	7	107	3	118
4	1	1	10	0	12
Total	11	127	131	3	272

Fig. 3. Contingency table for the prediction results.

Articles	Approaches	Prediction Accuracy
Pinches & Mingo (1973)	Quadratic Discriminant	64.6%
Altman & Katz (1976)	Quadratic Discriminant	77%
Kaplan & Urwitz (1979)	Ordered Probit analysis OLS	50% 55%
Ederington (1985)	Linear Regression Ordered Probit Unordered Logit Quadratic Discriminant	65% 78% 73% 64%
Dutta & Shekhar (1988)	Multiple Regression Neural Networks	63.2% 82.4%
Surkan & Singleton (1990)	Neural Networks	58%
Utans & Moody (1997)	Multi-layer Perceptron Network	31%

Fig. 4. Summary of previous studies on bond rating.

Input Description	Sensitivity	Importance
LNTA	0.8030	0.8197
CR	0.0782	0.0125
ROA	0.0606	0.0469
LDTA	0.0297	0.0532
FAT	0.0105	0.0101
OMAD	0.0096	0.0187
ICBT	0.0037	0.0109
LDTC	0.0024	0.0098
CCL	0.0023	0.0182

Fig. 5. Variable sensitivity and importance.

Altman and Katz (1976) used the quadratic discriminant analysis to securities in the public utility industry, and predicted correctly 77% of a testing sample. Their high prediction rate is believed to be attributable to their concentrations on companies in one industry: an electric public utility industry. Kaplan and Urwitz's (1979) correct prediction rate was only on 50%. Dutta and Shekar (1988) used a multiple regression model and neural network model to analyze 47 companies. They used 30 companies for training and 17 companies for testing. Their neural network model predicted 82.4% correctly while only 63.2% by the regression model. Their model performance by the neural network approach was better than most of studies in the literature, probably because their objective was to predict whether a bond's rating belongs to AA class or not. Some researchers (Dutta & Shekar, 1988; Pogue & Soldofsky, 1969) used a pair-wise classification while others used five classifications (Pinches & Mingo, 1973), but their results were not better than the ones in this study.

The final model selected a variety of nine variables including size (LNTA), profitability (ROA, OMAD, ICBT), financial leverage (LDTA, LDTC), cash reserves (CCL), asset management (FAT), and liquidity (CR) variables (Fig. 5). Among them the asset size is dominant in the sense that over 80% of bond rating changes are determined by total asset alone according to the sensitivity value that indicates the power of individual input variable on the model output value at a point of average input values. It amounts to the fact that over 97% of the rating changes were explained by the four major variables including LNTA, CR, ROA, and LDTA. The results provide important strategic implications for bond rating changes. One effective way of enhancing bond rating is to increase its asset size but it should be done without adversely affecting the company's ROA. The increase in assets must therefore be accompanied by an adequate amount of increase in net income, while the company maintains its liquidity and financial leverage at around the current level.

The Importance value in Fig. 5 indicates the expected overall contribution of each input variable to the predicted bond rating changes, which is standardized by the total output changes possible. For example, Fig. 5 indicates that about 82% of bond rating changes are caused by changes in total assets while only 5.3 and 4.7% of rating changes are attributable to changes in the size of the debt ratio (LDTA) and ROA, respectively. The important implication of the above results is that bond rating changes are primarily determined by the company's total asset size while the financial leverage and profitability of the company are also important in bond rating determination. There is not much difference between sensitivity and importance values except CR is less influential in importance level.

5. Concluding remarks

This study has shown that bond ratings could be assessed using publicly available data without referring to the expertise of the professional rating agencies. Furthermore, it has identified critical variables that determine bond ratings. Among them the total asset size of the company was the primary determinant. The variation of bond rating was also attributable to the current ratio and ROA among other variables although their influence levels are quite limited. When companies increase their capital investments on the profitable projects, their ROA and liquidity will automatically improve and the company will be able to build its value from operating activities.

As an extension of the study, it would be worthwhile to examine the prediction outcome of bond rating if this model is trained and tested within the confines of an industry or similar industries. Since individual industries have their own common characteristics in financial structure, it can be conjectured that the model performance could be enhanced significantly.

Appendix A. Input data to the ALN model

A.1. Size:

X_1: total assets (dollars) in natural log (TA)

A.2. Profitability:

X_{13}: return on total assets (ROA)
X_{14}: return on equities (ROE)
X_{15}: profit margin (PM)
X_{16}: operating margin after depreciation (OMAD)
X_8: earning growth (EG)

A.3. Leverage:

X_{10}: LT debts/total assets (LD/TA)
X_{17}: LT debts/total capitalization (LD/TC)
X_{11}: loan/total assets (L/TA)
X_7: interest coverage before tax (ICBT)
X_9: average TIE (AICBT)
X_{24}: total debts/cash equivalents (TD/CASH)
X_{21}: retained earnings/total assets (RE/TA)

A.4. Asset management:

X_{20}: sales/accounts receivables (RTO)
X_6: sales/total cash reserves (CaTO)
X_{26}: sales/fixed assets (F ATO)
X_{25}: sales/total assets (TATO)

A.5. Liquidity:

X_2: current ratio (CR)
X_5: net working capital/sales (NWC/S)
X_4: net working capital/total assets (NWC/TA)

A.6. Cash:

X_3: cash reserves/current liabilities (Ca/CL)
X_{22}: cash inflows/inventories (Cin/IV)
X_{23}: cash inflows/total debts (Cin!TD)

A.7. Value:

X_{19}: pricelearnings ratio (PE)
X_{18}: market value/long-term debts (MV/LTD)
X_{12}: market value/book value (ME)

A.8. Others:

X_{27}: subordination

Appendix B. Data descriptions of bond ratings

	Frequency of bond data			
	Original ratings	Model ratings	Frequency	Cum
AAA+	2	1	11	11
AA+	4	1	3	14
AA	5	1	13	27
AA−	6	1	22	49
A+	7	2	59	108
A	8	2	85	193
A−	9	2	76	269
BBB+	10	2	108	377
BBB	11	2	121	498
BBB−	12	2	101	599
BB+	13	3	59	658
BB	14	3	94	752
BB−	15	3	88	840
B+	16	3	103	943
B	17	3	60	1003
B−	18	3	41	1044
CCC+	19	4	19	1063
CCC	20	4	11	1074
CCC−	21	4	7	1081
CC+	23	4	10	1091
			1091	

Appendix C. Network equations for the final ALN model

C.1. Adaptive netowrk equations

C.1.1. Normalizers:

$$\mathrm{ROA} = -0.27 + 0.1116X_1$$

$$\mathrm{LNTA} = -15.2375 + 0.7007X_1$$

$$\mathrm{LDTA} = -1.4079 + 0.0419X_1$$

$$\mathrm{LDTC} = -0.9538 + 0.0187X_1$$

$$\mathrm{E\_ICBT} = -0.2534 + 0.0349X_1$$

$$\mathrm{C\_CL} = -0.4963 + 0.0122X_1$$

$$\mathrm{FAT} = 0.5232 + 0.1091X_1$$

$$\mathrm{CHE\_TA} = -0.5097 + 0.0071X_1$$

$$\mathrm{LOAN\_TA} = -1.4709 + 0.1201X_1$$

C.1.2. Triples:

$$\begin{aligned} T20 = {} & 0.0797 + 0.2968X_1 + 0.0010X_1^3 + 0.4618X_2 \\ & + 0.0557X_1X_2 + 0.0294X_1^2X_2 - 0.0881X_2^2 \\ & - 0.0712X_1X_2^2 - 0.4273X_3 - 0.0503X_2X_3 \\ & + 0.0701X_1X_2X_3 + 0.0893X_2^2X_3 - 0.0631X_3^2 \\ & - 0.0143X_1X_3^2 + 0.0349X_2X_3^2 + 0.0162X_3^3 \end{aligned}$$

$$\begin{aligned} T21 = {} & -0.1022 + 0.1759X_2 + 0.484X_1X_2 + 0.0906X_2^2 \\ & + 0.2081X_1X_2^2 + 0.0213X_2^3 + 0.4161X_3 \\ & + 0.2395X_1X_3 + 0.0385X_1^2X_3 + 0.8119X_2X_3 \\ & + 1.0632X_1X_2X_3 + 0.173X_2^2X_3 - 0.0687X_3^2 \\ & - 0.0534X_3^2 - 0.1411X_2X_3^2 + 2.0E-4X_3^3 \end{aligned}$$

$$\begin{aligned} T35 = {} & 0.0511 + 1.2712X_1 - 0.1471X_1^2 - 0.3168X_1^3 \\ & + 0.4655X_2 - 0.0831X_1^2X_2 - 0.1058X_2^2 \\ & + 0.0053X_2^3 - 0.2395X_3 + 0.0614X_1X_3 \\ & + 0.0786X_1^2X_3 - 0.4029X_2X_3 + 0.239X_1X_2X_3 \\ & + 0.0413X_2^2X_3 - 0.0083X_1X_3^2 + 0.0521X_2X_3^2 \\ & + 0.0020X_3^3 \end{aligned}$$

$$T46 = 1.2755X_1 - 0.0481X_1^2 - 0.296X_1^3 - 0.1525X_2 - 0.0106X_1X_2^2 + 0.0020X_2^3 + 0.1977X_3 - 0.0882X_1^2X_3 - 0.2276X_2X_3 - 0.0772X_1X_2X_3 + 0.0097X_2^2X_3 - 0.0311X_3^2 - 0.0273X_2X_3^2$$

$$T53 = 1.305X_1 - 0.3048X_1^3 - 0.1416X_2 + 0.208X_1^2X_2 - 0.1483X_1X_2^2 + 0.0464X_2^3 + 0.1391X_3 - 0.0398X_1X_3 - 0.0806X_1^2X_3 + 0.0365X_2X_3 - 0.0344X_2^2X_3 - 0.0262X_3^2 + 0.0026X_3^3$$

C.1.3. Unitizers:

$$R3 = 0.5575 + 0.497X_1$$

References

Altman, E., & Katz, S. (1976). Statistical bond rating classification using financial and accounting data. In M. Schiff, & G. Sorter (Eds.), *Topical research in accounting*. NYU Press.

Ang, J., & Patel, S. K. A. (1975). Bond rating methods: comparison and validation. *Journal of Finance*, *v30*(2), 631–640.

Barron, A. R. (1984). Predicted squared error: a criterion for automatic model selection. In S. J. Farlow (Ed.), *Self-organizing methods in modeling: GMDH type algorithms*. New York: Marcel-Dekker, Inc..

Barron, R. L., Mucciardi, A. N., Cook, F. J., Craig, J. N., & Barron, A. R. (1984). Adaptive learning networks: development and application in the u.s. of algorithms related to GMDH. In S. J. Farlow (Ed.), *Self-organizing methods in modeling: gmdh type algorithms*. New York: Marcel-Dekker, Inc..

Bhandari, S. B., Soldofsky, R. M., & Boe, W. J. (1979). Bond quality rating changes for electric utilities: a multivariate analysis. *Financial Management*, *v8*(1), 74–81.

Dutta, S., & Shekhar, S. (1988). *Bond rating: a non-conservative application of neural networks Proceedings of the IEEE international conference on neural networks* pp. II442–450.

Ederington, L. H. (1985). Classification models and bond ratings. *Financial Review*, *20*, 237–261.

Gibson, C., & Frishkoff, P. (1986). *Financial statement analysis using financial accounting information* (5th ed). Boston: Kent Publishing company.

Horrigan, J. L. (1966). The determination of long term credit standing with financial ratios, empirical research in accounting: selected studies, Supplement to V.4. *Journal of Accounting Research*.

Ivakhnenko, A. G., & Ivakhnenko, N. A. (1974). Long-term prediction by GMDH algorithms using the unbiased criterion and the balance-of-variables criterion. *Soviet Automation Control*, *7*(4), 40–45.

Kamstra, M., Kennedy, P., & Suan, T.-K. (2001). Combining bond rating forecasts using logit. *The Financial Review*, *v37*, 75–96.

Kaplan, R. S., & Urwitz, G. (1979). Statistical models of bond ratings: a methodological inquiry. *Journal of Business*, *52*, 261–321.

Kim, J. (1992). *A comparative study of rule-based, neural networks, and statistical classification systems for the bond rating problem, unpublished doctoral dissertation*. Richmond, VA: Virginia Common Wealth University.

Kwon, Y. S., Han, I., & Lee, C. (1997). Ordinal pairwise partitioning (OPP) approach to neural networks training in bond rating. *International Journal of Intelligent Systems in Accounting, Finance and Management*, *v6*(1), 23–40.

Martin, L., Henderson, G., Perry, L., & Cronan (1984). Bond ratings: Predictions using rating agency criteria. Working paper 85–3. Arizona State University, Tempe, AZ.

McAdams, L. (1980). How to anticipate utility bond rating changes. *Journal of Portfolio Management*, *v7*(1), 56–60.

Montgomery, G. J. (1989). *Abductive diagnostics Proceedings of the AIAA computers in aerospace conference, Monterey, CA*.

Pinches, G. E., & Mingo, K. A. (1973). A multivariate analysis of industrial bond ratings. *The Journal of Finance*, *v28*, 1–18.

Pogue, T. F., & Soldofsky, R. M. (1969). What's a bond rating? *Journal of financial and Quantitative Analysis*, *v4*, 201–228.

Salchenberger, L. M., Cinar, E. M., & Las, N. A. (1992). Neural networks: a new tool for prdicting thrift failures. *Decision Sciences*, *23*(4), 899–915.

Surkan, A., & Singleton, Jl. (1990). *Neural networks for bond rating improved by multiple hidden layers Proceedings of the IEEE international conference on neural networks* pp. II157–162.

Tam, K., & Kiang, M. (1992). Managerial applications of neural networks: the case of bank failure predictions. *Management Science*, *38*(7), 926–947.

Utans, J., & Moody, J. (1997). Selecting network architectures via prediction risk: application to corporate bond rating prediction. *IEEE*, *v6*, pp35–pp41.