Sabina Bakun, A. Lytvyniuk, Oleksandr Terentiev
(Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine)

USING OF CREDIT SCORING TO DECIDING ON THE LOAN

Introduction. Today, many banks faced with the problem of default issued by companies or individuals loans. Therefore, applying methods of evaluation solvency of individuals and risk of banks in lending in modern crisis of the financial sector by banking institutions is really necessary.

Scoring systems. Effective method that optimally solves problems of evaluating the solvency of individuals is credit scoring, it is a mathematical or statistical model that banks use to define the probability of returning the loan by potential client in the established term.

Can identify the following main stages of credit scoring:

1. Preparation of initial data to build scoring models. Include: collection of statistical data on bank customers, target definition, the definition of independent variables such as socio-demographic data of the client, information about the requested loan and so on.

2. Data processing. At this stage take place operations with the data, which may be: data filtering, substitution of missing values, transformation of parameters.

3. Definition of the most significant variables to be included in the model, using the Gini coefficient or Information Value.

4. Building of scoring models. In one project is possible to use different statistical methods for building scoring models: logistic regression, classification trees, neural networks, scoring cards.

5. Checking of the quality of the resulting model. To evaluate the quality of the model and its predictive power are used standard statistical coefficients, such as statistic of Kolmogorov-Smirnov, the area under the ROC-curve, coefficient Gini.

6. Analysis, determination of the optimal cut-off points.

7. Integration of models in the information bank infrastructure.

8. Retraining model. During using the scoring cards and accumulating new data on the loan portfolio data can be retraining scoring models.

Incoming data. The database consists of two data sets. The first data set contains information about 5837 clients who were issued loans, and who has already ended period of loans, the second data set contain 4233 records of applicants who have been denied loan. Each entry in the two datasets contains 18 attributes that characterize the customers, moreover the first data set contains additional indicator (target variable) which reflects the result of the return credit of customer.

Problem. Need to prepare the sample data to build the scoring card, select the most important characteristics of the clients to be included in the model, develop a scoring card and evaluate the quality of developed model.

The method of solution. SAS Enterprise Miner includes specialized components for solving the problem of scoring. To build a scoring model that is shown in fig. 1 was applied components such as:

1) Data Partition - allows you to split our sample into training and testing;

2) Interactive Grouping - provides a selection of the most important characteristics and forms groups of values for the continuous input variables;

3) Reject Inference - allows to add information on applicants who has been denied loan with automatic markup customers in positive and negative to a training sample data;

4) Scorecard - automatically builds scoring card on the results of the regression model, based on data from the training set.

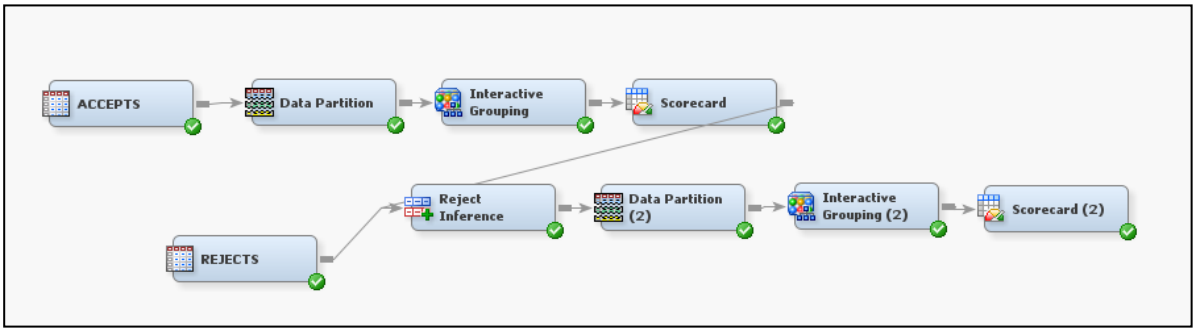

Fig. 1 Scoring model built by SAS Enterprise Miner

The results of system operation are displayed as built scoring card (fig. 2), which consists of parameters (attributes), ranges of values of each attribute and scoring mark for each of the ranges. To evaluate solvency of new applicant enough summarize marks for each indicator of scoring card. It also provides a number of reports with statistical indicators of quality built scoring card (fig. 3).

Scorecard

		Scorecard Points
age_oldest_tr	age_oldest_tr< 38	12
	38<= age_oldest_tr< 77	15
	77<= age_oldest_tr< 186	18
	186<= age_oldest_tr	25
	MISSING	12

Fig. 2 The fragment of scoring card built by SAS Enterprise Miner

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
bad		_AIC_	Информационный крите...	6579.409	.	
bad		_ASE_	Средний квадрат ошибки	0.141709	.	0.144467
bad		_AVERR_	Функция средней ошибки	0.438329	.	0.446885
bad		_DFE_	Число степеней свободы ...	7478.856	.	
bad		_DFM_	Число степеней свободы ...	8	.	
bad		_DFT_	Общее число степеней св...	7486.856	.	
bad		_DIV_	Делитель для асимптоти...	14973.71	.	1703.431
bad		_ERR_	Функция ошибки	6563.409	.	761.2373
bad		_FPE_	Итоговая ошибка прогноза	0.142012	.	
bad		_MAX_	Максимальная абс. ошиб...	0.9936	.	0.992683
bad		_MSE_	Средний квадрат ошибки	0.141861	.	0.144467
bad		_NOBS_	Сумма частот	7486.856	.	851.7156
bad		_NW_	Кол-во оценочных весов	8	.	
bad		_RASE_	Квадратный корень из су...	0.376443	.	0.380088
bad		_RFPE_	Корень из итоговой ошиб...	0.376845	.	
bad		_RMSE_	Стандартная ошибка	0.376644	.	0.380088
bad		_SBC_	Байесовский критерий Ш...	6634.777	.	
bad		_SSE_	Сумма квадратов ошибок	2121.915	.	246.0891
bad		_SUMW_	Сумма временных частот...	14973.71	.	1703.431
bad		_MISC_	Коэффициент неправиль...	0.200581	.	0.200697
bad		_AUR_	Area Under ROC	0.755207	.	0.743578
bad		_Gini_	Gini Coefficient	0.510414	.	0.487156
bad		_KS_	Kolmogorov-Smirnov Stati...	0.384649	.	0.383433

Fig. 3 Statistical estimates of quality developed model

Conclusions. The risk of credit operations is one of the most important risks that may effect on the functioning of financial institutions. Credit operations are the most profitable type of bank's asset, but at the same time it is the most risky. Implementation of the scoring cards allow banking institutions to examine more comprehensive not only the financial capacity of the solvency of client, but also take into account other socio-demographic characteristics such as age, sex, education etc., which can negatively affect for ability of the client repaying the loan. Applying scoring evaluation will also reduce the cost and time that bank has spent for processing statements and deciding on the loan.

References

1. Anderson, Billie S., and R. Wayne Thompson. Developing Credit Scorecards Using SAS Credit Scoring for Enterprise Miner 5.3. – Cary: SAS Institute Inc, 2009. – 41 p.

2. Siddiqi N. Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring – Cary: SAS Institute Inc, 2005. – 208 p.

3. Терентьев А.Н. SAS BASE: Основы программирования (научное издание) / Терентьев А.Н., Домрачев В.Н., Костецкий Р.И. – К: Эдельвейс, 2014. – 304 с. – ISBN 978-966-2748-49-9