Jarosław KUREK[1], Bartosz ŚWIDERSKI[1] , Stanisław OSOWSKI[2]

University of Life Sciences (1), Warsaw University of Technology and Military University of Technology(2)

Assessment of financial condition of companies using neural networks

Abstract. *The paper presents the new automatic method of assessment of the financial condition of companies using neural networks. The aim of this assessment is to determine the insolvency risk of company. This information is important for worldwide transactions at which the insurance institution must assess the level of insurance risk of the import-export transaction. On the basis of this assessment the insurer can determine the insolvency of each side of transaction and decide of the height of the insurance. The paper will present an automatic method of such assessment by applying neural networks as the classifiers. The important part of this task is the optimal selection of the diagnostic features, on the basis of which the neural network will undertake the final decision and final classification of the company to the proper group. This is done by the neural classifiers.*

Streszczenie. *Praca poświęcona jest opracowaniu metody oceniającej w sposób automatyczny kondycję finansową firmy za pomocą sieci neuronowych. Szacowanie kondycji finansowej firm pozwala na określenie ryzyka niewypłacalności firmy (jedna ze stron transakcji) czy też na decyzje ubezpieczyciela transakcji, czy taka transakcja warta jest ubezpieczenia oraz na jaką kwotę (instytucja ubezpieczająca transakcje międzynarodowe import-eksport). Istotny zatem jest przydział ryzyka niewypłacalności firmy do jednej z kilku klas, określającej stopień aktualnej kondycji finansowej firmy. W pracy zaproponowano system komputerowy, w którym na podstawie wielu cech diagnostycznych (nie tylko finansowych) klasyfikator neuronowy ocenia automatycznie poziom ryzyka związany z klasą.* ***(Szacowanie kondycji finansowej przedsiębiorstw przy pomocy sieci neuronowych)***

Keywords: financial condition assessment, insolvency risk, neural networks.
Słowa kluczowe: szacowanie kondycji finansowej, ryzyko niewypłacalności, sieci neuronowe.

Introduction

The well-know problem in business world is how to avoid customers, who may cause problems with financial liquidity or are the potential bankrupts (bad customers). The important point is to acquire and then use the information that can be helpful in recognition of such customers before transaction [4,6].

On the basis of all gathered information we can create the diagnostic features describing the financial state of the company, and then associate them with one of a few classes, representing the level of insolvency risk. The insurance companies apply different number of classes. In this paper we assume 5 classes of insolvency risk:

- excellent (without any risk)
- good
- satisfactory
- passable
- poor.

To assign the class to the company the expert knowledge or special IT program is needed. The IT program should conduct the user through all steps leading finally to the assessment of the insolvency risk assigning it to one of the mentioned classes. In this paper we propose special automatic method which can greatly support the insurance company to undertake the decision to which class the customer belongs.

To build such program we must rely on proved up-to-date information about company converted to the form of numerical representation (diagnostic features). To get reliable results we should take into account the features that form the basis for expert decision. Such automatic system will be of great help for insurance company to automate business process, accelerate the process of assessment and also to save a lot of money on hiring highly specialized experts able to analyse very complex data base in order to assign the insolvency risk class to the customer. On the basis of the analysis of such database we are able to propose the numerical descriptors, characterizing the company.

The next step is the selection of the most important descriptors, strictly associated with the class. These descriptors will form the diagnostic features that will be supplied as the input data to the final classifier. In our solution we propose to apply Support Vector Machine (SVM) and fuzzy K-Nearest Neighbour (FKNN) as the classifiers for comparison purposes. SVMs are known from efficient performance, especially in the classification mode and fuzzy k-NN generating the value of membership of the input data to each class. The output result of SVM and fuzzy k-NN will be the class of insolvency risk associated with the input data.

Database

In our numerical experiments we have used the database of international companies of different profiles of activities (industry, commerce, services). The created database consists of 2217 trials and every trial contains financial statement of the last three years with the same type of consolidation (non-consolidated, consolidated or group consolidated). The type of consolidation is very important to assess the size of the company. Group consolidated finance regards the subject and its parent company presenting finance from the parent company point of view.

Table 1. The quantity of data of different types of consolidation in the considered database

Group consolidated	720
Consolidated	637
Non-consolidated	860
Total	**2217**

Consolidated financing statement concerns the subject and its subsidiaries, while the non-consolidated finance show only the subject financing statement. The division of the subjects due to the mentioned types of financing consolidation in our database is as following (table 1).

Table 2. The quantity of cases belonging to different groups of insolvency risk

Excellent	211
Good	800
Satisfactory	783
Passable	332
Poor	91
Total:	**2217**

The subjects under consideration were belonging to different classes of insolvency risk. The exact distribution of subjects to different classes is given in Table 2.

Fields in a credit report

Credit reports contain huge information and some of it can be useful for assessment of the financial condition of the subject. Credit report is usually divided into dozen sections depicting information regarding specific areas. The most important sections always included in the credit report are as following.

1. The identification - address, phone, fax, head office, etc.
2. Registry data – date established, legal form, registered status, registered authority, registration number.
3. Legal filings - bankruptcy filings, court judgements, tax, liens, etc.
4. Management & Staff – key managers (name, job title), staff data.
5. Board of Directors – appointments (name, board function, ID, address, biography)
6. Share Capital - authorised capital, type of share, number of shares, paid-up/issued capital, etc.
7. Shareholders – composition (name, percentage of shares, address), how listed, etc.
8. Corporate Affiliation – structure (name, affiliation type, address).
9. Financial accounts containing such positions as
    - date of account
    - consolidation
    - period
    - sales turnover
    - gross profit
    - operating profit
    - profit before tax
    - profit after tax
    - current assets
    - non-current assets
    - total assets
    - current liabilities
    - long-term liabilities
    - total liabilities
    - shareholders' equity
10. Payments – the payment behaviour (previous payment experience with the subject).

Features generation

Some of mentioned fields (if have numerical representation) can be directly used as the potential features in our automatic system. In other cases we should use either coding or special transformation of the data. Sometimes a field contains finite set of values so we associate every item of this set with the appropriate class. For example the field payment behaviour can be assigned only to one of 12 comments, which might be associated with 12 numerically represented classes (from 1 to 12). In the report the twelfth class corresponds to the serious payment delays. In contrast the first class comments acquired from the local and international trade sources generally concur that the subject's payments are met promptly.

To compare generated features based on financial account of various currencies we have recalculated the financial values to the same currency (USD). This is possible, since the credit reports contain information regarding historical exchange rate of every currency corresponding to 1 USD at the date of creation of the credit report. The available credit reports include three historical financial statements regarding the years 2007, 2008, and 2009. Sometimes the financial information presents the financial statement for 3, 6 or 9 months (instead of the whole year). For this reason we have recalculated financial values corresponding to three months periods.

The choice of appropriate features used in automatic classification should reflect the opinions of experts doing manually such assessment of companies. After such consultations we have chosen the set of basic attributes which (according to experts) has an important impact on the assessment results of financial condition of the company. These potential set of attributes is depicted in the table 3.

Table 3. The set of attributes taken from the credit reports as the potential features

Consolidation class	Total Assets for the last year
Sales turnover for the last year	Current liabilities for the last year
Profit before tax for the last year	Total liabilities for the last year
Current Assets for the last year	Trend of sales turnover
Shareholders' equity	Trend of profit after tax
Phone connection	Trend of shareholders' equity
Fax connection	Age of company (in years)
Legal form class	Number of employees
Number of employees for group	Share capital (authorised or issued/paid-up)
Payment behaviour	Class of sales turnover
If profit after tax <0	Class of positive profit after tax
Sum of profit and equity (to cover losses)	Registry status

The coding of the class regarding the sales turnover is made at assumption of five classes. Assignment for one of these 5 classes is based on the following rules and show how big the company is:

- if sales is more than 1 billion USD - class 1
- if sales belongs to <100 million-1 billion) USD –class 2
- if sales belongs to <20 million-100 million) USD –class 3
- if sales belongs to <2 million-20 million) USD –class 4
- if sales is less than 2 million USD –class 5

The profit after tax is usually assigned to two classes. If this profit is below zero we assign the class 0 and otherwise 1.

On the basis of the presented above attributes we have generated additional features being the ratios of some of them. The most important are financial ratios which show the level of financial condition of company, although they don't take into account the size of the company (for example the level of sales turnover). Financial ratios used by us are calculated using the following formulas [4,6,7]:

- Gross margin (*GM*)

$$GM = \frac{Gross \quad profit}{Sales \quad turnover}$$

- Operating margin (*OM*)

$$OM = \frac{Operating \quad profit}{Sales \quad turnover}$$

- Net profit margin (*NPM*)

$$NPM = \frac{\Pr ofit \quad after \quad tax}{Sales \quad turnover}$$

- Return on equity (*ROE*)

$$ROE = \frac{\Pr ofit \quad after \quad tax}{Shareholders's \quad equity}$$

- Return on assets (*ROA*)

$$ROA = \frac{\Pr ofit \quad after \quad tax}{Total \quad assets}$$

- Current Ratio (*CR*)

$$CR = \frac{Total \quad assets}{Total \quad liabilities}$$

- Debt Ratio (*DR*)

$$DR = \frac{Total \quad liabilities}{Total \quad assets}$$

- Long Term Ratio (*LTR*)

$$LTR = \frac{Long\,term \quad liabilities}{Share\,holders' \quad equity}$$

The last part of generated features is connected with Altman's variables, known as z-score. The Z-score formula for predicting bankruptcy was published in 1968 by Edward I. Altman [1]. The formula may be used to predict the probability that a company will go into bankruptcy within two years. Taking them into account we have added the following variables derived from the z-score:

$$T_1 = \frac{Current \;\, assets \quad - \quad Current \;\, liabilities}{Total \quad assets}$$

$$T_2 = \frac{Operating \quad profit}{Total \quad assets}$$

$$T_3 = \frac{Shareholders' \quad equity}{Total \quad liabilities}$$

$$T_4 = \frac{Sale \quad turnover}{Total \quad assets}$$

The last potential feature proposed by us is the maximum credit limit for the subject (recalculated to USD). Information regarding calculation of this credit limit is confidential and can't be revealed in the data base.

Taking into account the original attributes and the relative numerical descriptors defined above we got the set of 37 potential features that can be used in classification of the insolvency risk assessment.

All attributes have been normalized in a statistical way by applying the general formula

$$x = \frac{x - x_{mean}}{std(x)} \qquad (1)$$

In this expression x_{mean} represents the mean value of the attribute *x* and *std*(*x*) is the standard deviation.

Feature selection

It is well known fact, that features have different impact on the process of classification [2,3]. Good feature should be very stable for examples belonging to the same class (the smallest possible variance) and at the same time it should differ significantly for different classes. Thus important problem in classification and machine learning is to find out the most important features, influencing the problem solution in a most distinctive way. Note that the elimination of some features leads to the reduction of the dimensionality of the feature space and improvement of performance of the classifier in the testing mode at the data not taking part in learning.

From the set of 37 potential features we have to choose the best ones, characterizing the insolvency risk of the company in a most efficient way. In our approach to feature selection we have applied the sequential feature selection method. This method selects a subset of features from the whole set (the data matrix **X**) that provides the best result of insolvency assessment (vector **y**) by sequentially selecting features until there is no improvement in classification accuracy [4]. It is a wrapper method, strictly associating the feature selection process with the classification of the learning data.

We can apply different classifier systems used in selection, for example multilayer perceptron (MLP), Support Vector Machine (SVM), K Nearest Neighbours (KNN), fuzzy networks, etc. At the beginning of the process we have to define the criterions used to select features and to determine when to stop. Starting from an empty feature set, sequential feature selection creates candidate feature subsets by sequentially adding each of the features not yet selected. For each candidate feature subset, sequential feature selection performs 10-fold cross-validation by repeatedly calling them with different training subsets formed from rows of input data gathered in matrix **X** and the appropriate destinations y associated with them. The training samples are called *XTRAIN* and *ytrain*, while the appropriate testing subsets are called *XTEST* and *ytest*. The result of application of the method is the logical vector indicating which features are finally chosen as the best. Typical loss measures include the number of misclassified observations for classification models at the validation mode. The feature selection process has been implemented in Matlab [8] at application of SVM as a classification tool.

After applying the mentioned above sequential feature selection we have obtained the optimal small set of features, indicated by the method as the best one for our classification problem. After these experiments we have got five most discriminative features:

1) Existence of phone connection to the company
2) Maximum credit limit
3) Operating margin
4) Class of sales turnover
5) Class indicating the profit of company after tax

The last step of the procedure is the true classification of the input vectors composed of the chosen features. We have tried two different classifiers. One of them is the Support Vector Machine and the second fuzzy k-nearest neighbours classifier.

SVM classifier system

After optimal feature selection the next step is to apply the neural SVM classifier able to associate the input data to 5 classes of insolvency risks of the company. We have applied SVM of the Gaussian kernel [10,11]. The SVM classifier is a simple circuit structure of one hidden Gaussian kernel layer and one output linear unit performing the weighted summation. We have chosen the Gaussian kernel due to its universal character and very good performance in comparison to other choices, like linear, polynomial or spline functions.

The learning process of SVM network is relatively easy and effective since the whole learning task is simplified to the solution of the quadratic problem with linear constraints. In our experiments we have used the modified Platt algorithm, implementing the sequential optimization [9]. The hyperparameters: σ of the Gaussian function and the regularization constant C have been adjusted by repeating the learning experiments for the set of its predefined values and choosing the best one at the validation data sets.

To deal with a problem of many classes we have used one against one approach [10]. In this approach the SVM networks are trained to recognize between all combinations of two classes of data. For *M* classes we train *M*(*M*-1)/2 individual SVM networks. In the retrieval mode the vector **x** belongs to the class of the highest number of winnings in all combinations of classes.

Fuzzy k-nearest neighbours

The basic algorithm of nearest-neighbour classifier assigns an input sample vector of unknown class, to the class of its nearest neighbour. This idea is easily extended to more than one nearest neighbour, applying in practice *k*-nearest neighbours. In practical implementation we have used its well known variant – the fuzzy *k*-nearest neighbours (FKNN). The fuzzy *k*-nearest neighbours algorithm associates a sample vector **x** with the value of the class membership rather than assigning the vector to a particular class [5]. On the basis of the membership function to each of five classes we can finally associate the input vector **x** with the class of the highest membership value.

In computer implementation we have used the typical membership function, assigning the vector **x** to its neighbour belonging to ith class according to the expression [4]:

$$u_i(\mathbf{x}) = \frac{1}{\sum_{j=1}^{k}\left(\frac{d(\mathbf{x},\mathbf{x}_i)}{d(\mathbf{x},\mathbf{x}_j)}\right)^{2/(m-1)}} \tag{2}$$

In this expression $d(\mathbf{x},\mathbf{x}_j)$ represents the Euclidean distance from vector **x** to $\mathbf{x}_j$, k is the number of considered neighbours, and m is the exponent shaping the sharpness of the membership function (in our case m=2). The vector **x** is assigned finally to the class which won the largest sum of the membership functions among the neighbours .

The numerical results of experiments

In our numerical experiments we have applied both mentioned above classifiers. To get the most objective results we have used the cross validation approach. The data set containing 2217 examples of data was split into two independent parts: 80% examples are used for learning purposes and 20% for testing only. The experiments of learning and testing have been repeated 100 times at the random choice of data forming both sets. In the case of SVM we have applied the same values of hyperparameters: the width of Gaussian kernel σ=0.4 and the regularization constant C=1000. In the case of fuzzy KNN approach we have tried different number of neighbours k= [5 10 15 20 25 30], registering the results for each of them.

The statistical results of numerical experiments using FKNN are presented in table 4. The mean values of error depend on the number of neighbours. The best result (13.57% of mean error) corresponds to 20 neighbours. The standard deviation of error was reasonably stable in all experiments and close to 2%.

Table 4. The results of insolvency risk prediction of companies for FKNN at chosen values of K

	Mean of error	Standard deviation of error
k=5	18.11%	2.2%
k=10	14.48%	2.0%
k=15	14.93%	2.4%
k=20	13.57%	2.1%
k=25	16.74%	2.3%
k=30	14.93%	2.0%

Table 5 presents the appropriate results of experiments at application of SVM. For the comparison we have added also the best results of FKNN. As it is seen FKNN is evidently more efficient. The mean error of classification of the data is much smaller.

Table 5. Comparative results of insolvency risk prediction of companies for SVM and the best FKNN

	Mean of error	Standard deviation of error
SVM	19.2%	2.7%
Fuzzy k-NN (k=20)	13.57%	2.1%

Conclusions

On the basis of many performed numerical experiments we may conclude that application of the neural classifiers for insolvency risk prediction of companies is an interesting proposition, that should be further developed. At the actual obtained accuracy the system may be treated as the subsidiary tool for the experts involved in assessment of the financial status of the companies. The results suggest better performance of the FKNN in comparison to SVM.

The additional advantage of FKNN is the form of output. Instead of crisp result indicating the class of insolvency risk we get the values of the membership of the input data of the company to all classes. These results are much more descriptive for the financial condition of company and represent better starting point for expert considerations.

Surprisingly the application of the feature selection process has chosen only 5 most discriminative features from the classification point of view. Some of important features usually taken into account (such as size of company, exact level of sales turnover or potential loss after tax) have been skipped by our feature selection algorithm. This might be the associated with the specific type of the considered data.

Of course the great influence on these results had the values of parameters controlling the sequential feature selection algorithm. We have accepted the parameters leading to the best results of classification of our set of companies. To create the universal program applicable to other types of companies we should repeat the whole procedure of feature selection and try to find the optimal set of controlling parameters.

REFERENCES

[1] Altman E., Managing Credit Risk, N.Y.: Wiley, 2008

[2] Gately E., Neural networks for financial forecasting. Wiley, New York, 1996

[3] Guyon I., Elisseeff A., An introduction to variable and feature selection. Journal of Machine Learning Research. 3, 2003, 1158 – 1182,

[4] Huang Z., Chen H., Hsu C.J, Chen W.H., Wu S., Credit rating analysis with SVM and neural networks: a market comparative study. Decision Support Systems, 2004, vol. 37, pp. 543-558.

[5] Keller J.M., Gray M.R., Givens J.A., A fuzzy k-nearest neighbour algorithm, IEEE Trans. Systems, Man and Cybernetics, vol. 15, no. 4, 580-585, 1985

[6] Lee K.C, Han I., Kwon Y., Hybrid Neural Network models for bankruptcy prediction. Decision Support Systems, 1996, vol. 18, pp. 63-72.

[7] Malhorta R., Malhorta D.K., Evaluating consumer loans using neural networks. Omega, 2003, vol. 31, pp. 83-96.

[8] Matlab user manual, MathWorks, Natick, 1999

[9] Platt, L. Fast training of SVM using sequential optimization, (in Scholkopf, B., Burges, B.,& Smola, A., Eds. Advances in kernel methods – support vector learning. Cambridge: MIT Press), 1998. pp. 185-208

[10] Schölkopf B., Smola A., Learning with Kernels, Cambridge, MA, MIT Press, 2002

[11] Vapnik V., Statistical Learning Theory, N.Y.: Wiley, 1998.

***Authors**: dr inż. Jarosław Kurek, Warsaw University of Life Sciences, Faculty of Applied Informatics and Mathematics, Email: jaroslaw_kurek@sggw.pl, dr Bartosz Świderski Warsaw University of Life Science, Faculty of Applied Informatics and Mathematics, Email: jbswiderski@wp.pl, prof. dr hab. inż. Stanisław Osowski Warsaw University of Technology, Electrical Engineering Department and Military University of Technology, Warsaw, Email: sto@iem.pw.edu.pl*