Expert Systems with Applications 42 (2015) 6277–6285

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

# Innovative default prediction approach

Július Bemš [a,*], Oldřich Starý [a], Martin Macaš [b], Jan Žegklitz [c], Petr Pošík [c]

[a] Department of Economics, Management and Humanities, Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, 166 27 Prague, Czech Republic
[b] Czech Institute of Informatics, Robotics, and Cybernetics, Czech Technical University in Prague, Technická 2, 166 27 Prague, Czech Republic
[c] Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, 166 27 Prague, Czech Republic

ARTICLE INFO

*Article history:*
Available online 29 April 2015

*Keywords:*
Magic
Square
Default
Prediction
Scoring

ABSTRACT

This paper introduces a new scoring method for company default prediction. The method is based on a modified *magic square* (a spider diagram with four perpendicular axes) which is used to evaluate economic performance of a country. The evaluation is quantified by the area of a polygon, whose vertices are points lying on the axes. The axes represent economic indicators having significant importance for an economic performance evaluation. The proposed method deals with magic square limitations; e.g. an axis zero point not placed in the axes origins, and extends its usage for an arbitrary (higher than 3) number of variables. This approach is applied on corporations to evaluate their economic performance and identify the companies suspected to default. In general, a company score reflects their economic performance; it is calculated as a polygon area. The proposed method is based on the identification of the parameters (axes order, parameters weights and angles between axes) needed to achieve maximum possible model performance. The developed method uses company financial ratios from its financial statements (debt ratio, return on costs etc.) and the information about a company default or bankruptcy as primary input data. The method is based on obtaining a maximum value of the Gini (or Kolmogorov–Smirnov) index that reflects the quality of the ordering of companies according to their score values. Defaulted companies should have a lower score than non-defaulted companies. The number of parameter groups (axes order, parameters weights and angles between axes) can be reduced without a negative impact on the model performance. Historical data is used to set up model parameters for the prediction of possible future companies default. In addition, the methodology allows calculating the threshold value of the score to separate the companies that are suspicious to the default from other companies. A threshold value is also necessary for a model true positive rate and true negative rate calculations. Training and validation processes for the developed model were performed on two independent and disjunct datasets. The performance of the proposed method is comparable to other methods such as logistic regression and neural networks. One of the major advantages of the proposed method is a graphical interpretation of a company score in the form of a diagram enabling a simple illustration of individual factor contribution to the total score value.



## 1. Introduction

The term *magic square* (see Fig. 1) was firstly used in macroeconomics by a German economist and former minister of finance Karl Schiller (Medrano-B & Teixeira, 2013). A magic square is a diagram with four perpendicular axes on which are depicted country main macroeconomics indicators – gross domestic product (GDP), growth rate, consumer price inflation rate, unemployment rate and a balance of trade to a GDP ratio (Fialová, 2006). The area of a quadrangle is used for a relative comparison between countries. The higher quadrangle area, the better economic performance of the examined country. The idea of economic performance evaluation based on the stated indicators was introduced by Nicolas Kaldor (born Miklós Káldor), Hungarian economist (Kaldor, 1971).

This chart (Fig. 1) is a special case of a spider (radar) chart, where the number of axes can be higher than four. The disadvantage is that the axes do not have the zero point in the intersection and a quadrangle area varies with a zero location on each axis. The example of another disadvantage is that the consumer price inflation rate below zero (deflation) is not desirable (Pontiggia, 2012), but the quadrangle area is growing in this case.

* Corresponding author. Tel.: +420 224 353 309.
*E-mail addresses:* julius.bems@fel.cvut.cz (J. Bemš), staryo@fel.cvut.cz (O. Starý), macasm1@ciirc.cvut.cz (M. Macaš), zegkljan@fel.cvut.cz (J. Žegklitz), petr.posik@fel.cvut.cz (P. Pošík).

http://dx.doi.org/10.1016/j.eswa.2015.04.053
0957-4174/

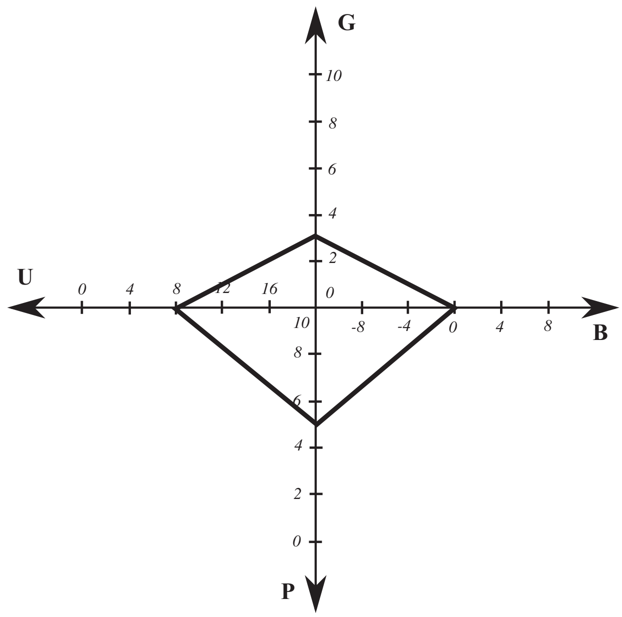


**Fig. 1.** Example of Magic Square (G - GDP growth rate, U - unemployment rate, P - inflation rate, B - trade balance), all numbers are in percentages.

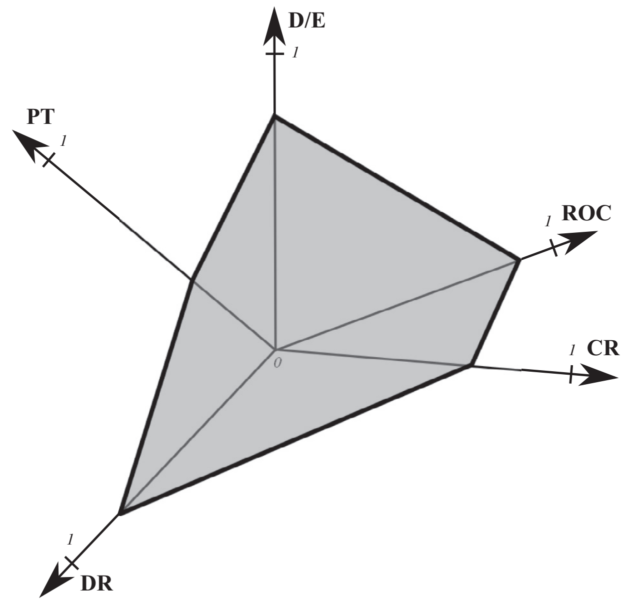


**Fig. 2.** Example of the diagram for the company. The grayed area represents a company score. (PT – payables turnover, D/E – debt to equity, ROC – return on costs, CR – current ratio, DR – debt ratio).

A similar principle of evaluation and comparison can be used also in case of companies. There are several axes with performance indicators (Fig. 2). A company with a larger polygon area has better performance (is rated higher) than a company with a lower polygon area. To avoid the above stated disadvantages, all axes have the same origin (zero point) and the values are transformed to the numbers in a 0–1 interval according to ''the higher, the better'' principle. The final polygon area is used as a company score in the same way as Altman did in his original work (Altman, 1968).

An Altman's innovative approach lay in applying a discriminant analysis on the data and the use of multiple variables for predicting a company default. Many other approaches based on multivariate regression, well described in Bishop (2006), have been used since that time. Default prediction techniques are very important for the banks that need the risk estimation of debtors. Logistic regression is mostly used in a bank sector for the probability of default estimation and is recommended to use by BASEL2.[1] For the default prediction, researchers have focused mostly on machine learning (closely related to statistics) algorithms recently. A new feature selection (FS) boosting procedure (Wang, Ma, & Yang, 2014) was introduced. Feature selection eliminates features with small predictive power, reduces dimensionality of feature space and removes irrelevant data. Boosting is a machine-learning algorithm for reducing variance and bias in supervised learning. FS-Boosting combines these two approaches and results in an alternative method for bankruptcy prediction. Genetic algorithms are being improved. The author in Kozeny (2015) introduces a new fitness function based on a variable bitmask. Research in the field of support vector machines (SVM) also moved forward and the clustered SVM were used in credit scoring (Harris, 2015). The advantage of clustered SVM lies in good performance and low computational complexity. The improvement of machine learning algorithms or innovative approach in their usage are evident in current research, where adaboost (Heo & Yang, 2014) genetic algorithms (Gordini, 2014) and neural networks (López Iturriaga & Sanz, 2015) are being used in the field of bankruptcy prediction. The dynamic models considering the time development of indicators are also represented by terminal failure processes (du Jardin, 2015). These models have better prediction performance in a long term. Improving prediction accuracy and identification of relevant predictive variables were two main objectives of the Least Absolute Shrinkage and Selection Operator (Tian, Yu, & Guo, 2015). Modern approach leads to often-complex technical improvements of machine learning methods.

A unique contribution of this paper is in the application and adaptation of macroeconomic approach (a magic square) in the field of company default prediction. The quadrangle is extended to a general polygon. In addition, the adjustment of angles between polygon axes, indicator weights and a scaling factor is innovative and here introduced parameters are easily imaginable in comparison to the parameters used in the above mentioned methods. The area of a polygon directly quantifies a risk factor in comparison to the magic square whose area could be used only for a relative comparison between countries. An easily adaptable system on different knowledge and databases provides a high-class approach for the bankruptcy prediction and overall economic company evaluation. Finally, yet importantly, a graphical interpretation helps in a quick orientation in the field of company risk factors and overall risk assessment.

An ideal scoring method evaluates each defaulted[2] company[3] by a score value lower than any non-defaulted company. It leads to an ordering performance evaluation of the scoring method quantified by the Gini and Kolmogorov–Smirnov indices. Another possible method is to find out the score threshold value, which means that the companies with the score below this value are treated as default and the companies with score above this value are treated as non-default. Subsequently, the true positive rate (sensitivity) and true negative rate (specificity) can be calculated. The threshold value discovery is inescapable for the default prediction. The difference between this scoring approach and classification models such as logistic regression or neural network is that Gini and Kolmogorov–Smirnov

[1] Recommendations on banking laws and regulations issued by the Basel Committee on Banking Supervision.

[2] The definition of the defaulted company varies. It can be either a company which is not able to fulfil its obligations in a specified time (e.g. 90 days) or a bankrupted company. However, it can be neglected as this paper aims at presenting a new scoring method.

[3] A company is not defaulted in the time of evaluation.