# A FEW NOTES ON DEPLOYMENT OF SUPERVISED CORPORATE FINANCIAL DISTRESS PREDICTION MODELS IN SMALL ENTERPRISES

**IGOR KOLLÁR, PAVOL KRÁĽ, PETER LACO**
Matej Bel University in Banská Bystrica, Faculty of Economics,
Department of Quantitative Methods and Information Systems,
Tajovského 10, Banská Bystrica, Slovakia
e-mail: igor.kollar@umb.sk, pavol.kral@umb.sk, peter.laco@umb.sk

## Abstract

*Prediction of corporate financial distress is often based on static classification models constructed using various supervised statistical methods, e.g. discriminant analysis, logistic regression, decision trees. Regardless of the selected classification method, a carefully checked data set containing information about the occurrence of financial distress and quantitative characteristics, most typically three or four years prior to the time of financial distress, of selected companies has to be available for a company in order to utilize this approach. Moreover, employees possessing at least moderate data analytic skills are vital to fit, interpret, validate, deploy and, eventually, upgrade models. These requirements make it quite infeasible for small businesses to construct such models on their own and utilization of freely available models from other sources is possibly preferred. In the paper we propose deployment of such models using Shiny, a web application framework for R. The deployed model is based on data obtained from CRIF – Slovak Credit Bureau, s.r.o. and we focus on simple automated interpretation of model outputs for a particular company.*

***Key words:*** *financial distress, prediction, application, shiny, small enterprises.*

## 1. Introduction

The financial health of the company is one of the crucial factors that defines whether company will survive in today's rapidly evolving and volatile world. As stated by Lesáková (2008, p. 607) "Small and medium enterprises (SME) are faced with the need for a strategic response to the changes in global business environment." Same author also stressed that impact of globalization on businesses sector is indisputable (see Lesáková, 2008, p. 607). "The future success of small and medium enterprises in the new world of global economics will be determined by

a) implementation of new businesses types in SMEs sector,
b) implementation of innovation, information and communication technologies by SMEs,
c) implementation of strategic management by SMEs."

In our opinion, the point b) in previous arguments encompasses also implementation of innovative theoretical ideas risen by academicians and the corresponding research outputs, e.g. prediction models, into simple, but still powerful, managerial decision support systems utilizing various data sources and up to date software solutions.

In our paper, focusing on Slovak economic environment, we further discuss this point illustrating it in the case of supervised corporate financial distress (bankruptcy) prediction models represented for example by (in)famous Altman's bankruptcy formula (Altman, 1968), well known as Altman's Z-score, and its later modifications and extensions (Altman et al.,

1977; Altman 1983). Although, as was demonstrated, for example, by Boďa and Úradníček (2016a, p. 70), innovative approaches can still emerge in corporate financial analysis, even in the case of Slovak economic environment, their inclusion into decision making processes in SMEs can be problematic in practice. The well-known and long existing financial distress and/or bankruptcy models still fully dominate their recent modifications or new counterparts, even though decision makers in SMEs are aware or should be aware of the possibility that assumptions used for fitting the original models are often not valid anymore due to changes in economic environments, law frameworks, incomparability of populations of interest etc.

Moreover, as pointed out by Bieliková et al. (2014, p. 38), due to the dynamics and uncertainty of the current economics, the number of studies which focus on financial situation diagnostics by using various statistical and data mining methods increases. However, without paying sufficient attention to accurate definition of risky company state. As a result, recently proposed modifications or new models can be only partially compatible or even worse, completely incompatible with the original ones.

In the case of SMEs, there are also many other factors influencing negatively adoption of corporate financial distress (bankruptcy) prediction models including employees lacking adequate data analytic skills, unavailable data and non-existent free software solutions allowing simple application of fitted models. These factors can make it quite infeasible for SMEs to construct and interpret such models on their own.

The presented obstacles are further amplified if SMEs decide to apply some current models presented in various research papers because of almost non-existent research reproducibility (nonpublic data sets, private source codes) and small to none attention of authors to deployment of their own research outputs.

The above mentioned reasoning leads us straightforwardly to the main aim of the paper, a proposal of a free, open source, modern, scalable and yet easily applicable software solution for deployment of supervised corporate financial distress (bankruptcy) prediction models focusing on SMEs. This software solution is designed in Shiny (Chang et al., 2016), a web application framework for R (R Core Team, 2016). Shiny and R are both open source and free, even for commercial activities, thus easily available for Slovak SMEs. Moreover, the main principles of friendly user graphics interface are applied to design the application as simple as possible to use in order to meet the needs and preferences of today's users.

The paper is outlined as follows. In Section 2 we describe the underlying statistical and data mining models and data sets serving as a basis for our application; in Section 3 we present basic principles of user experience we should follow while deploying statistical and data mining models; in Section 4 we discuss the developed application itself, both online and local versions.

## 2. Underlying Statistical Model and Data

We have decided to build our application using some corporate financial distress prediction models created within the framework of the scientific project supported by VEGA grant 1/0647/14 "Variant methods of corporate financial health prediction in the conditions of a dynamic economic environment". These classification models can be seen as alternatives to the well-known Altman's Z-score and they have been derived in (Kráľ et al., 2016) quadratic discriminant analysis, logistic regression, classification trees and random forests using different criteria for model selection, e.g. overall accuracy, specificity and AUC criterion. This particular set of statistical and data mining methods provide good balance of availability in R, classification ability, applicability and interpretability. In the first version of our application, we restricted ourselves to a model based on quadratic discriminant analysis (see

Appendix 1). We have chosen this particular method because of its simplicity and connection to the linear discriminant analysis used for Altmans's Z-score construction. On the other hand, this method is more general than linear discriminant analysis as it does not assume identical covariance matrices for different classes. More details can be found in (James et al., 2015). We exported the model from a local R installation and uploaded it to a Shiny server via the saveRDS() and readRDS() commands. Independent variables included in the model mimic those in the Altman's model, i.e. the set of independent variables consists of five financial ratios, namely

X1 – working capital / total assets,
X2 – retained earnings / total assets,
X3 – earnings before interest and taxes / total assets,
X4 – book value of equity / book value of debt, and
X5 – sales / total assets.

A dependent (qualitative) variable – a status of a company – takes two values "being in distress" and "not being in distress". Definition of the dependent variable has been taken from (Boďa and Úradníček, 2016b), where "an enterprise was considered financially distressed if (a) its equity was negative, (b) its EAT (earning after taxes) was negative, and (c) its current ratio attained a value lower than 1. All the three conditions had to be satisfied in order for an enterprise to be considered financial distressed".

The data sets used for model fitting were extracted originally for the paper (Boďa and Úradníček, 2016b) from the data base purchased from the leading Slovak corporate analytical agency CRIF – Slovak Credit Bureau, s.r.o, consisting of various financial indicators and covering economic activities 1110 – 96060 according to SK NACE classification. We have not applied any further data cleaning and data transformation procedures. The fact that assumptions of a particular method might be violated was ignored based on assumption that such a violation will degrade a prediction ability of the corresponding model and thus the model will be ignored.

More details on the model can be found in (Kráľ et al., 2016). More detailed description of data can be found in (Boďa and Úradníček, 2016b).

## 3. User Experience

As our aim is to develop application that could be used by managers of SMEs, it is important to create suitable user interface. That is why we discuss here the current trends and requirements of today's users.

According to Kurdi et al. (2014, p. 148): "Users desire friendly looking, highly interactive, dynamically visualized and naturally controlled user interfaces (UI) that allow smooth navigation."

The complex system of requirements and recommendations was originally developed by Nielsen and Molich (1990), later refined based on a factor analysis of 249 usability problems and published in Nielsen (1994) and Hertzum et al. (2016). General principles for interaction design can be summarized as follows. The application should keep users informed about what is happening, thanks to suitable feedback within reasonable time. The application should use the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Real-world conventions should be used, making information appear in a natural and logical way. Users sometimes choose options by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Provide undo and redo functions. Users should not have to wonder whether different words, situations, or actions mean the same. Better than good error

messages is a clear and precise design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action. Minimize the user's memory requirements by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the application should be visible or easily retrievable whenever appropriate. Accelerators, unseen by the novice user, may often speed up the interaction for the expert user such that the application can cater to both inexperienced and experienced users. Allow users to tailor frequent actions. Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility. Error messages should be expressed in plain language, precisely indicate the problem, and constructively suggest a solution. Even though it is better if the application can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

## 4. Application for Corporate Financial Distress Prediction

In this part we describe deployment of the application for corporate financial distress prediction. As we mentioned in the previous part of this article, the target audience for our application are mainly small and medium size enterprises.

The target concept of our application is the following:

- Prediction using statistical models: Application should be able to process statistical models and present results to the user. In this field there are more solutions. But if we consider only free available software, there are not so much possibilities.
- Usability: Simple user interface without need of any special knowledge in the field of information technology or statistics. For simple usability it is crucial that user interface corresponds to the users` knowledge and experiences. A web browser appears as the most suitable user interface, because it is daily used by our target group. There are a lot of different web browsers available, but the interface of all of them is almost identical.
- Availability: Available for 24 hours in a day, 7 days in week without any restrictions. Availability is crucial for our application and because of this we decided to deploy two versions of the application: online version and local version. Both versions must have the same user interface.
- Compatibility: Able to run on all platforms including desktops, notebooks, tablets or smartphones. Web browser as a user interface allows us to deploy cross-device and cross-platform application.
- Free of charge: Both versions are available for free, without registration.

From the functional point of view, our application has:

- Input: For prediction of possible corporate financial distress basic financial indicators are used. All of the indicators must be easy to understand by representatives of small and medium size enterprises. For calculation of probability of corporate financial distress, we use financial indicators included in Altman Z-Score.
- Output: Results from our model are represented by probability of corporate financial distress. The probability as the only result from the model may seems like too simplified, but in our case it was the purpose, as we need to make the results as simple as possible, because of easy interpretation by representatives of small and medium size enterprises.

As mentioned above, because of the availability requirements, we developed both online and local version of our application.

### 4.1 Online Version with Shiny Server

We decided to build online version of our application because of ubiquitous internet access and standard users' experience with web browsers. The most important benefits of online version are that users can use it directly in web browser without any installation. Many users prefer online applications for the reason that they do not have experiences with installation of local applications. Sometimes they even do not have administrator privileges on a computer or do not trust to unknown sources of software. Decision for online version of our application was also because of current trend - cloud computing. It is becoming very popular, especially by managers of SMEs, as they have access to various applications that are for free or it is possible to pay annual fee instead of buying them. Another advantages of cloud computing is also higher scalability, flexibility and security.

Our deployment is based on the statistical computing environment R with the package Shiny and software Shiny Server. We have chosen R because of simple implementation and processing of statistical models as this is essential for our application. Shiny package is elegant and powerful web framework for building web applications using R. Shiny Server is server-side application running on Linux servers. It is an environment for launching of R with Shiny package on a web server.

By combining of those components we achieved complex solution using only free and open source software. Another important benefit is the possibility to design layout of the application in details using standards for CSS and HTML.

Online version of our application can be found at https://efumb.shinyapps.io/deploy/.

### 4.2 Local Version

The main reason why we decided to create also local version of the application is that this version can work without internet connection. Internet connection is available almost everywhere, but in some situations local application can be very useful. For example, internet access failure or lack of internet access due to security reasons.

Design and functionality of both versions should be the same to increase users' experience. The online version has higher priority, because online applications are more popular then local ones. So we first deployed online application and the local version should use the same user interface and also the same software components.

As we already mentioned in the target concept, the basic requirement is capability to process statistical models. For offline version we use also R, because this program is very strong in processing of statistical models. We are aware of the fact that installation of R on a local computer of SME manager can discourage users from using of our application. Due to this fact, we include portable version of R in local version of our application. Portable versions of programs can work without installation and are optimal for our purposes. The local version of our application can be used on all operating systems supported by R and Shiny package (MS Windows, MAC OS, Linux).

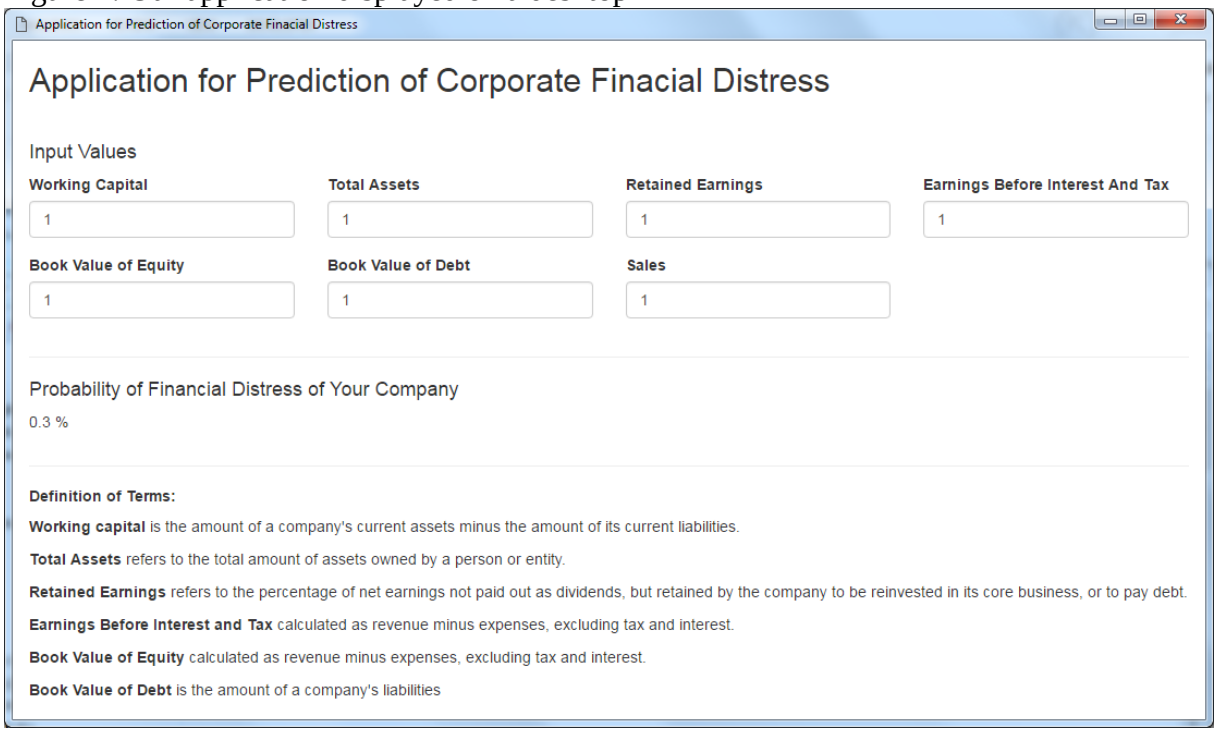
Graphics user interface is the same like in the online version and we need web browser to make it running. In case that our application will be installed on a computer that is not connected to the internet, version of a local web browser can be obsolete. Obsolete web browser does not support new application components and this fact can result malfunction of our application. To eliminate this possible problem, we included also web browser Google

Chrome Portable in the installation package of our application. Google Chrome is the most popular in Slovakia and users know this web browser very well.

Communication between R and Google Chrome is provided by Shiny package installed in R. We integrated all software into one application that can be run using shortcut located in the package. All required parameters are predefined.

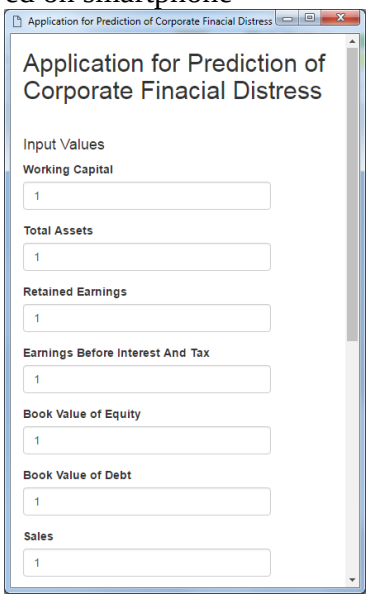
Figures 1 and 2 shows our application on desktop and smartphone.

Figure 1: Our application displayed on a desktop




Source: the authors.

Figure 2: Our application displayed on smartphone




Source: the authors.

In regard to target concept, we can conclude that the proposed application meets all defined requirements. From end users' point of view, the most important characteristics are usability, availability and free of charge use. From the software perspective, compatibility and prediction using statistical models are key characteristics. Source codes of both files ui.R and server.R can be found in appendices of the article.

## 5 Conclusion

According to the main goal of our paper we proposed application providing a supervised corporate financial distress prediction model for SMEs. It was designed in Shiny, a web application framework for R. This application was developed to run on commonly used technologies available for Slovak SMEs. A user graphic interface was designed in order to provide high user experience and usability, e.g. interpretation of results is also included in the application. Finally, our application is available on the Internet for free.

In the future, we plan to implement additional models and to improve the user graphic interface in accordance with the user feedback. We also plan to use the created tool to analyze the data available for Slovak companies to stratify them according to various criteria like sector or size.

We are convinced that this tool once finalized can help users, especially managers of small enterprises in Slovakia, to detect signs of financial distress. If they realize the current situation of their businesses, they can take measures that can possibly lead to reduction of problems caused by financial distress.

## Acknowledgements

Pavol Kráľ and Igor Kollár acknowledge support by VEGA 1/0647/14 grant: "Variant methods of corporate financial health prediction in the conditions of a dynamic economic environment".

## References

[1] ALTMAN, E. I. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. In The Journal of Finance, 2014, vol. 23, iss. 4, pp. 583-609.

[2] ALTMAN, E. I. 1983. Corporate financial distress : A complete guide to predicting, avoiding, and dealing with bankruptcy. New York : Wiley, 1983.

[3] ALTMAN, E. I., HALDEMAN, R. G., NARAYANAN, P. 1977. ZETA™ Analysis : A new model to identify bankruptcy risk of corporations. In Journal of Banking and Finance, 1977, vol. 1, iss. 1, pp. 29-54.

[4] BIELIKOVÁ, T., CÚT, S., ÚRADNÍČEK, V. 2014. The influence of the definition of risky company on financial situation diagnostic models in Slovak dynamic economic environment. In 7th International Scientific Conference on Managing and Modelling of Financial Risks, 2014, pp. 38-45.

[5] BOĎA, M., ÚRADNÍČEK, V. 2016a. Inclusion of weights and their uncertainty into quantification within a pyramid decomposition of a financial indicator. In Ekonomický časopis, 2016, vol. 64, iss. 1, pp. 70-92.

[6] BOĎA, M., ÚRADNÍČEK, V. 2016b. The portability of Altman's Z-score model to predicting corporate financial distress of Slovak companies. In Technological and Economic Development of Economy, 2016, vol. 22, iss 4, pp. 532-553.

[7] CHANG, W. et al.. 2016. Shiny: web application framework for R. R package version 0.13.2, https://CRAN.R-project.org/package=shiny.

[8] HERTZUM, M., MOLICH, R., JACOBSEN, N. E. 2014. What you get is what you see : Revisiting the evaluator effect in usability tests. In Behaviour & Information Technology, 2014, vol. 33, iss. 2, pp. 143-161.

[9] JAMES, G. et al. 2015. An introduction to statistical learning : With applications in R. New York : Springer, 2015.

[10] KRÁĽ, P. et al. 2016. Corporate financial distress prediction of Slovak companies : Z-score models vs. alternatives. In BOĎA, M., MENDELOVÁ, V. (eds.) 2016. Applications of Mathematics and Statistics in Economics 2016. Conference Proceedings. Banská Bystrica : OZ Financ, 2016.

[11] KURDI, H. A., HAMAD, S., KHALIFA, A. 2014. Towards a friendly user interface on the cloud. In Design, User Experience, and Usability : User Experience Design for Diverse Interaction Platforms and Environments, 2014, vol. 8518, pp. 148-157.

[12] LESÁKOVÁ, Ľ. 2008. Influence of globalisation on small and medium enterprises. In Ekonomický časopis, 2008, vol. 56, iss. 6, pp. 607-621.

[13] NIELSEN, J., MOLICH, R. 1990. Heuristic evaluation of user interfaces. In Proceedings of ACM CHI'90 Conference, 1990, pp. 249-256.

[14] NIELSEN, J. 1994. Usability inspection methods. New York : Wiley, 1994.

[15] R CORE TEAM 2016. R : A language and environment for statistical computing. Vienna : R Foundation for Statistical Computing, 2016. http://www.R-project.org/.

**Appendix 1** Confusion matrix of the implemented model based on quadratic discriminant analysis (AUC = 0.709)

| % of correctly classified enterprises | Training sample | Test sample |
|---|---|---|
| Overall | 77.22 % | 81.78 % |
| Non-distresses | 83.93 % | 89.98 % |
| Distresses | 29.74 % | 23.72 % |

Source: Kráľ et al. (2016).

**Appendix 2** Source code of ui.R file

```
# loading of required libraries
library(shiny)
library(MASS)

shinyUI(fluidPage(
  titlePanel("Application for Prediction of Corporate Finacial Distress"),
  br(),
  h4("Input Values"),
    fluidRow(
      column(3, numericInput("WK", label = "Working Capital", value=1)),
      column(3, numericInput("TA", label = "Total Assets", value=1)),
      column(3, numericInput("RE", label = "Retained Earnings", value=1)),
      column(3, numericInput("EBIT", label = "Earnings Before Interest And Tax", value=1))),
    fluidRow(
      column(3, numericInput("BVE", label = "Book Value of Equity", value=1)),
      column(3, numericInput("BVD", label = "Book Value of Debt", value=1)),
      column(3, numericInput("S", label = "Sales", value=1)),
      column(3, p())),
    fluidRow(
      column(12, hr(),
        h3("Probability of Financial Distress of Your Company"),
        p(textOutput("pred")))),
    fluidRow(
      column(12, hr(),
        h5(strong("Definition of Terms:")),
        p(strong("Working capital")," is the amount of a company's current assets minus the amount of its current liabilities."),
        p(strong("Total Assets")," refers to the total amount of assets owned by a person or entity."),
        p(strong("Retained Earnings")," refers to the percentage of net earnings not paid out as dividends, but retained by the company to be reinvested in its core business, or to pay debt."),
        p(strong("Earnings Before Interest and Tax")," calculated as revenue minus expenses, excluding tax and interest."),
        p(strong("Book Value of Equity")," calculated as revenue minus expenses, excluding tax and interest."),
        p(strong("Book Value of Debt")," is the amount of a company's liabilities")))))
```

**Appendix 3** Source code of server.R file.

```
# loading of all required libraries
library(shiny)
library(lattice)
library(ggplot2)
library(caret)

# loading of the external model from model1.RDS file
model1 <- readRDS("model1.rds")

shinyServer(
  function(input, output, session) {
    output$pred <- renderText({
      A1 <- input$WK/input$TA
      A2 <- input$RE/input$TA
      A3 <- input$EBIT/input$TA
      A4 <- input$BVE/input$BVD
      A5 <- input$S/input$TA
      data <-  data.frame(A1, A2, A3, A4, A5)
      data[1, ] <- c(A1, A2,  A3, A4, A5)
      predikcia <- predict(model1, data, type = "prob")
      paste(round(predikcia[1,2]*100,2)," %")
    })

    # close the R session when Chrome closes
    session$onSessionEnded(function() {
      stopApp()
      q("no")
    })
})
```