The Classification of Companies by Means of Neural Networks

Jörg Baetge, Clemens Krause

The Classification of Companies by Means of Neural Networks

by Jörg Baetge[1] and Clemens Krause[2]

1. Introduction

The readiness of banks to extend credits depends on the client's credit standing, i. e. the client's willingness and capability to discharge the credit as well as its rates of interest and redemption. Sometimes in cases of credit evaluation, formalized methods aiming at the objectification and rationalisation of that operation are made use of. More often than not, statistical methods[3] serve as formalized methods, but methods of pattern recognition[4] are also employed. So far, the statistical method of the multivariate linear discriminant analysis has frequently and successfully been used for the purpose of credit evaluation in Germany, e. g. by the Bayerische Vereinsbank AG[5], the Bundesbank[6], the Deutsche Bank AG[7] and the Allgemeine Kreditversicherung AG.[8]

---

1 Prof. Dr. Jörg Baetge, Director of the Institut für Revisionswesen, Westfälische Wilhelms-Universität Münster, Universitätsstr. 14-16, D-48143 Münster, Germany, Phone: (0251) 83-2961.

2 Dr. Clemens Krause, member of staff at the Institut für Revisionswesen, Westfälische Wilhelms-Universität Münster, Universitätsstr. 14-16, D-48143 Münster, Germany, Phone: (0251) 83-2961.

3 For the stastistical methods used for the insolvency diagnosis of companies see the survey in RÖSLER, J., Die Entwicklung der statistischen Insolvenzdiagnose, pp. 105-112.

4 See HENO, R., Kreditwürdigkeitsprüfung mit Hilfe von Verfahren der Mustererkennung, pp. 174-252.

5 See NIEHAUS, H.-J., Früherkennung von Unternehmenskrisen, p. 5.

6 See LAMPE, W., Aussagen quantitativer Kreditnehmeranalysen, p. 204.

7 See BREUER, R.-E., Bilanzanalyse aus der Sicht der Kreditinstitute, p. 154.

8 See FEIDICKER, M., Kreditwürdigkeitsprüfung - Entwicklung eines Bonitätsindikators, p. 4.

The methods of pattern recognition also include Artificial Neural Networks (Neural Networks) which, however, have hardly ever been used in cases of credit evaluation so far. In the following, the question shall be answered if and to what extent neural networks are suitable for credit evaluation. On the basis of a large amount of empirical data, the classification results of **backpropagation networks** (BPN) shall be ascertained and compared with the classification results of the multivariate linear discriminant analysis (MDA).[9]

2. The Data Material and the "Optimization" of the BPN-Parameters

6667 data records, each of which represents the annual financial statements of a company, formed the basis of the inquiry. The companies analysed by means of the data records are neither dependent on the government nor on a group, and their gross performances exceed DM 500,000 in each case.[10] Banks and insurance companies were excluded from the inquiry on account of their different balance sheet structure. The annual financial statements originate from a period of inquiry extending from December 31, 1973 to August 31, 1987, which means that the last financial year included is 1986. The annual financial statemants analysed were prepared in accordance with the regulations of the German Stock Corporation Law of 6 Sept. 1965. From the annual financial statements, 73 ratios were formed by means of which differences between **solvent** and **insolvent** companies could be analysed. Those companies which, e. g., were found to be in states of bankruptcy or debt composition proceedings will be labelled as insolvent in the following. The remaining companies were regarded as solvent.

For neural networks, the data material must be **divided into three groups of data records**.[11] We have decided in favour of the following allocation:[12]

---

9 The discriminant analysis was performed on behalf of the Allgemeine Kreditversicherung AG in Mainz/Germany. The results of this project have been described at length in: FEIDICKER, M., Kreditwürdigkeitsprüfung - Entwicklung eines Bonitätsindikators, and BAETGE, JÖRG/ BEUTER, HUBERT B./ FEIDICKER, MARKUS, Kreditwürdigkeitsprüfung mit Diskriminanzanalyse, pp. 749-761.

10 For the prerequisites also see NIEHAUS, H.-J., Früherkennung von Unternehmenskrisen, p. 59 f.; BAETGE, J., Möglichkeiten der Früherkennung, p. 799.

11 The division of the data material into three parts corresponds to the suggestion of HECHT-NIELSEN (see HECHT-NIELSEN, R., Neurocomputing, p. 115 f.).

Analysis sample		Test sample		Validation sample	
Solvent	Insolvent	Solvent	Insolvent	Solvent	Insolvent
336	336	2686	343	2654	312
Total=672		Total=3029		Total=2966	

Table 1: Allocation of 6667 records

With the **analysis sample**, the neural network is trained and FEIDICKER'S[13] discriminant function is calculated. The analysis sample includes 672 data records with the information from the annual financial statements of 112 insolvent and 112 solvent companies originating from three consecutive years in each case. The year in which the insolvency took place is later on marked as t-0 while, e. g., t-1 stands for the year directly prior to the state of insolvency. The data records of the analysis sample contain the annual financial statements of morbid companies from the periods t-1, t-2 and t-3. This temporal data structure also applies to the sound companies. With regard to the latter, the interval is not related to the date of insolvency but to the moment of data gathering: the most recent annual financial statements available are marked as t-1 while the least recent of the three annual financial statements are labelled as t-3. The classification results arrived at by means of neural networks from the analysis sample are confronted with the problem of "overtraining",[14] which has to be taken into account. Overtraining means that in an extreme case, the neural network might even learn the analysis sample completely and thereby lose the capacity to transfer the acquired knowledge to other data records not included in the analysis sample. Because of the problem of overtraining, the classification quality cannot be evaluated by means of the data records of the analysis sample but only by means of data of the validation sample and the test sample that were **not** used for network training.

The classification qualities of neural networks are influenced by **various parameters**. With the backpropagation network, seven network-specific and two

12 See KRAUSE, CLEMENS, Kreditwürdigkeitsprüfung mit Neuronalen Netzen, p. 128.

13 See FEIDICKER, M., Kreditwürdigkeitsprüfung - Entwicklung eines Bonitätsindikators, p. 161.

14 HECHT-NIELSEN calls this problem"overtraining" (see HECHT-NIELSEN, R., Neurocomputing, p. 115 f.). Other authors prefer the expression "overfitting" (see HERTZ, J./ KROGH, A./ PALMER, R. G., Introduction to the Theory of Neural Computation, pp. 145-147).

application-specific parameters are modified in this study.[15] As individual parameters may partially become continuants, it is, within a limited period of time, impossible to test all parameter constellations imaginable. By means of the **test sample**, the classification quality of various parameter adjustments of a neural network is tested. Since there are no analytical methods that lead to an optimum parameter adjustment, the parameters have to be "optimized" by means of a heuristic method. The opinion, however, that the adjustment of parameters should merely be based on the intuition of the network designer[16] is by no means convincing. The use of evolutionary or genetic algorithms provides a less subjective opportunity to optimize the parameters of a neural network.[17] However, the adjustment of such a search strategy to the problem at hand would require a lot of effort, for rules how to mutate and cross parameter adjustments must be developed. Accordingly, we did not decide in favour of evolutionary or genetic algorithms.

Instead of this, we proceeded analogous to a less demanding method of nonlinear optimization, i. e. the relatively simple **coordinate method.**[18] With the help of the coordinate method, multi-dimensional optimization problems are solved by making several single-dimensional, axis-parallel optimization steps. Proceeding from any starting point whatever, one variable (= parameter of the neural network) is modified by one step first. If the target figure has improved (the Beta-error[19] decreased), the step direction will be retained. Failing this, a step into the opposite direction will be tried. A successful step direction will be maintained unless a failure takes place. The point thus found becomes the starting point proceeding from which the next variable is changed. And yet, the coordinate method only arrives at the optimum safely with unimodal functions. Since the error function depends on the parameters of a neural network and thus unimodality cannot be taken for granted, an optimum solution is not guaranteed. Nevertheless, the coordinate method has proved its value in finding good solutions in the course of this inquiry within the test-sample.

In the process of searching for a suitable parameter adjustment, the quality of each of these parameter adjustments of the neural network has to be evaluated by means of a classification test with the **test sample**, and not with the **validation sample** because it may happen that a parameter adjustment is chosen best wich overpictures the special features of the sample. If the "optimization" is performed with the test sample this does not matter because the final classification results can be evaluated independently with the validation sample. With the

15 For the parameters of a backpropagation network, see chapter 4.

16 This is the opinion of KRATZER, K. P., Neuronale Netze, p. 167.

17 See BRAUSE, R., Neuronale Netze, p. 243.

18 For the coordinate method see ZANGWILL, W. I., Nonlinear Programming, p. 111 f.

19 For the definition of the Beta-error see chapter 3 and table 2 and figure 2.

parameter adjustment that proved best in the test sample, the classification quality in case of the **validation sample** is finally determined, for these results only are representative for data records of the parent population that are unknown to the network and must newly be classified.

FEIDICKER had only two samples, the analysis sample (672 records) and the validation sample (5995 records), because he did not need any test sample for MDA reasons. Therefore, we had to split up his validation sample into two groups of test sample (3029 records) and validation sample (2966 records). We had also to re-classify the validation sample. The result is seen in table 2.

3. The Evaluation of the Results of Classification

In the books and articles on the classification of companies by means of neural networks that have been published so far, the classification performances of different methods are often hard to compare.[20] In the course of classification the two error types Alpha-error (an insolvent company is classified as solvent) and Beta-error (a solvent company is classified as insolvent) are often differentiated. In this case the classification quality of two different neural networks cannot be compared if Alpha- as well as Beta-error differ from one another in the course of two classification tests with neural networks.[21] If, e. g., one neural network supplies an Alpha-error of 10 % and a Beta-error of 30 % while another network supplies Alpha- and Beta-errors of 20 % each, it does not become evident which of the two networks classifies in a better way. The differing Alpha- and Beta-errors come up automatically if a constant critical discriminant value is set in advance for the classification.[22] Hence, some authors arrive at a comparison of classification performances over an undifferen-

20 See ODOM, M. D./ SHARDA, R., A Neural Network Model for Bankruptcy Prediction, pp. 163-168.

21 TAM, K. Y./ KIANG, M., Predicting Bank Failures, pp. 265-282 and ODOM, M. D./ SHARDA, R., A Neural Network Model for Bankruptcy Prediction, pp. 163-168 differentiate Alpha- and Beta-error.

22 SCHUMANN/ LOHRBACH/ BÄHRS (see SCHUMANN, M./ LOHRBACH, T./ BÄHRS, P., Kreditwürdigkeitsprognose mit Künstlichen Neuronalen Netzen, p. 9 f.), who have checked the classification performances of different types of neural networks while evaluating the credit standing of consumers, and REHKUGLER/ PODDIG (REHKUGLER, H./ PODDIG, TH., Klassifikation von Jahresabschlüssen, p. 24), who classify companies on the basis of their annual financial statements, preset the critical discriminant value as an absolute term.

tiated hitting or error quota integrating Alpha- and Beta-error.[23] However, this method does not allow any comparison either, because the undifferentiated hitting or error quota also depends on the ratio of Alpha- to Beta-error.[24]

In our research work with an iterative search procedure, the critical discriminant value is exclusively determined in such a manner that, whenever a classification test is made, the Alpha-error remains at a constant level. We chose about 8.7 % as a constant level of the Alpha-error. This level of the Alpha-error is selected in such a way that the costs of error classifications are minimized. If, on the basis of empirically ascertained costing, there is an Alpha-error with DM 70,000 risk costs and a Beta-error with DM 100 reworking costs, FEIDICKER arrives at the conclusion that costs are minimized if the Beta-error is 5.26 times higher than the Alpha-error.[25] The classification quality of all our analyses is expressed through the Beta-error. The following table shows, for the test sample and the validation sample, the classification results of the discriminant function with the four annual financial statement ratios selected on the basis of the discriminant analysis.

---

23 Undifferentiated hitting or error quotas can be found in SCHUMANN, M./ - LOHRBACH, T./ BÄHRS, P., Kreditwürdigkeitsprognose mit Künstlichen Neuronalen Netzen, p. 35 and REHKUGLER, H./ PODDIG, TH., Klassifikation von Jahresabschlüssen, p. 7 f.

24 See KRAUSE, CLEMENS, Kreditwürdigkeitsprüfung mit Neuronalen Netzen, p. 118-124.

25 See FEIDICKER, M., Kreditwürdigkeitsprüfung - Entwicklung eines Bonitätsindikators, pp. 212-214.

	Test sample			Validation sample		
	Number	Classification as		Number	Classification as	
		solvent	insolvent		solvent	insolvent
Annual financial statements of companies actually insolvent	343	8,75 %	91,25 %	312	8,65 %	91,35 %
Annual financial statements of companies actually solvent	2686	52,79 %	47,21 %	2654	55,50 %	44,50 %

Table 2: Results of classification with FEIDICKER's discriminant function

FEIDICKER's discriminant function (MDA) contains the following four ratios, which we tested in one of our backpropagation networks (BPN-4):

Ratio	Definition
R_21	cash flow I: short-term outside capital
R_45	equity capital II: balance sheet total
R_47	(trade creditors + commitments + short-term due to banks): outside capital
R_55	total level of debt ("down payments for work in progress" included): sales revenues

Table 3: Ratios used by discriminant function and BPN-4

4. The Structure of the Backpropagation Network and the Way it Works

For reasons of efficiency, the classification tests start off with a very small backpropagation network. The backpropagation network made use of in this inquiry is structured as follows:

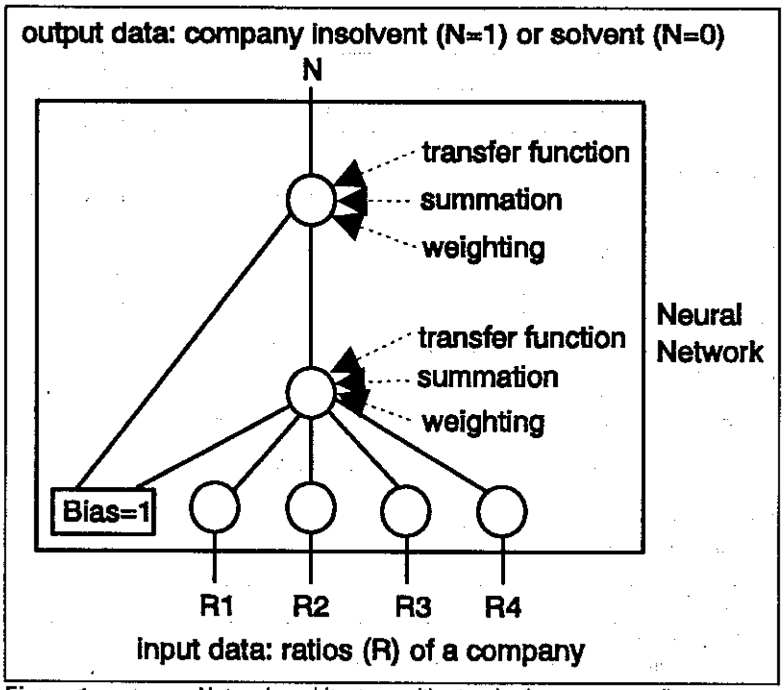

Figure 1: Network architecture with standard parameter adjustment

In the initial situation, i. e. with a standard adjustment of parameters, a three-layer backpropagation network is tested. At first, this network is trained with the four ratios calculated by means of the multivariate linear discriminant analysis. For each of these ratios, the network has got one neuron in the input layer. In the output layer, there is merely one neuron altogether, for one neuron proves sufficient for the classification into solvent (output = 0) and insolvent (output = 1) companies. The network has only got one hidden layer. Theoretically, one hidden layer suffices for the solution of a given problem.[26] This minimum equipment of the hidden layer with only one neuron restricts the capacity of learning and reduces the problem of overlearning, which we are dealing with in chapter 5 under 1.).[27] It will turn out whether this small learning capacity is sufficient if, in the course of additional tests, the network architecture is modified by using

26 See HECHT-NIELSEN, R., Neurocomputing, p. 131 f.

27 For the dimension of the hidden layer see KRATZER, K. P., Neuronale Netze, pp. 147-153.

more neurons in the hidden layer and by using more hidden layers. Following the recommendations of NEURALWARE, a so-called **bias** is generally used in addition to the conventional neurons. The bias functions as a cutoff. It is a special neuron steadily providing all neurons of the hidden layer and the output layer with the value 1.[28]

In the hidden layer as well as in the output layer the input signals of the neurons together with the bias are weighted (i. e. multiplied by the connection weights) and summed. Then, the transfer function is applied to the sum. The following transfer functions are possible: **sine, tangens hyperbolicus**, the **sigmoid** or the **linear function** (identity function). In case of standard parameter adjustment, a linear transfer function (the identity function) is used in the neurons of the input layer and the input data are passed on to the hidden layer without being modified. The following parameters have been changed while dealing with the structure of the backpropagation network described above:

Network-specific parameters:

- **Number of iterations**. The number of iterations shows how many data records are presented to the network during the learning process. The order in which the data records are presented to the network is determined by a random number generator. By means of this random sequence, cyclical weight changes can be avoided and the process of learning can be accelerated.
- **Rate of learning**. The learning rate influences the speed and the stride of learning. The standard adjustment of the **learning rate** μ of the generalized Delta-Rule amounts to 0.9.
- **Rule of learning**. Apart from the generalized Delta-rule, the cumulative generalized Delta-rule, which does not provide for a **learning step** with every iteration but only after a certain number of iterations, may also be used. The generalized Delta-rule as well as the cumulative generalized Delta-rule are always employed together with a so-called momentum. The continuity of learning is increased by the **momentum**, for the momentum leads to a rise in weight adjustments if several weight adjustments have the same algebraic sign and prevents the connection weights from going up and down with reciprocal weight adjustments.
- **Transformation of values.**[29] In many cases, a transformation of values is required before neural networks become able to supply satisfactory results. The transformation of values includes the processing of input data (input transformation), the processing of output data (output transformation) and the intra-network processing with the transfer function in the neurons. As a consequence of this transformation of values, the in- and output values are only presented to or put out by the network on a certain scale. With regard to the problem at hand, the output values need not be transformed, but it

---

28 See NEURALWARE INC. (eds.), Neural Computing, p. 106.

29 For the transformation of values see NEURALWARE INC. (eds.), Reference Guide, pp. 217-220.

will become obvious that, by means of a transformation of the input values, the classification results can be improved. The input transformation does not only lead to an increase or decrease in the interval the input data are scattering in, but also results in a uniform interval of dispersion for the various input data ratios. The values of a certain ratio in the analysis sample may, e. g., range between 0 and 0.3 while the values of another ratio may vary between 0.3 and 13.5. Here, the input transformation gives rise to the scattering of the values of the two ratios within an interval to be defined in advance. Moreover, in the neurons of the input layer, a sigmoid transfer function may be employed in order to reduce the annoying effect of mavericks in the data material.
- **Network architecture.** The network architecture includes the number of layers of the neural network as well as the number of neurons in the individual layers.

Application-specific parameters:
- **Selection of (t-1)-, (t-2)-, (t-3)-, or (t-1/t-2/t-3)-data sets.** The analysis sample includes data records from three different periods, t-1, t-2 and t-3. The neural network can either be trained with all 672 data records of the analysis sample or with the 224 data records of solvent and insolvent companies resulting from one of the three periods.
- **Number and selection of ratios.** The neural network is able to classify with different numbers of ratios. 4 to 73 ratios were tested.

5. The Results of the Parameter Tests

The parameters listed above were, in accordance with the coordinate method, tested and improved with the test sample. The results of the parameter tests may be summarized as follows:

1.) Number of Iterations. While testing this parameter, it became apparent that, with a rising number of iterations, the Beta-error with the test sample goes down for some time before it increases again. In order to avoid overtraining, the network ought not to be trained by too many **learning steps**. Hence, several numbers of iterations were tested with every parameter adjustment, and only the best result with the test sample was taken account of. In connection with the cumulative generalized Delta-rule and a learning rate of 0.45, the best results were achieved with about 100,000 iterations.

2.) Learning rate. The learning rate has relatively little influence on the results of classification. In connection with the cumulative generalized Delta-rule, a learning rate of 0.45 supplied optimum results.

3.) Rule of learning. Even though the cumulative generalized Delta-rule requires more iterations, it leads to slightly better results than the generalized Delta-rule. Moreover, the cumulative generalized Delta-rule has the advantage of the error curve running in a relatively flat and continuous manner so that it is easy to find the optimum number of iterations.

4.) Transformation of values. As a transfer function, the sigmoid function also proved useful in the hidden layer and in the output layer.

The input transformation has the greatest influence on the classification results. The best result will be obtained if the values of all ratios are transformed in such a way that they are scattering within an interval of -3 to +3, and the sigmoid transfer function is applied to the values thus transformed in the input layer. The sigmoid transfer function in the input layer transforms, in a second step, the ratio values into an interval of 0 to 1. By means of this second transformation with the sigmoid transfer function, the effect of mavericks in the data material is reduced, since the highest absolute values come closer together. The following figure shows the test results as to the sigmoid and linear transfer function with different upper and lower limits for the input transformation:

5.) Network architecture.[30] The following figure shows the test results with different pieces of network architecture. Behind the letters **I**, **H**, and **O**, the number of neurons in the **i**nput layer (I), the **h**idden layer(s) (H) and the **o**utput layer (O) is stated. In the following figure only results are included from tests of backpropagation networks with four input neurons (I4) and with one output neuron (O1) and different numbers of hidden layers with different numbers of neurons.

The number of neurons in the hidden layer had no significant influence on the classification results. When a test without a hidden layer (I4/O1) was performed, the results of classification only worsened slightly. With more than one hidden layer (I4/H4/H4/H4/O1) the training obviously required more iterations (3.44 million), and the classification results also changed for the worse.[31] The best results were obtained with a very simple and quickly to be trained network that only had one neuron in an intermediate layer (I4/H1/O1).

6.) Selection of (t-1)-, (t-2)-, (t-3)-, or (t-1/t-2/t-3)-data sets
The neural network could classify best when being trained with the 672 data records from all three periods. Results that were almost as good could be achieved with the 224 data records from the period three years prior to the insolvency. The data records resulting from the periods one year and two years prior to the insolvency were apparently less suitable for classification. The most interesting result is that almost the same (optimium) Beta-error is achieved if the t-3 (the oldest) data sets are used for the training as if the total analysis sample was used.

30 For the dimension of neural networks see KRATZER, K. P., Neuronale Netze, pp. 148-153. For the learning capacity see FIESLER, E./ CAULFIELD, H. J./ CHOUDRY, A., Some Theoretical Upperbounds on the Capacity of Neural Networks, pp. 51-58.

31 REHKUGLER/ PODDIG also arrive at the conclusion that the use of more than one hidden layer is without any advantage, see REHKUGLER, H./ PODDIG, TH., Klassifikation von Jahresabschlüssen, p. 15.

With regard to the annual financial statement ratio values of solvent and insolvent companies, the data records resulting from the periods t-1 and t-2 are obviously more different than the data records resulting from t-3. And yet, the less evident differences in the ratio values of the data records from t-3 are more suitable to increase the sensitivity of the backpropagation network to such an extent that even difficult cases can be classified correctly and as early as possible.

7.) Number and selection of ratios

With regard to this parameter that was finally tested, we present the classification results with the test and validation sample from three tests with different numbers of ratios. The following three selections of ratios were tested:

1. The best solution of the backpropagation network which was arrived at with the four ratios (**BPN-4**) selected by means of a discriminant function.[32] This test reveals the opportunity to improve the classification results of the discriminant analysis by means of a neural network in a second step.
2. The best solution of the backpropagation network achieved with all 73 ratios (**BPN-73**). These 73 ratios were merely arrived at on the basis of plausible thinking without preliminary statistical studies on their discriminant capacity.[33]
3. A selection of the backpropagation network of 13 ratios out of the 73 ratios (**BPN-13**). In the course of this test, it was checked by means of the backpropagation network how many of the 73 ratios were actually necessary for classification. Networks with 5 - 15 ratios became the subject of training.Those ratios which had had the highest connection weights in the network BPN-73 were determined as ratios in this test.

The 13 ratios selected by the backpropagation network itself were the following:

32 See FEIDICKER, MARKUS, Kreditwürdigkeitsprüfung - Entwicklung eines Bonitätsindikators, pp. 158-162 and our table 2.

33 See FEIDICKER, MARKUS, Kreditwürdigkeitsprüfung - Entwicklung eines Bonitätsindikators, pp. 55-69.

Ratio	Definition
R_06	cash flow I: gross performance
R_10	profit/net loss for the year: sales revenues
R_11	profit/net loss for the year: balance sheet total
R_13	(gross performance - raw materials and consumables - ordinary depreciation - sundry operating expense): balance sheet total
R_15	cash flow I: outside capital
R_28	(commitments + trade creditors) * 360: sales
R_43	short-term outside capital: balance sheet total
R_44	short-term due to banks: outside capital
R_52	equity capital I: (balance sheet total - liquid funds - land and buildings)
R_55	total level of debt: sales
R_59	(work in progress + finished goods/merchandise): sales
R_63	(trade creditors + commitments) * 12: raw materials and consumables
R_73	total level of debt (without "down payments for work in progress"): sales

Table 4: The 13 ratios selected by the backpropagation network

The networks with 4, 13 and 73 ratios only differ in the number of the financial statement ratios used and, accordingly, also in the number of neurons in the input layer. All other parameters adjustments did not have to be changed. The networks BPN-73 and BPN-13 show the classification results that can be obtained merely by making use of neural networks without any statistical inquiries as necessary for the MDA. In the following figure, the classification results with the discriminant function serve as a means of comparison (**MDA**):

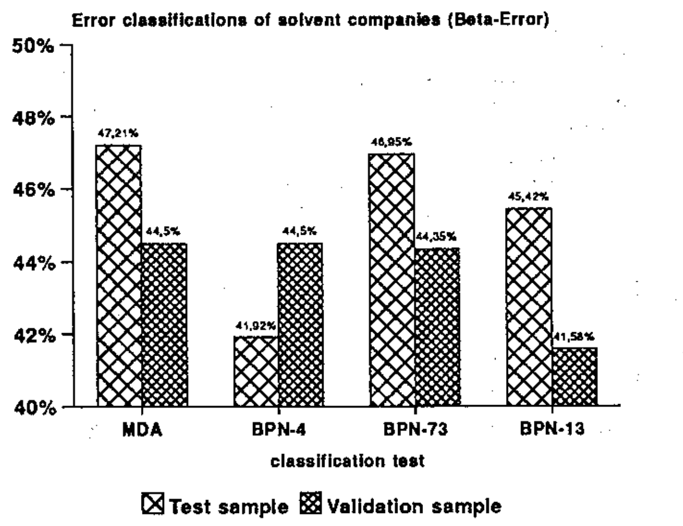

Figure 2: Survey of the best results

For the **test sample, BPN-4** supplies, with 41.92 %, the lowest Beta-error, which is by five per cent lower than the Beta-error with the MDA. However, the Beta-error of BPN-4 with the **validation sample** is 44.5 % and thus just as high as the Beta-error with **MDA**. The observation that the quality of the results of **BPN-4** differs with regard to the test and validation samples can possibly be attributed to the fact that the parameter adjustment of BPN-4 is the result of a long-lasting search process largely dependent on the results of the test sample. This parameter adjustment obviously depends on the test sample so strongly that the latter is clearly better classified than the validation sample. However, only the classification results obtained by means of the **validation sample** are representative for companies that are unknown to the network. Consequently, BPN-4 classifies just as well as the MDA.[34]

BPN-73 classifies the test sample as well as the validation sample to a slightly better degree than MDA. This result is remarkable because it shows that a sim-

34 ERXLEBEN ET AL. arrived at the same result, see ERXLEBEN, K./ BAETGE, J./ FEIDICKER, M./ KOCH, H./ KRAUSE, C./ MERTENS, P., Klassifikation von Unternehmen, S. 1256.

ple backpropagation network without preliminary statistical inquiries can classify just as well as the discriminant analysis with preliminary statistical studies on the selection of ratios requiring a lot of effort.

With the 13 ratios selected by the network itself the backpropagation network **BPN-13** was able to improve the classification of the **test sample** by about two per cent and the **validation sample** by about 2.9 per cent as compared to MDA.[35] This result shows that a backpropagation network selecting its ratios by means of network-specific methods can not only classify to a better extent than a backpropagation network without a selection of ratios but also to a better extent than the multivariate linear discriminant analysis selecting its ratios with the help of statistical methods. However, this result also shows that a backpropagation network is not absolutely resistent to unsuitable ratios, for BPN-13 classifies about 2.8 per cent better than BPN-73 even though the 73 ratios of BPN-73 also include the 13 ratios of BPN-13.[36]

As the restrictive assumptions (e. g. normally distributed ratio values in the parent population) concerning the used ratios that apply to the multivariate linear discriminant analysis do not hold true for the neural network, the neural network can classify better and is more universally applicable. In the case of the neural network, not only quantitative but also qualitative data can be taken into account. If such qualitative data were collected, the result of classification could once again be clearly improved by means of this integrated approach.[37] With regard to neural networks, it must however be criticized that the decision-making process of neural networks is not transparent and, in a lot of cases, not even plausible. The neural network that proved capable of supplying the best classification results in the course of this inquiry is relatively simple. And yet, the decision-making process of this system cannot be made transparent in such an uncomplicated way as it is possible with a multivariate linear discriminant function.

35 REHKUGLER/ PODDIG also achieve satisfactory results by making use of 13 ratios selected by the neural network, see REHKUGLER, H./ PODDIG, TH., Klassifikation von Jahresabschlüssen, p. 25.

36 The complete results of all parameter tests have been published in: KRAUSE, CLEMENS, Kreditwürdigkeitsprüfung mit Neuronalen Netzen, pp. 148-194.

37 FISCHER already argues in favour of a combined analysis of qualitative and quantitative data within a pattern recognition system, see FISCHER, J. H., Computergestützte Analyse der Kreditwürdigkeit auf Basis der Mustererkennung, p. 268.

Bibliography

BAETGE, JÖRG, "Möglichkeiten der Früherkennung negativer Unternehmensentwicklungen mit Hilfe statistischer Jahresabschlußanalysen", ZfbF, 41 (1989), pp. 792-810 (Möglichkeiten der Früherkennung).

BAETGE, JÖRG/ BEUTER, HUBERT B./ FEIDICKER, MARKUS, Kreditwürdigkeitsprüfung mit Diskriminanzanalyse, WPg, 45 (1992), pp. 749-761 (Kreditwürdigkeitsprüfung mit Diskriminanzanalyse).

BRAUSE, RÜDIGER, Neuronale Netze. Eine Einführung in die Neuroinformatik, (Stuttgart, 1991) (Neuronale Netze).

BREUER, ROLF-E., "Bilanzanalyse aus der Sicht der Kreditinstitute", ZfbF (special number), 43 (1991), pp. 151-155 (Bilanzanalyse aus der Sicht der Kreditinstitute).

ERXLEBEN, KARSTEN/ BAETGE, JÖRG/ FEIDICKER, MARKUS/ KOCH, HEIDI/ KRAUSE, CLEMENS/ MERTENS, PETER, Klassifikation von Unternehmen - ein Vergleich von Neuronalen Netzen und Diskriminanzanalyse, in: ZfB, 62 Jg., 1992, S. 1237-1262 (Klassifikation von Unternehmen).

FEIDICKER, MARKUS, Kreditwürdigkeitsprüfung - Entwicklung eines Bonitätsindikators dargestellt am Beispiel von Kreditversicherungsunternehmen, Schriften des Instituts für Revisionswesen der Westfälischen Wilhelms-Universität Münster, ed. Jörg Baetge, (Düsseldorf, 1992) (Kreditwürdigkeitsprüfung - Entwicklung eines Bonitätsindikators).

FISCHER, JÜRGEN H., Computergestützte Analyse der Kreditwürdigkeit auf Basis der Mustererkennung, Betriebswirtschaftliche Schriften zur Unternehmensführung, Vol. 23: Kreditwesen, (Düsseldorf, 1981) (Computergestützte Analyse der Kreditwürdigkeit auf Basis der Mustererkennung).

FIESLER, EMILE/ CAULFIELD, H. J./ CHOUDRY, A., Some Theoretical Upperbounds on the Capacity of Neural Networks, in: Proceedings of the First Workshop on Neural Networks: Academic/ Industrial/ NASA/ Defense, eds. NASA, Auburn (Alabama) 1990, pp. 51-58 (Some Theoretical Upperbounds on the Capacity of Neural Networks).

HECHT-NIELSEN, ROBERT, Neurocomputing, (Reading/ Massachussetts, 1990) (Neurocomputing).

HENO, RUDOLF, Kreditwürdigkeitsprüfung mit Hilfe von Verfahren der Mustererkennung, (Bern, Stuttgart, 1983) (Kreditwürdigkeitsprüfung mit Hilfe von Verfahren der Mustererkennung).

HERTZ, JAMES/ KROGH, ANDERS/ PALMER, RICHARD G., Introduction to the Theory of Neural Computation, (Redwood City, 1991) (Theory of Neural Computation).

KRATZER, KLAUS PETER, Neuronale Netze, (München, Wien, 1990) (Neuronale Netze).

KRAUSE, CLEMENS, Kreditwürdigkeitsprüfung mit Neuronalen Netzen, Schriften des Instituts für Revisionswesen der Westfälischen Wilhelms-Universität Münster, ed. Jörg Baetge, (Düsseldorf, 1993) (Kreditwürdigkeitsprüfung mit Neuronalen Netzen).

LAMPE, WINFRIED, "Aussagen quantitativer Kreditnehmeranalysen", Innovationen im Kreditmanagement. Entscheidungshilfen, Serviceaufgaben und Controlling im Firmenkundengeschäft der Banken, eds. Hans J. Krümmel and Bernd Rudolph, (Frankfurt a. M., 1985), S. 204-207 (Aussagen quantitativer Kreditnehmeranalysen).

NEURALWARE INC. (ed.), Neural Computing, (Pittsburgh 1991) (Neural Computing).

NEURALWARE INC. (ed.), Reference Guide, (Pittsburgh 1991) (Reference Guide).

NIEHAUS, HANS-JÜRGEN, Früherkennung von Unternehmenskrisen, Schriften des Instituts für Revisionswesen der Westfälischen Wilhelms-Universität Münster, ed. Jörg Baetge, (Düsseldorf, 1987) (Früherkennung von Unternehmenskrisen).

ODOM, MARCUS D./ SHARDA, RAMESH, A Neural Network Model for Bankruptcy Prediction, in: International Joint Conference on Neural Neworks, ed. IEEE, (San Diego, 1990), Vol. 2, pp. 163-168 (A Neural Network Model for Bankruptcy Prediction).

REHKUGLER, HEINZ/ PODDIG, THORSTEN., Klassifikation von Jahresabschlüssen mittels Multilayer-Perceptrons - Erste Ergebnisse und weiterführende Fragestellungen, Bamberger Betriebswirtschaftliche Beiträge Nr. 87/1992, (Bamberg, 1992) (Klassifikation von Jahresabschlüssen).

RÖSLER, JOACHIM, "Die Entwicklung der statistischen Insolvenzdiagnose", Krisendiagnose durch Bilanzanalyse, ed. Jürgen Hauschild, (Köln, 1988), pp. 102-114 (Die Entwicklung der statistischen Insolvenzdiagnose).

SCHUMANN, M./LOHRBACH, T./BÄHRS, P., Versuche zur Kreditwürdigkeitsprognose mit Künstlichen Neuronalen Netzen, Arbeitspapier Nr. 2, (Göttingen, 1992) (Kreditwürdigkeitsprognose mit Künstlichen Neuronalen Netzen).

TAM, KAR YAN/ KIANG, MELODY, "Predicting Bank Failures: A Neural Network Approach", Applied Artificial Intelligence, 4 (1990), pp. 265-282 (Predicting Bank Failures).

ZANGWILL, WILLARD I., Nonlinear Programming, A Unified Approach, (Englewood Cliffs, 1969) (Nonlinear Programming).