Anton V. Lytvyn[1]

APPLYING SUPPORT VECTOR MACHINES TO FINANCIAL CRISIS FORECASTING IN UKRAINIAN INSURANCE COMPANIES

In this study, a set of support vector machine models for predicting financial crisis in Ukrainian insurance companies has been developed. Based on the application of the proposed feature selection algorithm, a list of relevant indicators has been formed. The comparative analysis of classification accuracy and sensitivity measures for two consecutive periods supported the choice of the best performing support vector model applicable to forecasting tasks.

Keywords: *support vector machines; kernel functions; insurance companies; financial crisis forecasting.*

Антон В. Литвин

ЗАСТОСУВАННЯ МЕТОДУ ОПОРНИХ ВЕКТОРІВ ДЛЯ ПРОГНОЗУВАННЯ ФІНАНСОВОЇ КРИЗИ В УКРАЇНСЬКИХ СТРАХОВИХ КОМПАНІЯХ

У статті розроблено набір моделей опорних векторів для прогнозування фінансової кризи в українських страхових компаніях. На підставі застосування запропонованого алгоритму вибору ознак сформовано перелік релевантних індикаторів. Порівняльний аналіз показників точності та чутливості класифікації для двох послідовних періодів дозволив обрати модель опорних векторів, найбільш прийнятну для завдань прогнозування.

Ключові слові: *метод опорних векторів; функції ядра; страхові компанії; прогнозування фінансової кризи.*

Форм. Рис. 1. Табл. 13. Літ. 25.

Антон В. Литвин

ПРИМЕНЕНИЕ МЕТОДА ОПОРНЫХ ВЕКТОРОВ ДЛЯ ПРОГНОЗИРОВАНИЯ ФИНАНСОВОГО КРИЗИСА В УКРАИНСКИХ СТРАХОВЫХ КОМПАНИЯХ

В статье разработан набор моделей опорных векторов для прогнозирования финансового кризиса в украинских страховых компаниях. На основе применения предложенного алгоритма выбора признаков сформирован перечень релевантных индикаторов. Сравнительный анализ показателей точности и чувствительности классификации для двух последовательных периодов позволил выбрать модель опорных векторов, наиболее применимую для задач прогнозирования.

Ключевые слова: *метод опорных векторов; функции ядра; страховые компании; прогнозирование финансового кризиса.*

Introduction. Ensuring business stability and continuity of financial companies has always been one of the main goals for both their managers and governmental bodies. Alongside with commercial banks and pension funds, insurance companies are the key entities at the financial market providing unique services to households and enterprises, as well as realizing different important functions at the national level. Due to the nature of insurers' activities, they have to deal with greater risks and in general face more significant financial threats. Despite the fact that many insurers have adopted crisis management procedures and the state regulator has been implementing contemporary prudential practice, effectiveness and objectivity of the finan-

[1] National University of "Kyiv-Mohyla Academy", Kyiv, Ukraine.

cial crisis management processes, and in particular, financial crisis prediction, in Ukrainian insurance companies remain low because of various factors such as low quality and accessibility of data or application of inappropriate methods. While it is difficult to address directly the data-related problems, the efficacy of financial crisis forecasting in insurance companies can be improved by choosing and applying justified mathematical modelling methods that show reasonable performance on scarce and statistically heterogeneous data.

Literature review. Financial crisis forecasting topics have been extensively addressed as applied to various businesses in both Ukrainian and foreign literature (Altman, 1968; Balcaen, 2006; Charitou et al., 2004; Korol, 2013; Matviychuk, 2010; Ohlson, 1980; Sun et al., 2014; Tereshchenko, 2004). However, fewer (and mainly Western) works have been dedicated to building crisis prediction models for insurance companies in particular (Brockett et al., 2006; Chiet et al., 2009; Kleffner and Lee, 2009; Lee and Urrutia, 2006; Sharpe and Stadnik, 2007). As for Ukrainian literature, there is an evident gap in research on the methods for financial crisis prediction in insurance companies: many authors either tend to rely on previously developed mathematical models without enough reasoning, or limit themselves to overviewing and systemizing them (Dobosh, 2009; Kryvytska, 2012), whereas only several Ukrainian scholars have made attempts to design original models for insurer bankruptcy prediction, financial stability evaluation etc. (Klepikova, 2011; Olkhovska, 2013; Shpitzhluz, 2013). At the same time, many new and promising methods still have not been applied to the problem of predicting financial crisis in Ukrainian insurance companies. One of such methods is the support vector machines (SVM).

Goal statement. The goal of the research is to develop a set of support vector machine models for financial crisis prediction in Ukrainian general insurance companies and choose the most acceptable support vector classifier based on model characteristics and performance measures.

Data and method. The data for this research was based on the 2010–2011 annual financial reports of Ukrainian general insurance companies gathered from various sources, including the official web-site of Stock market infrastructure development agency of Ukraine, the Public information database of the National securities and stock market commission of Ukraine, printed copies of the "Ukraine Business Review" journal, and corporate web-sites of the insurance companies. Information on the status of insurers was taken from the Complex information system of the National commission for regulation of financial services markets of Ukraine. The 2010 data (of 308 functioning and 31 financially distressed insurers) was used in model building (including the cross-validation process), whereas 2011 (of 314 functioning and 40 financially distressed insurers) data served for prediction power testing purpose.

Some important remarks have to be made. First, we assume that an insurer experienced financial crisis (or appeared "financially distressed") if it either filed for bankruptcy or ceased posting annual financial reporting data, which may be an indication of activity cessation. Secondly, the unavailability of information on the dates of insurance company exclusion from the register due to reorganization of State commission for regulation of financial services markets of Ukraine into National commission for regulation of financial services markets of Ukraine did not allow us create the exact list of companies excluded from the register in 2011. Hence, the crisis insurers in year 2011 are the companies that were excluded from the register in 2011–2012.

The initial set of financial indicators of insurance companies was gathered out of the reviewed principal works dedicated to bankruptcy prediction in insurance companies (Brockett et al., 2006; Chiet et al., 2009; Kleffner and Lee, 2009; Lee and Urrutia, 2006; Sharpe and Stadnik, 2007). During the modelling process, the indicators were purged of outliers (double-sided 10% crop), redundant indicators were excluded (based on mutual correlations), and the initial set was reduced to smaller subsets of variables with the use of two different feature selection algorithms. Table 1 contains the list and the formulae of indicators that appeared in one or more of the used feature sets.

Table 1. **Selected business indicators of general insurance companies,** *systematized by the author*

Variable name	Indicator (formula)
[C + CFI + R] / IR	*Liquid Assets / Insurance Reserves*
[C + CFI + R] / S	*Liquid Assets / Sales*[1)]
[C + CFI + R] / TA	*Liquid Assets / Total Assets*
[C + CFI + R] Gr	$(Liquid\ Assets_t - Liquid\ Assets_{t-1}) / Liquid\ Assets_{t-1}$
C / CL	*Cash / Current Liabilities*
CFO	*Cash from Operation Activities*
CL / TA	*Current Liabilities / Total Assets*
E / S	*Equity / Sales*
E / TL	*Equity / Total Liabilities*
GP / E	*Gross Premiums Written*[2)] */ Equity*
GP Gr	$(GPW_t - GPW_{t-1}) / GPW_{t-1}$
II / [C + CFI + LTFI]	*Investment Income / Average Invested Assets*
NI / S	*Net Income / Sales*
NI / TA	*Net Income / Total Assets*
QR	*Liquid Assets / Current Liabilities*
RE / TA	*Retained Earnings / Total Assets*
S / TA	*Sales / Total Assets*
TL / TA	*Total Liabilities / Total Assets*
WC / S	*Working Capital / Sales*

[1)] The ratios in which the numerator presents a state indicator and the denominator presents a flow indicator (and vice versa) imply that the state indicator is taken at its mean value to ensure unit consistency.
[2)] Gross Premiums Written may further be abbreviated as GPW.

The method applied in this research is the support vector machines, which is a relatively new artificial intelligence method based on the principle of structural risk minimization (in contrast to the methods that deal with empirical risk minimization). It has been developed by V. Vapnik (1995), C. Cortes and V. Vapnik (1995). Numerous applications of this method has shown its high effectiveness in classification and regression tasks (Wang et al., 2005: 821).

Application of support vector machines in corporate crisis prediction implies developing a classifier via finding the equation of a hyperplane that has the largest distance (margin) to the elements of each class. Most frequently, support vector machines are applied in linear classification tasks, but they can also be employed in greater dimension non-linear classification problems using kernel functions.

The general algorithm of building a support vector classifier is the following (Burges, 1998: 123–124; Cortes and Vapnik, 1995). The goal is to learn a mapping $X \mapsto Y$, where $x \in X$ is some object with attributes, $y \in Y$ is the class label. In the simplest two-class case, $x \in R^n, y \in \{\pm 1\}$. The desired classifier is $y = f(x,a)$, where a is the set of function parameters. The function is chosen from the set of hyperplanes in R^n space and can be presented as $f(x,\{w,b\}) = sign(w \times x + b)$, where w is the normal to the hyperplane, w x x is the dot product, and b is the shift of hyperplane relative to its normal.

The choice of $f(x,a)$ function is made via minimizing the overall risk (testing error) after imposing an upper bound to the overall risk. The bound includes a classifier complexity parameter and a free parameter which denotes the probability of risk being within the mentioned bound. The classifier building process implies balancing between margin maximization and error minimization. In the separable linear case, $y_i(w \times x_i + b) \geq 1$ must hold; otherwise (in the non-separable case), a soft margin $y_i(w \times x_i + b) \geq 1 - \xi_i$, with $\xi_i \geq 0$, is used to introduce a slack variable. The graphical interpretation of the described two-class linearly separable case is presented in Figure 1.

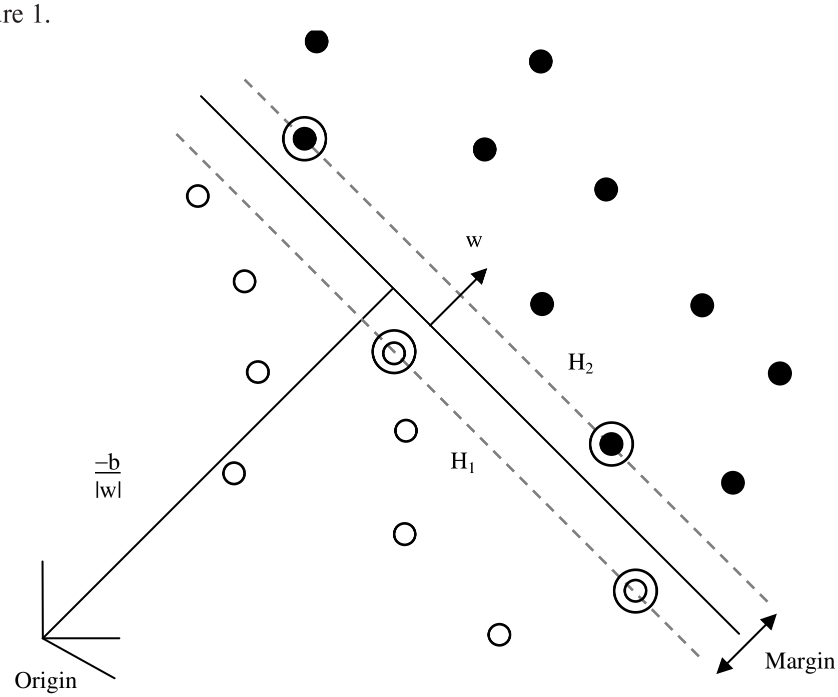

Figure 1. **Linear separating hyperplanes for the separable case; the support vectors are circled** (Cortes and Vapnik, 1995: 129)

As mentioned previously, support vector machines can be applied to non-linear cases by using kernel functions. The most commonly applied are polynomial $K(x,x') = (x^T x' + 1)^d$ and radial-basis $K(x,x') = e^{\gamma\|x - x'\|^2}$ non-linear kernel functions, where x and x' are different class samples, and d, γ are the parameters of the respective functions.

In foreign research, support vector machines have been widely applied to classification of companies depending on their financial health. For instance, Min and Lee (2005) were among the first to employ support vectors to predict corporate bankruptcy. The developed models showed better performance than multiple discriminant analysis and logistic regression. Shin et al. (2005) used support vector machines to predict failure of Korean enterprises and concluded better results; moreover, the models appeared to have outmatched artificial neural networks used for similar tasks.

Key research findings. Prior to building a support vector classification model, a set of relevant features had to be formed. With this aim, several methods could be applied. Firstly, recursive conditional correlation weighing was used on the year 2010 data in order to select variables most correlated to their class labels. This procedure was applied to both the initial sample and the purged sample to track possible changes in correlations. The results are presented in Table 2.

Table 2. **Top 10 features (recursive conditional correlation weighing, the year 2010),** *own calculations*

Sample			
Initial		Outliers excluded	
Indicator	Weight	Indicator	Weight
[C+CFI+R] Gr	1	[C+CFI+R] / TA	1
CFO / TA	1	CFO / S	1
CL / TA	1	E / S	1
E / TA	1	E / TL	1
GP / E	1	II / [C + CFI + LTFI]	1
CFO	1	ln (TA)	1
NI / TA	1	NS / TA	1
RE / TA	1	QR	1
TD / TA	1	TL / TA	1
PR	1	WC / S	1
Other	0	Other	0

It is obvious that the correlations between the features and the label indicator changed after outlier exclusion procedure because none of the top 10 relevant features matched. There was little confidence that sample purging did not discharge of any crucial information about the financial condition of companies. Therefore, the variable sets were formed based on both lists by excluding indicators that have mutual correlation higher or equal to |0.4|. Table 3 contains the first 4 sets of indicators used in model building.

Table 3. **Feature sets (recursive conditional correlation weighing, the 2010 year data),** *own calculations*

Feature set No.	Indicators
1	[C+CFI+R] / TA; E / S; II / [C+CFI+LTFI]; S / TA; QR; TL / TA
2	[C+CFI+R] / TA; E / TL; II / [C+CFI+LTFI]; S / TA; TL / TA; WC / S
3	[C+CFI+R] Gr; CL / TA; GP / E; CFO; NI / TA
4	[C+CFI+R] Gr; CL / TA; CFO; RE / TA

The application of support vector machines method was implemented using a RapidMiner 6.1 software package. Developing a support vector model consisted of choosing a kernel function and setting up its parameters; yet the theory of selecting optimal kernels and parameters is still limited. Thus, the selection of the best model was performed by iterating through 3 common kernel functions (linear, polynomial, and radial-basis) and the proposed ranges of parameter values.

Sensitivity was chosen as the main criterion of model appropriateness, while the overall accuracy of classification served as the second criterion. The reason for this lies in the characteristics of the training sample used in model building. The number of financially distressed insurance companies was nearly 10 times less than the number of financially healthy companies, which could induce classification bias towards the prevalent class. In such a case, the accuracy of the model might appear high even with a large number of incorrect classifications of financially distressed insurers. In our case, classifying a financially distressed company as a healthy one posed greater threat than classifying a healthy company as financially distressed. One of the possible ways to resolve this issue was to use class-specific precision measures such as specificity and sensitivity. Sensitivity, specificity, and accuracy formulas, as well as some other elements of the confusion matrix are presented in Table 4.

Table 4. **Confusion matrix,** *based on* Fawcett (2006)

Classification	Observation	
	0 (healthy)	1 (distressed)
0 (healthy)	*a*	*b* (type I error)
1 (distressed)	*c* (type II error)	*d*
Class-specific classification precision indicators	Specificity $a/(a+c)$	Sensitivity $d/(b+d)$
Overall classification precision indicator	Accuracy $(a+d)/(a+b+c+d)$	

Several support vector models were built using different kernel functions and parameter settings with application of 10-fold cross-validation. The cross-validation procedure implies random sampling of training data into 10 subsamples with each serving as the testing sample once, while others are used as a training sample. This allows for averaging the results and limits the overtraining risk. It is also worth noting that cross-validation in this case involves stratified sampling that accounts for class disproportions. In addition, built-in scaling of data and misclassification cost balancing were performed.

First, a linear support vector machine was built on feature set No. 1. Table 5 contains the matrix of model performance measures based on parameter C and ε values. Parameter C characterizes the width of the margin (balance between model complexity and error rate), while parameter ε denotes the insensitive zone (inversely related with the number of support vectors in the model).

The precision measures of the selected model and respective parameter values are shown in bold, and the maximum sensitivity level is additionally underlined. As it can be seen in Table 5, most models with high overall accuracy had rather low sensitivity and, thus, poorly classified the financially distressed insurers. The best performing linear model on feature set No. 1 showed the accuracy of almost 90%, yet relatively low sensitivity of nearly 54%, which could not be regarded as an acceptable outcome.

Table 5. **Modelling results for different *C* and ε parameter settings for feature set No.1 (linear kernel function),** *%, own calculations*

		Model precision measure	*C*				
			0.001	0.01	**0.1**	1	5
ε	0.001	Specificity	100.00	96.50	93.63	91.72	92.04
		Sensitivity	23.08	46.15	50.00	50.00	50.00
		Accuracy	94.12	92.65	90.29	88.53	88.82
	0.01	Specificity	100.00	96.50	93.63	91.40	91.72
		Sensitivity	23.08	46.15	50.00	50.00	50.00
		Accuracy	94.12	92.65	90.29	88.24	88.53
	0.1	Specificity	100.00	95.86	**92.04**	90.13	91.40
		Sensitivity	26.92	46.15	**53.85**	53.85	50.00
		Accuracy	94.41	92.06	**89.12**	87.35	88.24
	1	Specificity	100.00	100.00	100.00	100.00	100.00
		Sensitivity	0.00	0.00	0.00	0.00	0.00
		Accuracy	92.35	92.35	92.35	92.35	92.35

The logical step was to check non-linear kernels. Thus, based on feature set No. 1 polynomial kernel support vector models were built. In this case, besides the common *C* parameter, the polynomial degree d had to be screened. The ε parameter was kept constant at the value of 0.001. The modelling results are presented in Table 6.

Table 6. **Modelling results for different *C* and *d* parameter settings for feature set No. 1 (polynomial kernel function, ε = 0.001),** *%, own calculations*

		Model precision measure	*C*				
			0.001	0.01	0.1	**1**	5
d	0.5	Specificity	100.00	96.50	93.63	91.72	92.04
		Sensitivity	23.08	46.15	53.85	50.00	50.00
		Accuracy	94.12	92.65	90.59	88.53	88.82
	2	Specificity	100.00	94.90	92.68	87.90	84.71
		Sensitivity	30.77	38.46	57.69	69.23	69.23
		Accuracy	94.71	90.59	90.00	86.47	83.53
	3	Specificity	97.13	94.27	92.99	**88.54**	88.85
		Sensitivity	42.31	53.85	61.54	**80.77**	65.38
		Accuracy	92.94	91.18	90.59	**87.94**	87.06
	4	Specificity	96.82	95.54	96.50	97.13	96.18
		Sensitivity	42.31	57.69	57.69	69.23	65.38
		Accuracy	92.65	92.65	93.53	95.00	93.82
	5	Specificity	96.18	96.50	83.12	91.08	93.63
		Sensitivity	61.54	61.54	61.54	57.69	57.69
		Accuracy	93.53	93.82	81.47	88.53	90.88
	6	Specificity	95.86	95.22	98.73	95.54	86.31
		Sensitivity	50.00	53.85	50.00	57.69	61.54
		Accuracy	92.35	92.06	95.00	92.65	84.41

It is obvious that using a polynomial kernel function allowed us obtaining better results. The highest sensitivity reached was almost 81%, with the overall accuracy of the same model being almost 88%. At the same time, models with higher accuracy values showed lower sensitivity levels.

The next step was to build support vector models using a radial-basis kernel function. Here, the parameter ε was also kept constant, whereas *C* and γ parameters were alternated. Table 7 contains the modelling results.

Table 7. **Modelling results for different *C* and γ parameter settings for feature set No. 1 (radial-basis kernel function, ε = 0.001),** *%, own calculations*

		Model precision measure	*C*				
			0.001	0.01	0.1	**1**	5
γ	0.1	Specificity	100.00	100.00	94.90	89.17	90.13
		Sensitivity	0.00	0.00	57.69	73.08	61.54
		Accuracy	92.35	92.35	92.06	87.94	87.94
	0.2	Specificity	100.00	100.00	89.49	88.54	93.63
		Sensitivity	0.00	0.00	65.38	80.77	53.85
		Accuracy	92.35	92.35	87.65	87.94	90.59
	0.25	Specificity	100.00	100.00	89.49	**90.45**	93.63
		Sensitivity	0.00	0.00	65.38	**80.77**	61.54
		Accuracy	92.35	92.35	87.65	**89.71**	91.18
	0.3	Specificity	100.00	100.00	86.94	91.40	94.27
		Sensitivity	0.00	0.00	65.38	76.92	53.85
		Accuracy	92.35	92.35	85.29	90.29	91.18
	0.5	Specificity	100.00	100.00	82.80	92.36	94.90
		Sensitivity	0.00	0.00	80.77	57.69	53.85
		Accuracy	92.35	92.35	82.65	89.71	91.76
	0.75	Specificity	100.00	100.00	78.66	93.63	95.86
		Sensitivity	0.00	0.00	88.46	46.15	38.46
		Accuracy	92.35	92.35	79.41	90.00	91.47

As it can be observed, the use of radial-basis function yielded support vector models with better characteristics. While the best level of sensitivity was equal to the previous model, the overall accuracy increased by almost 2%. Assuming that radial-basis function had performed better, support vector models on the remaining feature sets were built using this kernel function. Table 8 contains the precision measures and respective parameter settings of the best models for all feature sets.

As it can be observed from the table, the most reasonable model based on the performance measures was the one on feature set No. 2: despite the fact that sensitivity level was relatively low in comparison to the models on feature sets No. 3 and No. 4, the overall accuracy levels of the latter appeared to be unacceptable.

In order to conduct the testing of predicting characteristics of the classifiers, the built support vector models were applied to the subsequent year data. The results are shown in Table 9.

The obtained results denote that the developed support vector models demonstrated weak predictive qualities. In spite of the relatively high level of accuracy of the model on feature set No. 2, the sensitivity levels were rather low. At the same time, model built on feature set No. 3 appeared to have a high sensitivity value, yet lacked overall accuracy.

Table 8. **Modelling results for different *C* and γ parameter settings for feature sets No. 1–4 (radial-basis kernel function, ε = 0.001),** *own calculations*

Model characteristics		Feature set			
		No. 1	No. 2	No. 3	No. 4
Precision measure	Specificity, %	90.45	92.04	58.92	71.02
	Sensitivity, %	80.77	80.77	88.46	92.31
	Accuracy, %	89.71	91.18	61.18	72.65
Parameters	C	1	5	0.1	1
	γ	0.25	0.2	0.1	0.1

Table 9. **Precision measures for selected models based on subsequent year data,** *%, own calculations*

Feature set, kernel function, model parameters	Specificity	Sensitivity	Accuracy
Feature set No. 1, radial-basis function, γ = 0.25, C = 1	70.70	32.50	66.38
Feature set No. 2 radial-basis function, γ = 0.2, C = 5	77.07	35.00	72.32
Feature set No. 3, radial-basis function, γ = 0.1, C = 0.1	43.31	70.00	46.33
Feature set No. 4, radial-basis function, γ = 0.1, C = 1	58.28	55.00	57.91

One of the possible causes of such an outcome may be the change in the conditions of insurance business functioning, which, in turn, affected the relationships between financial indicators in the following year. In order to address the abovementioned issue, a more complex modelling approach was followed.

The alternative feature selection algorithm, which is proposed, involves using the minimum redundancy – maximum relevance feature ranking method combined with accounting for the change in feature ranks between the periods. At the same time, purged samples were used due to better performance of the built support vector models on the data cleared of outliers. Table 10 contains the top 10 relevant and least redundant features for both years; bold font denotes matches between the years.

Table 10. **Top 10 features (minimum redundancy – maximum relevance method, feature matches presented in bold),** *own calculations*

Year			
2010		2011	
Indicator	Weight	Indicator	Weight
[C+CFI+R] / S	1	**[C+CFI+R] / S**	1
C / CL	1	**C / CL**	1
CA / S	1	CFO / TL	1
E Gr	1	EBIT / TA	1
GP Gr	1	E / TL	1
[C+CFI+R] / NIR	1	**GP Gr**	1
NI / S	1	**[C+CFI+R] / NIR**	1
TD / TA	1	**NI / S**	1
WC / S	1	QR	1
PR	1	**WC / S**	1
Other	0	Other	0

Apparently, the selected features differed between the years 2010–2011. Nevertheless, matching features were present and were applied in further modelling.

As in the previous case, only uncorrelated indicators were used to build support vector classifiers. The correlation matrix was used to spot highly correlated pairs, after which several possible uncorrelated subsets were formed. The selected feature sets are presented in Table 11.

Table 11. **Feature sets (minimum redundancy – maximum relevance method, both years data, numbering continued),** *own calculations*

Feature set No.	Indicators
5	[C+CFI+R] / S; C / CL; GP Gr; [C+CFI+R] / IR
6	C / CL; GP Gr; [C+CFI+R] / IR; NI / S
7	C / CL; GP Gr; [C+CFI+R] / IR; WC / S

Thereby, 3 additional sets of indicators were created to be used in support of vector model developing. Since it had been previously concluded that radial-basis kernel function yielded better results, only this kernel function was applied for new feature sets. Table 12 contains the precision measures and parameter settings for built models.

Table 12. **Modelling results for different** *C* **and** γ **parameter settings for feature sets No. 5–7 (radial-basis kernel function,** ε **= 0.001),** *own calculations*

Model characteristics		Feature set		
		No. 5	No. 6	No. 7
Precision measure	Specificity, %	90.13	87.58	89.17
	Sensitivity, %	96.15	96.15	96.15
	Accuracy, %	90.59	88.24	89.71
Parameters	C	1	1	0.1
	γ	1.25	0.75	0.75

It can be observed that all 3 models had high performance measures: sensitivity was above 95%, whereas the overall accuracy was around 90%. It can be concluded that the second feature selection algorithm facilitated the model quality improvement. However, the developed support vector classifier had to be also tested on the subsequent year data. The results are demonstrated in Table 13.

Table 13. **Precision measures for the selected models based on subsequent year data,** *%, own calculations*

Feature set, kernel function, model parameters	Specificity	Sensitivity	Accuracy
Feature set No. 5, radial-basis function, $\gamma = 1.25$, $C = 1$	89.49	57.50	85.88
Feature set No. 6 radial-basis function, $\gamma = 0.75$, $C = 1$	87.26	57.50	83.90
Feature set No. 7, radial-basis function, $\gamma = 0.75$, $C = 0.1$	88.22	60.00	85.03

Although sensitivity levels were relatively lower in the models on feature sets No. 4–6 in comparison to those of the support vector classifiers built based on the first 4 indicator sets, the overall accuracy values appeared to be considerably higher. Moreover, the greatest gain could be seen in the sensitivity values. Taking into account the presented precision measures, support vector classifier based on the feature set No. 7 could be regarded as the best preforming and most acceptable among the developed models. This feature set includes the following measures of insurance companies' financial health:

- liquidity (cash to current liabilities ratio);
- revenue dynamics (growth of gross premiums written);
- inverse working capital turnover (working capital to sales ratio);
- insurance reserve coverage (liquid assets to (net) insurance reserves ratio).

Conclusions and implications for further research. It can be concluded that the application of support vector machines to financial crisis prediction in Ukrainian insurance companies appears to be sufficiently effective. The choice of features that can be used to discriminate between financially distressed and financially healthy insurance companies should be made considering the change of relevant indicators between periods. Among the 3 studied kernel functions of support vector machines, radial-basis kernel function allowed for best performing models with sensitivity and accuracy levels close to 90% on the testing data within the cross-validation procedure. The models testing revealed that the predicting power of the developed support vector classifiers are lower, especially in terms of sensitivity. Nevertheless, high overall accuracy level allowed concluding that the model built on the feature set No. 7 (with parameter C value of 0.1 and γ value of 0.75) can be applied for financial crisis prediction purposes in Ukrainian insurance companies.

There are several possible directions of improving the effectiveness of support vector machines application for financial crisis prediction. For instance, developing method ensembles (both purely support vector ensembles and different classification method ensembles) can greatly facilitate the increase in prediction accuracy. Additionally, continuous relevant database expansion should alleviate the problems related to scarce samples and class disproportions, which will also positively affect model precision measures.

The research results can be applied in Ukrainian insurance companies in order to objectify and improve the effectiveness of financial crisis management processes.

References:

Добош Н. Оцінка фінансової стійкості страховика // Формування ринкової економіки в Україні.– 2009.– №19. – С. 207–212.

Клепікова О. Прогнозування головних показників страхової компанії методом системної динаміки // Праці Одеського політехнічного університету.– 2011.– №37. – С. 74–79.

Кривицька О. Систематизація підходів до сутності фінансової стійкості страхової компанії // Наукові записки Національного університету «Острозька академія».– Серія: Економіка.– 2012.– №20. – С. 129–132.

Матвійчук А. Моделювання фінансової стійкості підприємств із застосуванням теорії нечіткої логіки, нейронних мереж і дискримінантного аналізу // Вісник НАН України.– 2010.– №9. – С. 24–29.

Охловська О. Моделювання фінансового стану страхової компанії із застосуванням апарату нечіткої логіки // Нейро-нечіткі технології моделювання в економіці.– 2013.– №2. – С. 119–134.

Терещенко О. Антикризове фінансове управління на підприємстві: Монографія. – К.: КНЕУ, 2004. – 268 с.

Шпіцглуз С. Моделювання раннього попередження нестабільності страхової компанії // Економіко-математичне моделювання та інформаційні технології.– 2013.– №1. – С. 296–301.

Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The Journal of Finance, 23(4): 589–609.

Balcaen, S. (2006). 35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems. The British Accounting Review, 38: 63–93.

Brockett, P., Golden, L., Jang, J., Yang, Ch. (2006). A comparison of neural network, statistical methods, and variable choice for life insurers' financial distress prediction. The Journal of Risk and Insurance, 73(3): 397–419.

Burges, C. (1998). A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery, 2: 121–167.

Charitou, A., Neophytou, E., Charalambous, C. (2004). Predicting corporate failure: empirical evidence for the UK. European Accounting Review, 13(3): 465–497.

Chiet, N., Jaaman, S., Ismail, N., Shamsuddin, S. (2009). Insolvency prediction model using artificial neural network for Malaysian general insurers. Proceedings of IEEE World congress on Nature and Biologically Inspired Computing (NaBIC) (pp. 584–589), Coimbatore, India.

Cortes, C., Vapnik, V. (1995). Support-Vector Networks. Machine Learning, 20: 273–297.

Fawcett, T. (2006). An introduction to ROC analysis. Pattern recognition letters, 27(8): 861–874.

Kleffner, A., Lee, R. (2009). An examination of property & casualty insurer solvency in Canada. Journal of Insurance Issues, 32(1): 52–77.

Korol, T. (2013). Early warning models against bankruptcy risk for Central European and Latin American enterprises. Economic Modelling, 31: 22–30.

Lee, S., Urrutia, J. (2006). Analysis and prediction of insolvency in the property-liability insurance industry: a comparison of logit and hazard models. The Journal of Risk and Insurance, 63(1): 121–130.

Min, J., Lee, Y.-C. (2005). Bankruptcy prediction using support vector machines with optimal choice of kernel function parameters. Expert systems with applications, 28: 603–614.

Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research, 18(1): 109–131.

Sharpe, I., Stadnik, A. (2007). Financial distress in Australian general insurers. The Journal of Risk and Insurance, 74(2): 377–399.

Shin, K.-S., Lee, T., Kim, H. (2005). An application of support vector machines in bankruptcy prediction model. Expert Systems with Applications, 28: 127–135.

Sun, J., Li, H., Huang, Q.-H., He, K.-Y. (2014). Predicting financial distress and corporate failure: a review from the state-of-the-art definitions, modeling, sampling, and featuring approaches. Knowledge-Based Systems, 57: 41–56.

Vapnik, V. (1995). The Nature of Statistical Learning Theory. Springer, New York.

Wang, Y., Wang, S., Lai, K. (2005). A new fuzzy support vector machine to evaluate credit risk. IEEE Transactions on fuzzy systems, 13(6): 820–831.

Стаття надійшла до редакції 17.02.2015.