# AN GLM MODEL FOR PREDICTION OF CRISIS IN SLOVAK COMPANIES

**Peter Adamko**[77]
**Tomáš Klieštik**[78]
**Mária Kováčová**[79]
DOI: https://doi.org/10.31410/EMAN.2018.223

**Abstract:** *The goal of this paper is to create and evaluate a GLM model that predict whether a Slovak company will find itself in a crisis in the following year. Paper also provide an overview of metrics used to evaluate the quality or effectiveness of a model.*
*The data used to create and test the model are from years 2014/2015 and 2015/2016, respectively. The data was obtained from amadeus - a database of comparable financial information for public and private companies across Europe (www.bvdinfo.com). For each year we had data about 100 000 Slovak companies. To test the performance of the model we used a standard metrics (AUC, Sensitivity, Confusion Matrix, RMSE, ...).*

**Key words:** *Default, prediction model, company in a crisis*

## 1. INTRODUCTION

In 2016 in Slovakia the new provisions of Law No. 513/1991 Coll. Commercial Code on companies in crisis has entered into force. The company is in a crisis when it is in default or at risk of imminent default.

The default of a company is defined in law no. 7/2005 Coll. On Bankruptcy and Restructuring as amended. Under the Act a company is in default if it has liabilities to at least two entities and the value of its liabilities exceeds the value of its assets or if it is unable to pay at least two financial liabilities to at least two creditors 30 days after the due date.

A company is at risk of imminent default when has a low ratio of equity and liabilities. For 2016, a company is at risk of imminent default when the ratio is less than 4 to 100. For 2017 and 2018 ratios are 6 to 10 and 8 to 100 respectively.

If a company is in a crisis due to a negative equity and has at least two creditors, then it is obliged to file for bankruptcy within 30 days. If a company is in crisis due to insolvency and it has asked (written form) to pay its obligation, then creditors may file a bankruptcy petition.

In the Slovak Republic, individually submitted claims are entered into the list of claims, while each registered claim is compared with the list of liabilities, the accounting and other documentation of the bankrupt subject. The claims that are disputed are subsequently denied within a 30-day period. At the same time, the law adjusts the objective responsibility of actively

[77] University of Zilina, Faculty of Operation and Economics of Transport and Communications, Department of Quantitative Methods and Economic Informatics, Univerzitna 1, 010 26 Zilina, Slovak Republic
[78] University of Zilina, Faculty of Operation and Economics of Transport and Communications, Department of Economics, Univerzitna 1, 010 26 Zilina, Slovak Republic
[79] University of Zilina, Faculty of Operation and Economics of Transport and Communications, Department of Economics, Univerzitna 1, 010 26 Zilina, Slovak Republic

authorised persons (i.e., the creditor and the administrator) for the damage of the creditor caused by the negation of his claim [2] - [4].

The goal of this paper is to create and evaluate a generalized linear model (GLM) that predict whether a Slovak company will find itself in a crisis in the following year. The family of GLM models are some of the most commonly-used models for many types of data analysis use cases. GLM estimate regression models for outcomes following exponential distributions in general. Along with the Gaussian distribution, these include Poisson, binomial, gamma and Tweedie distributions. Each serves a different purpose and depending on distribution and link function choice.

***Peter Adamko***

*The main scope of Mgr. Bc. Peter Adamko, PhD. pedagogic activities is informatics. Research activities are focused mainly on the mathematics of quantitative methods in financial management and corporate governance. He is involved in many research tasks relating to that issue, and their outputs have been published in many domestic and foreign publications. He was a member of the research team of four VEGA and one APVV projects. He is the author or co-author of one university textbook, two technical publications, 4 lecture notes and 58 articles published in domestic and international journals, conference proceedings of conferences. 15 of them are indexed in WoS or Scopus. His publications have 59 citation responses in WoS or Scopus, his h-index is 4.*

## 2. DATA

The data used to create and test the model are from years 2014, 2015 and 2016. The data was obtained from amadeus - a database of comparable financial information for public and private companies across Europe (www.bvdinfo.com). To create the model, we used data from 2014 and 2015 to calculate predictors and to calculate the response variable, respectively. These data were divided into training data (77 591 observations) and validation data (25 752 obs.). To test the model, we used data from 2015 (predictors) and 2016 (response) (104 842 obs.).

Summary statistics of training set and test set are shown in Table 1 and Table 2, respectively.

| | ***X1*** | ***X2*** | ***X3*** | ***X4*** | ***Y*** |
|---|---|---|---|---|---|
| SD | 176.89 | 50.18 | 439.45 | 226.45 | 0.43 |
| VAR | 31 289.96 | 2 517.78 | 193 117 | 51279.33 | 0.18 |
| Zeros | 6748 | 2 944 | 0 | 44 | 59 093 |
| Min | -35 810.25 | -9 967.3 | -6 708 | -6 212.95 | 0 |
| 1stQ | -0.01 | -0.02 | 0.27 | 0.02 | 0 |
| Mean | -1.64 | -0.62 | 9.13 | 9.15 | 0.24 |
| Median | 0.07 | 0.05 | 0.67 | 0.47 | 0 |
| 3rdQ | 0.29 | 0.18 | 0.98 | 2.61 | 0 |
| Max | 799.86 | 2 733.17 | 64 939.25 | 39 836.2 | 1 |

Table 1: Training data – Summary statistics

| | X1 | X2 | X3 | X4 | Y |
|---|---|---|---|---|---|
| SD | 123.27 | 94.22 | 1 486.43 | 1 158.90 | 0.42 |
| VAR | 15 194.43 | 8 878.03 | 2 209 465 | 1 343 052 | 0.18 |
| Zeros | 9023 | 3 983 | 0 | 44 | 80 606 |
| Min | -21 396.86 | -17 245 | -5 549 | -5 461.79 | 0 |
| 1stQ | -0.01 | -0.01 | 0.27 | 0.02 | 0 |
| Mean | -1.33 | -0.54 | 13.60 | 12.55 | 0.23 |
| Median | 0.07 | 0.05 | 0.67 | 0.48 | 0 |
| 3rdQ | 0.29 | 0.19 | 0.98 | 02.60 | 0 |
| Max | 919.51 | 10 855 | 462 230.5 | 339 035.1 | 1 |

Table 2: Test data – Summary statistics

For our model, we chose the following four predictors:
X1 = Working capital / Total assets
X2 = Operating P/L (=EBIT) / Total assets
X3 = (Non-current liabilities + Current liabilities) / Total assets
X4 = Solvency ratio
Y = 1 - company will find itself in a crisis in the following year. 0 - otherwise.

To identify company in a crisis we checked the following conditions:

1. Equity/Liabilities < 0,4
2. Quick ratio < 1
3. Earnings after Taxes < 0
4. Total assets < Non-current liabilities + Current liabilities)

If an enterprise meets first three conditions or fourth condition, it is classified as "in crisis". These criteria treat potential/real indebtedness of a company. (Due to accumulation of losses from previous years, which would indicate an inability to generate profit in the longer term, insolvency and the current inability to make a profit) [5] - [6].

It should be emphasized that debt ratio of equity and liabilities varies from industry to industry [1]. A ratio that could be a standard for an industry may be unmanageable for another one.

## 3. MODEL

We created a logistic regression model to predict whether a company will find itself in the crisis following year (under the conditions listed above). Logistic regression is frequently used because of its less restrictive assumptions. Mainly: predictors don't have to be normally distributed, or have equal variance in each group, there is no

***Tomáš Klieštik***

*Scientific and research activities of prof. Ing. Tomáš Klieštik, PhD. are mainly focused on the area of application of quantitative mathematical and statistical methods in financial management and decision making, Data envelopment analysis, neural networks, genetic algorithms, fuzzy logic, risk quantification and diversification, etc. He is or was a leader of a research team of three VEGA projects and a research team member of other seven VEGA projects. He is an author or a co-author of three monographs, 5 university textbooks, 16 articles in national scientific journals, 23 articles in international scientific journals, 38 contributions published in the proceedings of national scientific conferences, 49 contributions published in the proceedings of international scientific conferences. The research work of prof. Ing. Tomáš Klieštik, PhD. has more than 450 responses. 214 out of them are indexed in the database of Web of Science and 9 of them in Scopus. 24 publications of prof. Ing. Tomáš Klieštik, PhD. are indexed in the database of Web of Science and his H-index is 6.*

homogeneity of variance assumption, normally distributed error terms are not assumed. The model was created and tested in R with H2O package.

Model coefficients are shown in Table 3. From last column we can see that X3 and X4 have main impact on the result.

| ***Names*** | ***Coefficients*** | ***Standardized coefficients*** |
|---|---|---|
| Intercept | -1.1494 | -1.1766 |
| X1 | 0.0027 | 0.4838 |
| X2 | -0.0036 | -0.1828 |
| X3 | 0.0034 | 1.4733 |
| X4 | -0.0061 | -1.3745 |

Table 3: Model coefficients

| | ***Validation data*** | ***Test data*** |
|---|---|---|
| MSE | 0.180508 | 0.174054 |
| RMSE | 0.424862 | 0.417198 |
| r2 | 0.019489 | 0.020677 |
| logloss | 0.546524 | 0.532060 |
| AUC | 0.914996 | 0.920230 |
| Gini | 0.829993 | 0.840459 |
| mean_per_class_error | 0.153339 | 0.147593 |

Table 4: Model performance metrics

Table 4 shows that our model yields very good results, for example AUC is 0.92.

Table 5 shows that all balanced metrics are maximised when the threshold is approximately the same = 0.241.

***Ing. Mária Mišanková, PhD.*** *graduated at the Faculty of Operation and Economics of Transport and Communications of the University of Zilina in the field of study Corporate Economics and Management in 2012, where subsequently in 2015 in the same field of study she completed her PhD studies with a dissertation thesis: "Implementation of the strategy into the business plan". Appointed currently works at the Department of Economics of the Faculty of Operation and Economics of Transport and Communications of the University of Zilina as a researcher and specializes on financial markets, credit risk and international business. In these areas she also publishes in domestic and foreign scientific journals, on international domestic and foreign conferences, whose outputs are involved in internationally respected databases (nowadays 16 publications and 14 citations in the database Web of Science and 1 publication in database SCOPUS). So far she was an investigator on project VEGA 1/0656/14 "Research of Possibilities of Credit Default Models Application in Conditions of the SR as a Tool for Objective Quantification of Businesses Credit Risks".*

| Metrics | Validation data | | Test data | |
|---|---|---|---|---|
| | Threshold | Value | Threshold | Value |
| max f1 | 0.2413 | 0.7587 | 0.2413 | 0.7574 |
| max f2 | 0.2410 | 0.8092 | 0.2410 | 0.8079 |
| max f0point5 | 0.2414 | 0.7715 | 0.2416 | 0.7754 |
| max accuracy | 0.2414 | 0.8833 | 0.2414 | 0.8885 |
| max precision | 0.8822 | 0.8929 | 0.9999 | 0.9322 |
| max recall | 0.0035 | 1.0 | 0.0001 | 1.0 |
| max specificity | 0.9999 | 0.9999 | 0.9999 | 1.0 |
| max absolute_mcc | 0.2413 | 0.6803 | 0.2413 | 0.6811 |
| max min_per_class_accuracy | 0.2411 | 0.8543 | 0.2412 | 0.8499 |
| max mean_per_class_accuracy | 0.2411 | 0.8582 | 0.2412 | 0.8582 |

Table 5: Maximum metrics at their respective thresholds

## 3. CONCLUSION

We created the logistic regression model to predict whether a Slovak company will find itself in the crisis following year. We decided for logistic regression because it is easy to interpret and because of its not very restrictive assumptions. Also, it is also easy to use, unlike models such as Gradient boost model, Random forest, Deep learning model and so on.

To determine the result, simply put X1, X2, X3, X4 into Formula 1 and place the result ($t$) in Formula 2. Then compare the result with the threshold (0.241, based on table 5). If $c$ is bigger than the threshold, the company will find itself in crisis next year.

$$t = -1.1494 + 0.0027 \cdot X1 - 0.0036 \cdot X2 + 0.0034 \cdot X3 - 0.0061 \cdot X4 \quad (1)$$

$$c = \frac{1}{1+e^{-t}} \quad (2)$$

Figure 1 and Table 6 shows that our model yields very good results. The table shows specific results when threshold is 0.2416.

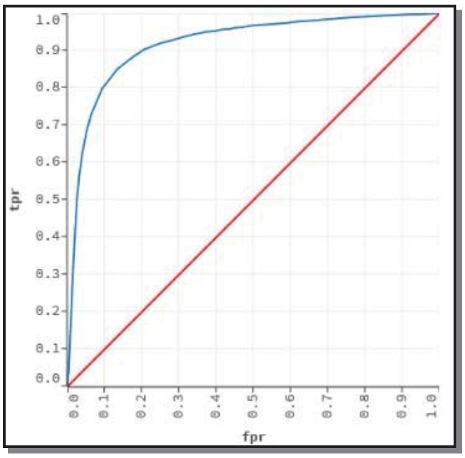


Figure 1: ROC - test data

| ***Actual \ Predicted*** | ***0*** | ***1*** | ***Error*** |
|---|---|---|---|
| ***0*** | 77 396 | 3 210 | 0.0398 |
| ***1*** | 9 091 | 15 145 | 0.3751 |
| ***Total*** | 86 487 | 18 355 | 0.1173 |

Table 6: Confusion matrix, test data, threshold = 0.2416

**Acknowledgment**
This research was financially supported by the Slovak Research and Development Agency – Grant NO. APVV-14-0841: Comprehensive Prediction Model of the Financial Health of Slovak Companies.

## REFERENCES

[1] Cisko, S., Kliestik, T. (2013). *Financny manazment podniku II*, EDIS Publishers, Zilina, Slovakia.
[2] Kliestikova, J., Misankova, M., Kliestik. T. (2017). Bankruptcy in Slovakia: international comparison of the creditor´s position. *Oeconomia Copernicana*, 8(2), pp. 221-237.
[3] Kliestikova, J., Misankova, M., Kliestik. T. (2016). *European insolvency law harmonisation in terms of global challenges*. Globalization and its socio-economic consequences, 16th international scientific conference proceedings, pts i-v. p. 914-921.
[4] Majerova, J., Kliestik T., Misankova, M. (2015). *Analysis of Convergences and Divergences of Bankruptcy Law Institutes in the Slovak Republic and in the World*. 5th international conference on applied social science (ICASS 2015), PT 1 Book Series: Advances in Education Research, Volume: 80.
[5] Valaskova, K. and Zvarikova, K. (2014). *Prediction of Credit Risk by Selected Theoretical Rating Models*. 2nd international conference on social sciences research (SSR 2014), PT 1 Book Series: Advances in Social and Behavioral Sciences Volume: 5 pp. 160-166.
[6] Svabova, L., Kral, P. (2016). S*election of predictors in bankruptcy prediction models for Slovak companies*. Prague: 10th international days of statistics and economics. pp. 1759-1768.