CONSTRUCTION ELEMENTS OF BANKRUPTCY PREDICTION MODELS IN MULTI–DIMENSIONAL EARLY WARNING SYSTEMS

Jarosław Kaczmarek*

Abstract: A consequence of the inevitability of the occurrence of internal crises in companies is the taking of preventive action in place of purely remedial measures. In this respect a significant role is played by *Early Warning Systems* (EWS), which provide early warning information and financial threat assessments relating to the continuation of operations and bankruptcy not only for individual companies as such but also for companies as a whole. The limitations of existing models used for EWS purposes have led to the elaboration of new models, estimated on one of the largest hitherto drawn up teaching sets, constituting more than five hundred bankrupt companies. These models also distinguish themselves through the application of innovative methods and precise instruments; the structural concept of these models for multi–dimensional EWS purposes, accompanied by elements used for predicting, is presented in this article.

Key words: bankruptcy, bankruptcy prediction, early warning system, logistic regression model

JEL Codes: G33

Introduction

Thinking of early warning systems one has in mind a solution which takes into account the systematic diagnosis of a company's financial standing, appropriately early identification and interpretation of warning signals and assessment of the risk of financial threat to the continuation of activities, company bankruptcy and companies as a whole (classes, groups, departments, sections).The result of such assessments may be the taking of action in support of company restructuring processes; this is because if restructuring is performed effectively and in a timely manner, it can save a company, and the expenses incurred by the State for this objective are much lower than the economic and social costs caused by the collapse of companies [11].

Bearing the above in mind a concept of an institutionalised instrument was elaborated; its basic function in the first component, constituting the core of the *Early Warning System*, is the current monitoring and analysis of company standing

* **PhD. Jarosław Kaczmarek**, Cracow University of Economics, Poland;
✉ corresponding author: kaczmarj@uek.krakow.pl

as a whole and informing about the need to take remedial action in terms of entities (as a whole), in the face of their deteriorating economic and financial standing.

The remaining elements are: programming, implementation and evaluation components.

Basic structure of research procedure

The approved main stages of research procedure in terms of the monitoring component are performed in the form of four analytical paths and the creation of company classification sets (isolated approach) and of an integrated set (integrated decision rule), submitting it next for analysis and systematising the companies from the point of view of level of threat to the financial continuation of operations and bankruptcy.

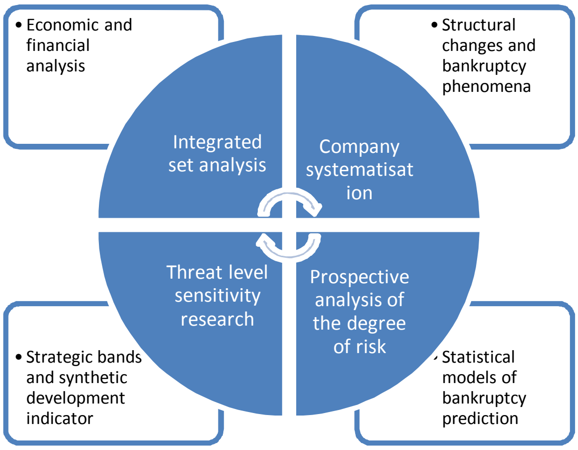

Figure 1. Analytical routes, creation and analysis of an integrated set

Source: own work.

Up until now a shortcoming of research conducted in Poland concerning early warning was the considerable lapse between the appearance of signs of threat to the financial continuation of operations and bankruptcy, and the obtaining of analyses and assessments in this area. The proposed solution permits one to switch the conducted analyses to quarterly frequency, which considerably shortens the time required to recognise changes in the state of companies and the taking of the appropriate decisions in the area of structural policy.

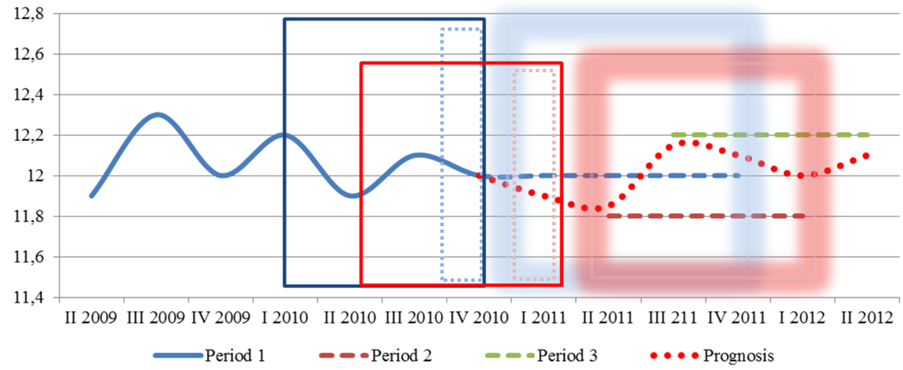

Fig. 2. Presentation of a rolling method of determining the period of bankruptcy prediction

Source: own work.

Research carried out on a quarterly basis [13] means that the estimated threat of bankruptcy retains "contractual validity" for a period of one quarter; this is because the next conducted analysis provides new information on the value of the analysed measure, referring to the next started annual prediction period (see fig. 2).

Bankruptcy prediction model creation layout

Measures taken as part of the research route of bankruptcy prediction statistical models designate the taking up of the following analytical activities [15]:

- identification of dependency of layout:company standing – threat of bankruptcy,
- susceptibility testing: level of threat of bankruptcy – abrupt changes in the determinants of company standing,
- testing of efficiency (identification and dynamic prediction methods),
- prospective analysis and assessment of the state of threat of bankruptcy.

The standard instruments for predicting the threat to company bankruptcy include discriminatory models and logistic regression models. Compared to the new generation methods – such as neural networks – these are less expensive, more transparent, and their results are easier to interpret and compare. Moreover they allow to introduce the prior knowledge to the regression analysis, [AG]. The latter ability often leads to prediction error reduction, [AGRV]

In order to meet the requirement that the model, as far as possible, relate to the true conditions under which the researched companies operate, focus has been placed on Polish bankruptcy prediction models [8, pp. 713–726; 12, pp. 129–172]. Their common trait, however, is the application of a very small number of teaching sets, which may result in overestimation of their very high predictive ability.

The consequence is the restricted possibility of using these models on a large scale in the national Early Warning System, before the occurrence of risk of financial threat to the continuation of activities. Following research four bankruptcy prediction models were constructed, based on Firth's logistic regression (for companies in general, including production, trade and services companies), elaborated on the basis of a set of 530 bankrupt companies and 2,650 non–bankrupt companies. In terms of Polish specialist literature, and not only, this is one of the largest hitherto drawn up company data sets considered in the context of modelling the level of bankruptcy threat. Estimated models are characterised by a high capacity to predict the state of company threat of bankruptcy and the application of innovative methods and construction tools.

At this point the main structural elements of bankruptcy prediction models shall be presented together with the prediction of company threat of bankruptcy.

Methods of creating the teaching set

At this stage the most frequently adopted approach involves the collection of a data set on bankrupt companies, followed by the expert matching of companies which have not gone bankrupt ("1 to 1" method). This approach, however, is not possible in the case of large data sets (this is the kind of set which constitutes the basis of estimations relating to the discussed bankruptcy prediction models).

For this reason the *case*–control technique has been adopted for the purpose of selecting companies. This involves defining a number of key characteristics of the units and the matching of each unit with a distinguishing trait to a unit without such a trait, but which is most similar to it in terms of the variables used for matching. In this manner it was accepted that each bankrupt company would be accompanied by companies which have not gone bankrupt but which are similar in terms of the value of assets and net revenue on sales; matching would take place in consideration of PKD (Polish Classification of Activities) compliance and the legal and organisational form of the company.

The "1 to 1" matching approach is most often used, but from a theoretical point of view it is justified to even perform matching on a "1 to 5" basis (7, pp. 145–162); this approach was in fact used, in that each non–bankrupt company received during the bankruptcy model estimation process a weighted value equivalent to 1/5.

The use of the alternative technique of creating a data set for comparative purposes involved *frequency matching*. This technique does not require individual matching of non–bankrupt companies to bankrupt companies. However, what is required – much the same as in the case of the *case–control* method – is a definition of matching variables (e.g. the value of assets, level of employment, net revenue), followed by the matching of non–bankrupt companies so as to allow the layouts of variables to be as close as possible to each other in both groups of companies.

Choice of the basic bankruptcy prediction model

The prediction model creation stage is preceded by analysis of the considered one–dimensional explanatory variable distributions. These distributions were analysed both on the basis of their descriptive characteristics (average, deciles, dispersion measure), as well as appropriate graphs (histograms, estimators of density functions determined by means of nuclear estimators). The distributions were generated separately for bankrupt and non–bankrupt companies. Because of the sufficiently comprehensive set of data, the explanatory variable distributions were also indicated in terms of type of business activity (production, trade, services).

The next step involved the matching of empirical explanatory variable distributions and theoretical distributions. The quality of matching elementary distributions for one–dimensional variables, such as normal, log–normal and Weibull distribution, was verified. In the event of noting the absence of the matching of standard distributions application is made of certain general distribution classes – generalized lambda distribution and generalized beta distribution [15].

Naive Bayes classifier

The naive Bayes classifier was constructed on the basis of the above described one–dimensional analyses of explanatory variable distributions generated for bankrupt and non–bankrupt companies. It has been accepted that i–th object (company) will be described by means of p–dimensional vector of explanatory variables x_i. It has further been assumed that $f_k(x_{ik}|bankrupt)$ and $f_k(x_{ik}|non-bankrupt)$ are respectively density of likelihood functions for k–explanatory variable in the group of bankrupt and non–bankrupt companies.

The decision rule shall allocate i–th company to the group of bankrupt companies if:

$$\prod_{k=1}^{p} f_k(x_{ik}|bankrupt) > \prod_{k=1}^{p} f_k(x_{ik}|non-bankrupt)$$

Discriminant model

Fischer's linear discriminant analysis is one of the oldest classification methods which entails the finding of a linear combination of explanatory variables, which in the best possible manner differentiates between two (or more) groups of objects [15].

Discriminant analysis assumes that i–th object may be described by means of p–dimensional vector of explanatory variables x_i. Let $\bar{x}_1$ and $\bar{x}_2$ define average vector values appropriately in the first and second group of objects, and Sdefine the common co–variance matrix for both groups. The basic component of the classifying rule (decision rule) is vector β which is provided with the following analytical formula:

$$\beta = S^{-1}(\bar{x}_2 - \bar{x}_1)$$

Object x_i is classified to the second group if the following inequality is true:

$$\beta'(x_i - (\bar{x}_2 + \bar{x}_1)/2) > 0$$

The main assets of the discriminant model is the ease of calculation connected with the assignment of the model. Furthermore, in many cases the accuracy of prediction of the discriminant model is comparable with more advanced models, which is particularly evident for relatively small data sets [2].

Logistic regression model

For the needs of the conducted research comprehensive studies were carried out on the logistic regression model [1] and specialist software was created [3]. Logistic regression is one of the most popular tools for analysing binary data. Most often it is used for assessing the impact of independent variables on the likelihood of a certain event occurring.

In the classic logistic regression model it is assumed that the dependent variable $y_i \in \{0,1\}$ $(i = 1, \ldots, n)$ is subject to Bernoulli distribution with success likelihood of $F(x_i'\theta)$, where $F(\cdot)$ is the distribuant of logistic distribution, x_i is the p–dimensional vector of explanatory variables, and $\theta \in \mathbb{R}^p$ is (containing the intercept) p–dimensional vector of structural parameters [10, pp. 56–68].

In order to estimate θ parameters vector one may use the greatest likelihood method which involves the maximisation of the credibility function in relation to the structural parameters vector $\theta \in \mathbb{R}^p$ [7, pp. 145–162]:

$$L(\theta) = \prod_{i=1}^{n} F(x_i'\theta)^{y_i}[1 - F(x_i'\theta)]^{1-y_i}$$

The results of findings show that the assessments of logistic regression model parameters attained by means of the greatest credibility method are characterised by significant burden. Furthermore, in the event of complete separation, this method does not lead to the attaining of finished parameter assessments [6, pp. 2409–2419]. The standard approach in this case involves removal from the model of the variable or variables causing the problem, but their removal as the best way of describing the condition of the dependent variable, appears to be illogical.

The comparative analysis of two methods of estimating parameters which eliminate the problem of complete separation and which provide assessments of parameters of lower burden in comparison to the highest credibility method, has become the basis of subsequent conclusions in the estimation of the bankruptcy prediction model [2].

Logistic regression model with the latent explanatory variable (Hidden Logistic Model)

It is possible to assume the existence of a certain additional stochastic mechanism which prevents the true values of the dependent variable from being observable – the dependent variable is the latent variable. Expressed in another manner, it could happen that a given case is a success (failure), but because of the presence of the mentioned stochastic mechanism will be recognised as a failure (success) [14, pp. 315–332].

A basic matter is to define the method of estimating the parameters of such a defined model. The practical implementation of estimation, however, does not cause difficulty. A key element is the creation of so–called pseudo–observation $\tilde{y}_i$ for each of the original observations $y_i$according to the following formula:

$$\tilde{y}_i = (1 - y_i)\delta_0 + y_i\delta_1$$

The value of likelihoods δ_0 and δ_1 may only be recognised for very large data sets, and in practical terms one may assume that $\delta_0 = 0{,}01$, $\delta_1 = 0{,}99$ [14, pp. 315–332]. As an alternative, one may match values δ_0 and δ_1 in such a manner that the pseudo–observation sum be equal to the sum of original values of the variable being explained.

Next, one may adapt to the pseudo–observations formed in such a manner, a classic model of the greatest credibility, whose purpose is maximisation of the credibility function in terms of the $\theta \in \mathbb{R}^p$ structural parameter vector.

Firth's logistic regression model

Concerning Firth's logistic regression model the $s(\theta)$ function is replaced by its part modification:

$$s^*(\theta) = \sum_{i=1}^{n} \left(y_i - F(x_i'\theta) + h_i(\frac{1}{2} - F(x_i'\theta)) \right) x_i$$

where h_i is the diagonal elements of matrix $H = W^{\frac{1}{2}} X(X'WX)^{-1} X' W^{\frac{1}{2}}$, X is the data matrix, and W is the diagonal matrix with dimensions $n \times n$, whose i–th diagonal element is equal to $F(x_i'\theta)(1 - F(x_i'\theta))$.

Modification of the system of equations $s(\theta)$ is identical to modification of the credibility function in the following manner:

$$L^*(\theta) = L(\theta)|I_\theta|^{1/2}$$

where I_θ is the information matrix, whilst function $L^*(\theta)$ is referred to as the penalized likelihood function [6, pp. 2409–2419; 5, pp. 4216–4226]. The described approach significantly reduces the burden of parameter assessments of models [4, pp. 27–38], furthermore, modification of the original likelihood function contains interpretation from the Bayes point of view. This is the classic logistic regression model with Jeffreys prior distribution imposed on the parameters.

The comparative analysis of models and the demonstrated assets of Firth's logistic regression model have become the basis of accepting this model as fundamental in the analysis of company bankruptcy prediction for the needs of the *Early Warning System*. This measure, obtained by adopting the model, assumes a value of (0,1), in that its higher values indicate a higher likelihood of bankruptcy and, in principle, the possibility of bankruptcy, where chance is defined as the likelihood ratio of bankruptcy to the likelihood ratio of non–bankruptcy (in the context of one year). This permits, in a quantitative manner, to describe the scale of bankruptcy change in a dynamic sense and to compare the degree of threat between various classes and groups of companies. All of the previously constructed models do not display such properties [13].

Defining the efficiency of bankruptcy prediction models

Assessment of the effectiveness of the bankruptcy prediction method was performed with the use of typical measures used with object recognition systems or diagnostic systems, as shown below. Calculated measures of efficiency were as follows:

- sensitivity – likelihood of detecting a bankrupt company i.e. A/(A+B),
- specificity – likelihood of recognising a non–bankrupt company as not bankrupt i.e. D/(C+D),
- positive predictive capacity – likelihood that the company recognised by the model as bankrupt will in fact go bankrupt i.e. A/(A+C),
- negative predictive capacity – likelihood that the company recognised by the model as non–bankrupt will not go bankrupt i.e. D/(C+D),
- general model efficiency – (A+D)/(A+B+C+D).

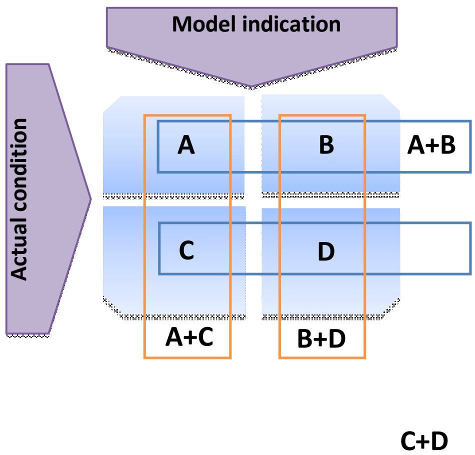

Fig. 3. Defining the efficiency of bankruptcy prediction models

Source: own work.

Comparison of various classifiers, financial indicator prediction capacity and the searching of so–called optimal cut–off points is made possible by the ROC (*Receiver Operating Characteristic*). This is a two–dimensional graph, which on the axis of ordinates presents *sensitivity*, and on the axis of abscissa presents *1 – specificity*, calculated for various cut–off points. One of the most often used classifier quality measures connected with the ROC is the area under the curve, referred to as AUC (*Area Under Curve*) defined as (where $y(x)$ is a function defining the ROC curve):

$$AUC = \int_0^1 y(x)\, dx$$

The AUC measure assumes values ranging [0,1], in that the higher the value the better the assessed decision rule [15]. In such cases it must be noted that the totally random model is represented by a straight line running through points (0,0) and (1,1); and what follows on is that the area under its graph is equal to 0,5. In this sense the structure of the classifier, for which the AUC is less than 0,5 is not practically justified. When two classifiers are compared, if the ROC which refers to the first of these is located above each point of the second ROC curve, the value of the AUC for the first classifier must be bigger than the value of the AUC for the second one. This dependency does not operate in reverse because there exists the possibility of tow ROC curves intersecting.

Predicting variables determining the degree of risk of bankruptcy

One of the traits of the *Early Warning System* involves prolongation of the period of predicting the degree of risk of bankruptcy. This means the application of a prediction method, however application is made not of the prediction of the risk of measurement values, but of prediction model variables. The predicted values of these variables have been introduced into prediction models, which provided the basis for concluding about the shaping of future levels of risk of bankruptcy and permitted the prolongation of the time line for possible conclusions concerning the next annual period (see fig. 2.).

Variables defining the degree of risk of bankruptcy have been predicted with the use of one–dimensional time series models because of the current length of the researched time series (for quarterly frequencies this length was 16, and for six–month frequencies it was 8.). This decision is justified by the findings of numerous empirical studies which demonstrate that multi–dimensional models used for short time series usually generate predictions burdened by greater error than those stemming from one–dimensional models.

Each studied time series has matched to it the best classification error level method SARIMA (*Seasonal Autoregressive Integrated Moving Average*) model and ETS (*Exponen Tial Smoothing*) model according to Hyndman. The predicted values were calculated as an average arithmetical of SARIMA model and ETS model predictions (predictions burdened with smaller error, confirmed through research, are obtained).

One–dimensional models

As a basic approach consideration was given to the application of the ARIMA class models (*Autoregressive Integrated Moving Average*) in predicting variables defining the degree of risk of bankruptcy. The y_t stochastic process is referred to as the ARIMA process (p, d, q) if d–fold differentiation brings process y_t to ARIMA form (p, q), which is defined according to the following formula:

$$y_t = \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} + \varepsilon_t + m_1 \varepsilon_{t-1} + \cdots + m_q \varepsilon_{t-q}$$

where ε_t is white noise with zero expected value and invariant in terms of time variance, α_i are parameters standing by the ARIMA process autoregression component, whilst m_j are parameters facing the moving average component [15].

Because of the seasonality of the data application was made of the SARIMA class models (*Seasonal ARIMA*). Determining the full specification of the SARIMA model is, in terms of concept, very similar to the process for determining the VAR model (*Vector Autoregression*).

Exponential smoothing is a certain class of prediction methods used for attaining a scoring prediction of the studied phenomenon for the future period.

An undoubted asset of the ETS models is the accessibility of studies describing the construction of prediction intervals not only based on the assumption of normality, but also the simulation method (Hyndman's work).

In order to specify the exponential smoothing model it is necessary to define its three components [15]:

- the component responsible for trend – this covers the "level coefficient" and the "growth coefficient". They can be connected in a number of ways and as a result they can define the process which has no trend or has a trend:
- additive, damped additive, multiplicative, damped multiplicative,
- seasonal fluctuations component – may be in the additive or multiplicative version or may not occur,
- random fluctuations component – may be in the additive or multiplicative version.

ETS model constituent parts expressed in this manner define dozens of various models, as a result of which it is possible to carry out an automatic manner of predicting time series (estimation of each series of possible models, choice of model based on the AIC, generation of scoring predictions on the basis of the best model).

Multi–dimensional models

As the elaborated *Early Warning System* is supposed to be a permanent solution, providing an increasing amount of information (longer time series), for this reason multi–dimensional models for predicting the threat of bankruptcy (VAR – *Vector Autoregression* and VECM – *Vector Error Correction Model*) were anticipated and elaborated for future use. However, because of the absence of their empirical verification at this stage of research, the description of adapting these models was omitted.

Summary

The elaboration of the concept of an institutionalised instrument, whose core is the national *Early Warning System*, required a new approach to three basic problems. The first involved the elimination of delay between the appearance of signs of threat to the financial continuation of operations and company bankruptcy, and the obtaining of analyses and assessments in this area. The solution which permits a switch to quarterly frequency analyses leads to more frequent updating of the estimated level of threat, giving the possibility of taking implementation measures.

The second solved problem involved the prolongation by one year of the period of predicting the level of threat of bankruptcy by applying prediction methods in estimating the probability of bankruptcy. This was achieved by forecasting prediction model variables, and not the measure of threat itself.

Thirdly, limitation through imitation transfer and the disadvantages of known bankruptcy prediction models have led to the basis of estimation of new models based on Firth's logistic regression on a unique set of more than three thousand companies, with the application of innovatory methods and tools of own construction and high levels of prediction capacity, which speaks in favour of the uniqueness of this solution concerning the modelling of level of risk of bankruptcy.

The results of empirical studies have been presented in separate publications [e.g. 9, pp. 27–31].

References:

[1] Fijorek K., *Aproksymacja modelu regresji logistycznej Firtha za pomocą ważenia obserwacji*, materiały z XLVII Konferencji Statystyków, Ekonometryków i Matematyków, Uniwersytet Ekonomiczny w Krakowie, Osieczany 2011
[2] Fijorek K., Fijorek D., *The effectiveness of best subset variable selection in Firth logistic regression,* Applied Informatics Methods 2011, accepted for publication

[3] Fijorek K., Sokołowski A., *Separation–resistant and bias–reduced logistic regression*, Statistica macro, 2011, accepted for publication
[4] Firth D., *Bias reduction of maximum likelihood estimates*, Biometrika 80/1993
[5] Grzybowski A.Z., *On some new method of incorporation prior information along with its uncertainty in regression estimation*, Scientific Research of the Institute of Mathematics and Computer Science, (5)2006, 23-31.
[6] Grzybowski A.Z., Vokorosova R., *Reducing a prediction error risk in some industrial and market problems – a cross validation analysis of some method of incorporation a prior information*,[in:] Strengthening the Educational and Scientific Collaboration Among Faculties of Economics within V4 and Countries of South Eastern Europe, R.Vokorokosova, J. K.Grabara. (eds.) , Journal of Scientific Works, Faculty of Management, Czestochowa Technical University, series Monographs No 17, 2009, 185-191.
[7] Heinze G., *A comparative investigation of methods for logistic regression with separated or nearly separated data*, Statistics in Medicine, 25/2006
[8] Heinze G., Schemper M., *A solution to the problem of separation in logistic regression*, Statistics in Medicine, 21/2002
[9] Hosmer D.W., Lemeshow S., *Applied Logistic Regression*, Wiley, New York 1989
[10] Juszczyk S., *Prognozowanie upadłości przedsiębiorstw,* "Ekonomista", 5/2010
[11] Kaczmarek J., *Przedsiębiorstwa w okresie kryzysu. Retrospektywna analiza zmian stopnia zagrożenia*, Przegląd Organizacji, 10/2011
[12] Long Scott J., *Regression models for categorical and limited dependent variables*, SAGE 1997
[13] Mączyńska E., Zawadzki M., *Dyskryminacyjne modele predykcji bankructwa przedsiębiorstw*, Ekonomista, No. 2/2006
[14] Prusak B., *Nowoczesne metody prognozowania zagrożenia finansowego przedsiębiorstw,* Difin, Warsaw 2005
[15] *The Report of the assessment of enterprises bankruptcy threat – micro economy component*, Kaczmarek J. (ed.), authors: Fijorek K., Kaczmarek J., Kolegowicz K., Krzemiński P., MSAP UEK, June 2011 (duplicated typescript)
[16] Rousseeuw, P. J., Christmann, A., Robustness against separation and outliers in logistic regression, Computational Statistics & Data Analysis, 43/2003
[17] *The proposals for the Rapid Reaction Facility and the strategy for its implementation in Poland – Early Warning System*, Kaczmarek J. (ed.), authors: Fijorek K., Kaczmarek J., Kolegowicz K., Krzemiński P., MSAP UEK, maj 2011 (duplicated typescript)

ELEMENTY KONSTRUKCJI MODELI PREDYKCJI UPADŁOŚCI W WIELOWYMIAROWYCH SYSTEMACH WCZESNEGO OSTRZEGANIA

Abstrakt: Konsekwencją nieuchronności występowania kryzysów wewnętrznych w przedsiębiorstwach jest podejmowanie działań zapobiegawczych w miejsce działań tylko o charakterze sanacyjnym. Istotną rolę spełniają w tym względzie *Systemy Wczesnego Ostrzegania* (SWO), dostarczając wyprzedzających informacji oraz ocen zagrożenia finansowego kontynuacji działalności i upadłości nie tylko dla pojedynczych przedsiębiorstw, ale i ich zbiorowości.

Ograniczenia zastosowania istniejących modeli dla potrzeb SWO, skłoniły do opracowania nowych modeli, estymowanych na jednym z największych jak dotychczas zbiorze uczącym, liczącym ponad pięćset przedsiębiorstw upadłych. Modele te wyróżnia także zastosowanie innowacyjnych metod i narzędzi szczegółowych, a ideę ich konstrukcji dla potrzeb wielowymiarowego SWO wraz z elementami prognozowania zawarto w niniejszym artykule.

破产预测模型元素在多维预警系统中的构建

摘要：在公司发生内部危机的一个不可避免的后果是采取预防措施代替纯的补救措施。在

这方面早期预警系统（EWS）发挥重要作用。它不仅能为个体公司，也能够为作为一个整

体的公司能提供有关运行以及破产的早期预警信息和财务风险的评估。

EWS的目前使用的模型的局限性导致新模型的产生，这是建立在对迄今超过五百家破产企

业的研究上。这些新模型还通过创新的方法和精密的手段来区分它们自己。本文介绍关于

多维预警系统模型的结构概念，以及预测中使用的元素。