44 DeRis: Information System Supporting the Prediction of Default Risk in Companies

Cláudio Augusto Silveira Lélis and André Luiz Silveira Lopardi

Abstract

The risk of default has grown as a concern for financial institutions. In a scenario of uncertainties, the correct decision is essential in the granting of credit. A predictive model of default risk and the linking of conflict management strategies can be critical in reducing financial losses and in decision-making doubts. This article presents an information system, called DeRis (Default Risk Information System), designed to support activities in the management of default risk in the context of a bank focused on the granting of credit. It covers a default prediction model based on conflict indicators, management, and financial indicators, a reasoner and visualization elements. Collecting historical data and sorting indicators is also possible. Through an experimental study, quantitative and qualitative data were collected. The feasibility of using DeRis was verified through an experimental study.

Keywords

Conflict indicators · Financial management · Knowledge-based decision-making · Default prediction model · Data visualization

44.1 Introduction

Changes in the world financial scenario since the 1990s, such as deregulation of interest rates and exchange rates, increased liquidity and increased competitiveness, especially in the banking sector, have increased the concern of financial institutions with the risk of default. Although there are different concepts, default in the context of this research can be understood as a delay of more than 90 days in the liabilities assumed with a financial institution [1].

According to [2], credit risk is associated with the risk of a borrower or counterpart being defaulted. Thus, in the position of financial intermediaries, banks must act in a way that minimizes risk and enables fairer terms of credit acquisition. The difficulty of performing guarantees and recovering credit has led to uncertainty and instability in the market, making default the biggest cost of a bank's financial margin.

Over time, companies have been adapting to changes while remaining competitive and profitable in an increasingly crowded market. Applying constant investments in the area of Information Technology, financial institutions seek to offer products to their customers in a fast, safe and high technological value. Always attentive to high performance and information security, especially with the large volume of data. On the other hand, customers can count on the trust, performance and safety expected of a financial institution [3].

In this period, efficiency gained prominence and in companies where there were conflicts, there were also losses, impairing efficiency. According to the American Management Association (AMA), managers spend at least 24% of their time dealing with conflicts. Such conflicts represent, for example, a tension, a disagreement or polarization between two, or more people, or groups. It is part of the management and if it is treated in a constructive way, the conflict is considered as an opportunity [4].

The Bank Zak who became partner and offered their data for this research, for example, focuses its activities in generating resource and credit analysis. In addition, it fosters the consumption and investment needs of individuals and companies. Recognized the importance of the default risk has become important for the Bank Zak check if companies are shaping the economic environment, assessing their conflicts in business scope and influence on a possible default.

C. A. S. Lélis (✉)
Scientific Initiation Program, IMES/Faculty ImesMercosur, Juiz de Fora, Brazil
e-mail: claudioaugustolelis@imes.org.br

A. L. S. Lopardi
Master's Program in Business Administration, FUMEC University, Belo Horizonte, Brazil

© Springer International Publishing AG, part of Springer Nature 2018

S. Latifi (ed.), *Information Technology – New Generations*, Advances in Intelligent Systems and Computing 738,
https://doi.org/10.1007/978-3-319-77028-4_44

The use of default prediction models serves to measure, monitor and predict the financial situation of companies, reducing uncertainties and doubts in decision making [5]. The models are constructed with the support of statistical techniques and applied to analyze their dependent variables.

A financial institution should identify risks in lending situations, draw conclusions as to the borrower's ability to repay, and make recommendations regarding the best structuring and type of loan to be granted in the light of the applicant's financial needs [6]. In a scenario of uncertainties and incomplete information, risk analysis involves the ability to establish a decision rule to guide the granting of credit.

For the survival of financial institutions, the correct decision to grant credit is essential [7]. Any error in the decision to grant the credit may mean that the gain on other successful transactions is lost in a single transaction. Therefore, it is important to anticipate and reduce default [8], since the losses from unsuccessful credits should be covered by charging high interest rates on new concessions. Therefore, using a default risk forecasting model for a financial institution and linking management strategies to the reality of the borrower can be critical in assessing credit risk and reducing financial losses.

Considering that Zak Bank has management conflicts and the possibility of linking conflict reduction to efficiency and productivity gains, there is a need to evaluate possible organizational variables that impact the bank's course and its perpetuity. In addition, to verify the influence of conflict indicators in a default forecast model from Zak Bank's point of view.

For proper monitoring of the indicators involved, a significant amount of data must be collected, processed and stored over time. The process of discovering knowledge through these data may be associated with an information system that offers this information through interactive visualizations, aiming the recognition of new knowledge and assisting decision making. Visualizations enable the manager to interact and gain insight into the data they have, gaining new insight through distinct views across different perspectives. Such systems minimize the occurrence of misinterpretations when compared to analysis performed through a single view [9]. This demonstrates the importance of a support system in this context.

Therefore, this paper presents an information system, called DeRis, aimed to support activities in the management of default risk in the context of Zak Bank. It encompasses a default prediction model based on conflict indicators, management, and financial indicators, a reasoner and visualization elements. Through the storage of decisions a knowledge database is generated. Collecting historical data and sorting indicators is also possible.

This article is structured by this introduction and Sect. 44.2 shows the background in which the proposal is inserted and some related work. Section 44.3 details the components of the DeRis system. In Sect. 44.4, the experiment carried out and Sect. 44.5 the final considerations.

44.2 Background

Through a survey of the specialized technical literature, it was possible to perceive that the researchers' interest in default risk models dates back to the 1930s [10, 11]. Over the years, the pioneering work of Beaver [12] and especially Altman [13], boosted research in the 1970s with accounting indicators [14–18]. Since the mid-1990s, issues such as the emergence of new modeling techniques, the growing importance of credit risk management and the prevailing economic conditions, again aroused the interest in the area [19–22].

There are several techniques applied to credit risk forecasting models. They can be classified as discriminant analysis used in the model proposed by [13], neural networks, multiple linear regression, linear programming, genetic algorithms, decision tree, logistic regression used in the DeRis system model, and more recently the analysis of survival.

Bellovary et al. [23] investigated the main financial indicators used in studies to predict default and found the current liquidity present in 51 studies among those analyzed.

Bonfim [24] examined the determinants of corporate defaults in the banking sector in Portugal through the Logit or Probite Models of Survival Analysis. The study found that default is affected by specific characteristics of companies such as: capital structure, company size, profitability and liquidity, recent sales performance and investment policy. However, there was a significant improvement in the quality of the models, with the introduction of variables, especially the growth rate of all the riches produced in the country, the growth of lending, the average lending rate and the variation of stock market prices.

The model presented by [25] used the two most important macroeconomic factors that affect corporate default which are the nominal interest rate and the output gap. As financial variables specific to each company, the authors used the Earnings Before Interest, Taxes, Depreciation and Amortization (EBITDA) per Total Assets ratio, the interest coverage ratio, the leverage ratio, the total liability ratio and revenues, the ratio net assets and total liabilities and, finally, inventory turnover.

From the identified models, as well as the concepts of credit and risk, it should be noted that the default forecast models, mostly, have financial indicators as explanatory variables. Therefore, creating a model without taking such indicators into account would put its effectiveness in question.

However, the models identified in the survey conducted were not associated with information systems with visualizations capable of aiding decision making. In addition, indicators of conflicts in management have not been the object of study of the researchers, which justifies the interest for the present work.

44.3 DeRis System

The proposed DeRis system arose from the need to predict the financial situation of companies to avoid default, as well as support managers and financial institutions in making decisions regarding the granting of credit. Figure 44.1 shows an overview of the DeRis system architecture with its main components and the basic flow of information in the decision-making process.

The information flow while using the system starts in the repository that stores the historical data of the indicators. Additional information such as a brief description and the indicator classification are also stored. Indicator management is performed to determine which indicators will be used in the prediction model. This choice is made considering the possibility of calculating the indicator with the available data. Then, the default prediction model is triggered. Reasoner helps managers interpret the model results. Finally, the visualizations show the results and analyzes made by Reasoner and through interaction elements managers can associate preventive measures with the indicators. A knowledge database is also maintained as decisions are taken, that feeds the system and this information is recorded to support future decisions. The other components of the system are detailed below.

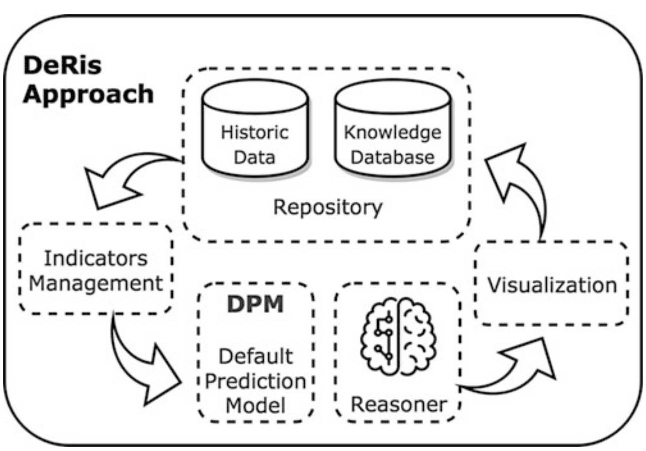

Fig. 44.1 DeRis architecture

44.3.1 Indicators Management

This is an important component of the system. It identifies the indicators whose data is stored, its additional information and mainly the classification that assists the Reasoner in the analyzes. These are the indicators that are candidates to participate in the model. It is necessary to make the selection of the indicators when starting the monitoring of a company, or through previous decisions, the own component preselects the indicators.

Another feature of this component is to relate the indicator to the way of collecting, or calculating, its value. For example, in the context of the Zak Bank considered in this study the indicators were classified between: conflict, management or financial as described below.

- **Number of analyzable business partners:** Represents the partners analyzed that have at least more than 1% of the company. **Acronym:** NABP. **Class:** Conflict.
- **Number of analyzable business partners by total:** Represents the ratio of partners who have at least more than 1% of the company by the total number of partners. **Acronym:** NABPT **Class:** Conflict
- **Age of oldest business partners:** Represents the age of the older partners over the age of 18. **Acronym:** AOBP **Class:** Conflict
- **Age of oldest leader business partners:** Represents the age of the older leaders over the age of 18. **Acronym:** AOLBP **Class:** Conflict
- **Gross Annual Revenue/1,000,000:** Represents all revenues earned by companies during the year. For scale purposes, it is divided by 1 million. **Acronym:** GAR/1,000,000 **Class:** Financial
- **Average balance in account and investment/Exposure in the last 12 months/1000:** Defined by the amount in cash and financial application with immediate liquidity that the company owns. It is divided by short-term obligations and represents current liquidity [23]. **Acronym:** (BAI/E)/1000 **Class:** Financial
- **Average balance in account and investment/GAR in the last 12 months:** Represents the amount in cash or financial application, divided by gross annual revenue. **Acronym:** BAI/GAR **Class:** Financial
- **Usage of Investment Lines Indicator:** Defined to indicate the management of the companies cash flow. **Acronym:** UIL(YES) **Class:** Financial

- **Account time in years:** It considers the time between the opening of the current account in the financial institution, until the current date. **Acronym:** ATY **Class:** Management
- **Activity time counted from the operation start:** Defined from the date on which the company actually started its operational activities. **Acronym:** ATOS **Class:** Management
- **Operating time at last address:** Defines the time of permanence in the last registered address. **Acronym:** OTLA **Class:** Management
- **Number of employees:** Defined by the number of employees informed to the financial institution when registering. For reasons of scale, the indicator is shown in the ratio of 10 employees. **Acronym:** NE/10 **Class:** Management

Although the indicators have been collected and are in the context of the Zak bank, the system is prepared to add new indicators and classifications, making it possible to adapt it to other contexts.

44.3.2 Prediction Model

The proposed model is based on conflict, management and financial indicators, classified in the indicators management stage.

The technique applied by the model is logistic regression [26] that allows analyzing the effect of one or more independent variables on a dichotomous dependent variable, representing the presence or absence of a characteristic. In this way, it describes the relationship between several independent variables. According to this theory, the model calculates the probability of default, given by Eq. (44.1):

$$ProbDefault(yes) = \frac{e^{\eta}}{1 + e^{\eta}} \tag{44.1}$$

where η depends on the indicators and data available for the logistic regression calculation. In the next section, the two equations used for the η calculation during the evaluation are shown.

44.3.3 Reasoner Phase

Assuming that there are indicators A, B and C. However, only the data for indicators A and C are available. Given this, *would it be possible to replace B? Which indicator could replace it?* Questions like these that Reasoner tries to answer with their analysis.

These questions can be answered by the Reasoner due to information such as: the class to which the indicator belongs (Conflict, Management or Financial), the unit of indicator measure, its degree of influence on the default and decisions taken previously after the exchange of indicators.

Another important function is to relate a moment of the past with a description of the decision made and what were the critical indicators for default, based on historical data and the knowledge base.

Although the review process is transparent to the user, the decision to replace an indicator is performed by the manager, when necessary, whenever a company's monitoring begins.

44.3.4 Visualization

The DeRis system provides two views developed from the JavaScript d3js library considering its flexibility of use and availability for editing. This library has been used successfully in the context of visualizations in information systems [27].

The Indicators Evolution View (Fig. 44.2) uses a line chart to represent the evolution of the default probability over time. Furthermore, it uses a gauge metaphor to highlight the influence percentage of each indicator considered on the probability of default. This decision was made since it would not make sense to compare its absolute values since the indicators are on different scales.

The view is generated after the user selects a range of days and a time period to be analyzed. As shown in Fig. 44.2, the line and the gauges are arranged as a dashboard for the decision maker.

Through interaction elements, it is possible to select a point in the line graph and obtain contextual information, such as: the future trend of the default probability, the values of each indicator, the model used and the decision taken at the time, if any. Moreover, the percentage of each indicator, colored according to the variation to the previous occurrence. This feature allows the analysis of the variability in the influence of indicators.

44.4 Evaluation

This section presents the experimental study conducted. According to the Goal/Question/Metric approach (GQM) [28] the goal can be stated as: **Analyze** the DeRis system **in order to** verify the feasibility of use **with respect to** the comprehension of conflict, management and financial indicators as a support tool to default prediction model **from the point of view of** managers and financial institution professionals **in the context of** Zak Bank through a decision support system.

In this sense, the metrics defined to verify the fit quality of the models were the result of the Hosmer-Lemeshow test

Fig. 44.2 Indicators evolution view

[26] based on p-value with significance level of 0.05. The sensitivity that is the ratio of true positives, assesses the ability of the model to classify that event occurred to an individual, since the event actually occurred. Finally, the specificity that is the ratio of true negatives, assesses the ability of the model to predict that the event did not occur.

The experiment was proposed based on a set of real data collected from documentary sources, from January to December 2015. The sample comprised 15,000 companies, between defaulters and debt free companies, from various sectors of the economy located in the database of Zak Bank, intentionally selected, choosing management, conflict and financial indicators.

It is worth noting that there was a concern about the validity of the results found due to the fact that the sample was unbalanced, with 952 companies in default and 14,048 companies free of debt. Therefore, it was decided to repeat the process with a new sample that was balanced and thus obtain a revalidation of the model. For the construction of this sample, all the defaulters were selected (952) and a random sample was made in the 14,048 so that 952 were selected. In order to guarantee the external and internal validity of the results found, a cross validation was used.

44.4.1 Descriptive Analysis

The collected data indicates that 6.35% of the companies were in default and 28.65% used the investment lines. Considering family businesses, it was certified that 21.35% belong to the universe surveyed. It can be said that the companies had on average 1.48 analyzable partners, with a standard deviation of 0.63. It was observed that 48.09 years is the age of oldest business partners, with a standard deviation of 12.99. It is important to note that the average account time of the companies was 5.62 years, with a standard deviation of 5.02 years. The results indicate that firms had on average 7.05 employees, with a standard deviation of 11.93.

44.4.2 Statistical Analysis

The Logistic Regressions adjusted for each possible variable predicting the occurrence of default allowed to verify, in an isolated way, the impact of each variable on the default.

Data analysis showed that the variables: *Average balance in account and investment/GAR in the last 12 months* and *Number of employees* were potential risk factors for the occurrence of default. The other indicators were considered potential protection factors for the occurrence of default.

44.4.2.1 Model 1: Unbalanced Sample

From the data collected and analyzed, the equation of the selected model was presented to calculate the default probability, whose η value is calculated by:

$$\begin{aligned}\eta =& -1.752 - 0.166 * NABP - 0.340 * \frac{(BAI/E)}{1000} \\ & - 0.191 * UIL(YES) - 0.074 * ATY \\ & - 0.019 * ATOS + 0.069 * NE/10\end{aligned} \tag{44.2}$$

Considering the prevalence of default (6.35%), a sensitivity of 73.22%, a specificity value of 45.33%, was observed. Regarding the Hosmer-Lemeshow test, the p-value equals 0.085, which indicates a well-adjusted model.

The results show that the variable, "average balance in current account and investment/exposure in the last 12 months" was the most important to determine the default, and then, account time in years, use of investment activity counted from the start of operation, number of analyzable members and number of employees.

44.4.2.2 Model 2: Balanced Sample

When considering a balanced sample, the model did not indicate as significant the variables: Number of analyzable business partners (NABP) and number of employees (NE). Other variables and interpretations were the same. Thus, the relationships between default and NABP and NE, found in the model that considers the sample unbalanced, are not validated.

It can also be verified that the model with the balanced sample was capable of pointing 78.18% of sensitivity and predicting 41.00% of specificity. The Hosmer-Lemeshow test indicates that the model is well-adjusted (p-value = 0.200).

The equation of the model based on the balanced sample has η calculated by:

$$\begin{aligned}\eta =& 0.775 - 0.469 * \frac{(BAI/E)}{1000} - 0.263 * UIL(YES) \\ & - 0.081 * ATY - 0.016 * ATOS\end{aligned} \tag{44.3}$$

The results showed that the "Average balance in account and investment/Exposure in the last 12 months" was the most important to determine default. Then, are the following variables: "Account time in years", "Usage of Investment Lines indicator" and "activity time counted from the operation start".

44.4.3 Lessons Learned

By means of the data collection, 15,000 companies were analyzed, using indicators provided by Zak Bank. Once the companies to be searched are chosen, they come up with some interesting discoveries.

By identifying the main variables of the available conflict indicators that influence a default forecast model, such as: number of members, age of members and family companies, it was found, based on field research, that the larger the number of members the lower the probability that the company will default. One aspect that may have been captured by the model is that of divided responsibility, that is, with more partners there is more equity and subdivisions of the responsibilities assumed. Therefore, unlike the premise used in the research, conflicts related to the number of partners do not affect the risk of default by companies. The age of the partners also did not show enough influence to explain the default.

The family business indicator was not consistent to influence the risk of default. Thus, the fact that the company is familiar and having conflicts does not affect its default risk.

With regard to the influence of conflict indicators, it can be seen that, among the variables used in the default model of the research, only one (number of analyzable members) among the three in the conflict block was consistent to explain the default rate.

It was found that the higher the liquidity of the company, the lower its probability of default. The facilitating variable for the development of the model (use of investment lines) was efficient to explain the default, that is, the greater the use of these lines by the companies, the less likely they were to default.

The fact that all management variables are representative in the model is particularly significant. The variables (time of account in years and time of activity counted from the beginning of the operation) explained that the longer the company's experience in the market, the lower the probabilities of default. It should be noted that the variable (number of employees/10) showed an inverse relationship, that is, the higher the number of employees, the greater the probability of default. A more accurate analysis allows to infer a possible existence of conflicts within these companies.

44.5 Closing Remarks

This paper introduced the DeRis system to support predict the financial situation of companies to avoid default, as well as support managers and financial institutions in making decisions regarding the granting of credit. An experimental study was conducted and through the statistical analysis of the data, the system feasibility has been verified. Considering the results, it was concluded that the financial, management and conflict indicators are useful as an aid tool to the default prediction model.

As a contribution, this work offers financial institutions an approach to encourage the use of management, conflict and financial indicators, as a specification for the development of a default forecasting model, in the search for continuous improvement. A limitation of the study is associated to the universe of variables previously defined by the financial institution. It is also worth mentioning the impossibility of obtaining data from more than one financial institution. Thus, the results of the research are limited to the particularities of the financial institution under study. Its application in other institutions is subject to market conditions and may present different results.

As future work we intend to elaborate a study comparing other statistical techniques, such as: neural networks, discriminant analysis, among others. Moreover, integration with other visual tools to diversify the presented data possibilities.

References

1. J. Bessis, *Risk Management in Banking* (Wiley, New York, 2011)
2. S. Westgaard, N. Van der Wijst, Default probabilities in a corporate bank portfolio: a logistic model approach. Eur. J. Oper. Res. **135**(2), 338–349 (2001)
3. C.G. Bernardo, Loan system in Brazilian financial institution-a SOA application, in *2012 Ninth International Conference on Information Technology: New Generations (ITNG)* (IEEE, New York, 2012), pp. 293–298
4. G. Kohlrieser, Six essential skills for managing conflict. Perspect. Manag. **149**, 1–4 (2007)
5. M.T. Dugan, C.V. Zavgren, How a bankruptcy model could be incorporated as an analytic. CPA J. **59**(5), 64 (1989)
6. D. Duffie, K.J. Singleton, *Credit Risk: Pricing, Measurement, and Management* (Princeton University Press, Princeton, NJ, 2012)
7. L.J. Gitman, R. Juchau, J. Flanagan, *Principles of Managerial Finance* (Pearson Higher Education AU, Boston, MA, 2015)
8. M. Steiner, C. Carnieri, B. Kopittke, P.S. Neto, Probabilistic expert systems and neural networks in bank credit analysis. Int. J. Oper. Quant. Manag. **6**(4), 235–250 (2000)
9. S. Ainsworth, The functions of multiple representations. Comput. Educ. **33**(2), 131–152 (1999)
10. P.J. FitzPatrick, A comparison of the ratios of successful industrial enterprises with those of failed companies. Certif. Public Account. **12**, 598–605 (1932)
11. A. Winakor, R. Smith, Changes in the financial structure of unsuccessful industrial corporations. Bulletin **51**, 44 pp. (1935)
12. W. H. Beaver, Financial ratios as predictors of failure. J. Account. Res. **4**, 71–111 (1966)
13. E.I. Altman, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. J. Finance **23**(4), 589–609 (1968)
14. R.O. Edmister, An empirical test of financial ratio analysis for small business failure prediction. J. Financ. Quant. Anal. **7**(2), 1477–1493 (1972)
15. E.B. Deakin, A discriminant analysis of predictors of business failure. J. Account. Res. **10**, 167–179 (1972)
16. M. Blum, Failing company discriminant analysis. J. Account. Res. **12**, 1–25 (1974)
17. R. A. Eisenbeis, Pitfalls in the application of discriminant analysis in business, finance, and economics. J. Finance **32**(3), 875–900 (1977)
18. R.C. Moyer, Forecasting financial failure: a re-examination. Financ. Manag. **6**(1), 11 (1977)
19. E.I. Altman, G. Marco, F. Varetto, Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks (the Italian experience). J. Bank. Finance **18**(3), 505–529 (1994)
20. P. Brockett, W. Cooper, L. Golden, X. Xia, A case study in applying neural networks to predicting insolvency for property and casualty insurers. J. Oper. Res. Soc. **48**, 1153–1162 (1997)
21. C.Y. Shirata, Financial ratios as predictors of bankruptcy in Japan: an empirical research. Tsukuba College Technol. Jpn. **1**(1), 1–17 (1998)
22. C. Lennox, Identifying failing companies: a re-evaluation of the logit, probit and da approaches. J. Econ. Bus. **51**(4), 347–364 (1999)

23. J.L. Bellovary, D.E. Giacomino, M.D. Akers, A review of bankruptcy prediction studies: 1930 to present. J. Financ. Educ. **33**, 1–42 (2007)
24. D. Bonfim, Credit risk drivers: evaluating the contribution of firm level information and of macroeconomic dynamics. J. Bank. Finance **33**(2), 281–299 (2009)
25. T. Jacobson, J. Lindé, K. Roszbach, Firm default and aggregate fluctuations.J. Eur. Econ. Assoc. **11**(4), 945–972 (2013)
26. D.W. Hosmer Jr., S. Lemeshow, R.X. Sturdivant, *Applied Logistic Regression*, vol. 398 (Wiley, New York, 2014)
27. C.A. Lélis, M.A. Miguel, M.A.P. Araújo, J.M.N. David, R. Braga, Ad-reputation: a reputation-based approach to support effort estimation, in *Information Technology-New Generations* (Springer, Berlin, 2018), pp. 621–626
28. V.R. Basili, D.M. Weiss, A methodology for collecting valid software engineering data. IEEE Trans. Softw. Eng. **SE-10**(6), 728–738 (1984)