ECONOMY AND FINANCE • VOL. 12. ISS. 1 • MARCH 2025 • DOI: 10.33908/EF.2025.1.2

IMPACT OF DYNAMIC VARIABLES IN BANKRUPTCY PREDICTION MODELS BASED ON LOGISTIC REGRESSION

Tünde Katalin Szántó[1]

ABSTRACT

The study focuses on logistic regression, a method applied with the majority of scorecards used in bank lending processes. The author is trying to answer the question whether considering the trends of financial indicators can improve ranking accuracy for bankruptcy prediction models. The research used the sample of 1527 enterprises in the construction industry. The findings of the study suggest dynamic variables have improved the accuracy of bankruptcy prediction particularly improving the ranking accuracy of operational enterprises. Thus, banks are advised to analyse the trends of financial indicators to make lending decisions. There is an important difference here between lending to start-ups or to enterprises with a long operational history. Obviously, you cannot study earlier values of indicators with a start-up company; for this reason, the spread of the practice may further deepen (restrict) the options of start-ups to have access to resources. Therefore, the method should rather be employed for companies with a long history. The findings of the study also indicate that large enterprises are less threatened in terms of bankruptcy.

JEL codes: G33, C6, G17

Keywords: corporate bankruptcy prediction, default risk, logistics regression, dynamisation.

1 *Tünde Katalin Szántó,* PhD student. E-mail: szanto.tunde.katalin@o365.u-szeged.hu.

1 INTRODUCTION

1966 can be regarded to be the commencement of modern bankruptcy prediction for companies, as it was the year when Beaver published his bankruptcy prediction model based on single-variable discriminant analysis. Methods and techniques for bankruptcy prediction have proliferated over the decades since, while the indicators used have also multiplied, however, no consensus has been reached as to the methods and indicators to be applied to analyse the survival ability of enterprises.

Researchers mostly use static indicators for bankruptcy prediction, i.e., they try to come to conclusions regarding the future of enterprises on the basis of their position at a given point in time. Bankruptcy, however, is usually the outcome of a lengthy process in time, and that is why additional information can be gained for the models if the trends of the indicators are analysed (Nwogugu 2007). This study is trying to answer the question whether considering the trends of financial indicators can improve ranking accuracy for bankruptcy prediction models. Studies relating to the national economy as a whole have already been made on the topic. This study, however, is trying to answer the question whether dynamic indicators can result in higher accuracy of ranking for the bankruptcy prediction of construction companies.

The research used a sample of enterprises in the Hungarian construction industry. There are 1527 enterprises in the sample in total, of which 1188 are operational while 339 are under liquidation proceedings. It also shows that in reality properly operating healthy companies represent a higher rate than those in bankruptcy.

The methodology of logistic regression was used in the study. The reason for its application is that banks favour the methodology in their lending practices (Rajka–Pollák, 2024). Logistic regression is still a wide-spread method although recent studies tend to pair it with the application of decision trees (Márton et al., 2023). Logistic regression also has an advantage as it does not require large computation capacity, so enterprises can use it to analyse their survival chances. The models have been compared by the size of the area below the ROC curve. Their effectiveness has also been checked on an independent test sample including 105 solvent and 45 insolvent companies.

2 LITERATURE REVIEW

There is no agreement in the professional literature as to when you can speak of economic bankruptcy in the life of an enterprise. Bankruptcy, insolvency, economic failure, or missing payments are all terms often used to describe unsuccessful companies, and their meanings are frequently confused in the relevant studies. According to Constand and Yazdipour (2011), there is no consensus about the concept of bankruptcy in the professional literature. Sharma and Mahajan (1980) are of the opinion that diagnosing economic failure is the most difficult step in bankruptcy prediction. In Greenwald's economic dictionary (1973), economic failure means the event when an economic enterprise – voluntarily or as a result of legal proceedings – gives up its business activities causing losses to its creditors in that way. Giving up business activities may be the result of several causes, such as capital loss, insufficient profit or retirement. According to Dun and Bradstreet's study published in 1978, a company winding up cannot be considered busted if creditors' accounts receivable have fully been satisfied. Many types of economic failure and bankruptcies can be differentiated, however, bankruptcy in legal terms is a clearly defined type, so it is the one this study uses.

In legal terms, bankruptcy means insolvency, i.e., the event when an enterprise cannot meet its payment obligations by deadline. Bankruptcy, however, is not something appearing out of the blue. It is a lengthy process and a potential outcome of financial difficulties. An enterprise will face financial difficulties if there is impairment as a result of the unsatisfactory effectiveness of assets or a poorly established asset portfolio. Because of impairment, the market value of the asset portfolio declines, which will increase financial leverage in the company. The above effects lead to liquidity problems, and - as a result of the process - a state of insolvency may arise, which means bankruptcy in the legal sense (Pálinkó–Svoób, 2016). A bankruptcy situation typically takes some time to evolve, which allows its prediction for the companies involved.

In Hungary, there are two procedures in the event of insolvency. One, termed bankruptcy proceedings, is a type of reorganisation. Its objective is that having come to an agreement with its creditors the debtor should reorganise its operations and continue its business. The final goal of the procedure is to sign an agreement between debtor and creditor, so that the debtor is allowed a debt moratorium to settle its debts. Such proceedings can only be launched on the debtor's request. They cannot be launched if the undertaking involved is in compulsory liquidation. If no bankruptcy agreement is reached, the proceedings are automatically converted into compulsory liquidation. Liquidation proceedings, on the other hand, do not aim to reorganise the debtor party's business effectively, but to terminate it with no legal successor during which efforts are made to satisfy the claims of as large part of the creditors as possible. Both the debtor and the creditor may initiate the launch of liquidation proceedings (Piller, 2013).

The importance of corporate bankruptcy prediction has been recognised more and more lately. Most bank crises in the Japanese and Scandinavian banking systems were caused by the bankruptcy of credited companies, which shed light to the importance of analysing customers' survival ability as part of the lending process. Accordingly, banks are the most important users of bankruptcy prediction models, nevertheless, they may also be important for accounting enterprises or even for bonds assessment agencies (Virág, 2004). There are two basic types of corporate bankruptcy prediction models. One involves models based on mathematical statistics, while the other includes methods based on simulation experiments and machine learning (Shi–Li, 2019).

Analysing the survival ability and creditworthiness of enterprises has been a longstanding topic among economists, but there were no highly developed statistical methods that could have allowed for the effective prediction of company bankruptcies at the beginning of the 20th century. At that time, different indicators of surviving and bankrupt companies were tried to be compared and conclusions to be drawn regarding the solvency of the enterprises (Kristóf–Virág, 2019). Those statistical methods neglected the use of any statistical method; researchers simply tried to reveal differences (Fitzpatrick, 1932).

The first modern corporate bankruptcy prediction model was created by Beaver in 1960. His work was based on single-variable discriminant analysis. Using the method, you analyse a single financial indicator to decide if a given company should be categorised as insolvent or a survivor. Compared to earlier examples, the model had outstanding success, since it could categorise enterprises with an accuracy of 90 percent (Beaver, 1966). A downside of the single-variable discriminant analysis, however, is that it can often lead to contradictory results, i.e., one financial indicator will predict survival while another will warn about the risk of bankruptcy (Virág, 2004).

Edward I. Altman published his model in 1968, which was based on multi-variable analysis. He examined a total of 22 financial indicators to build his model using five of them in the end for prediction. It is a linear function analysis where five variables weighted with objective ratios are added to provide a value "Z". The value of "Z" compared to a pre-set cut-off point will decide if an enterprise is a survivor or it will become insolvent (Altman, 1968). The model had a ranking accuracy of 95 percent and has been used to this day in research mainly as a basis for comparison (Ágoston, 2022).

Multi-variable discriminant analysis has been a trailblazer in the field of corporate bankruptcy prediction as its application has led to non-contradictory results. Nevertheless, some problems have also arisen, namely the variables it used had to be statistically independent, but financial indicators often display multicollinearity, which fails to meet the above condition. Another important requirement is that indicators should follow normal distribution. By not requiring normal distribution of the variables, corporate bankruptcy prediction based on logistic regression solves the problem. Accordingly, the method of maximum likelihood is used to fit a function to the observations (Virág–Kristóf, 2006). It was Ohlson who has first used logistic regression for corporate bankruptcy prediction. His model is regarded to be a pioneering effort, because he has been the first to demonstrate a negative relationship between company size and insolvency. Mihalovic (2016) compared logistic regression and multi-variable discriminant analysis. He built a model for each method, then compared their effectiveness both in terms of ranking accuracy and cumulative ranking accuracies (ROC curve). He has found the model based on logistic regression had a higher hit rate for categorising enterprises than that based on multi-variable discriminant analysis.

The use of decision trees was the next milestone in the development history of corporate bankruptcy prediction. Frydman, Altman and Kao were the first to use decision trees in 1985. Using decision trees is extremely popular, as you do not need to meet the statistical requirements discussed above (Kristóf–Virág, 2019). Recursive partitioning is a frequently applied method based on a decision tree. It works on single-variable separation cutting the data in two at every step to build the branches of the tree. The initial data series is a pattern where you know which companies belong to the solvent and which ones to the insolvent category. By the method, variables are examined one by one, and the tree is built along the variables having the highest dividing value so that the resulting categories should be as homogeneous as possible. Categorising the data from the aspect of the dependent variable, the method tries to minimise variance within the groups and maximise it among the different groups (Virág–Kristóf, 2006). Another highly popular method based on decision trees is automatic interaction detection based on the chi-square (CHAID). In the process, the value stock of an explanatory variable is broken down into intervals. Next, it analyses the bins in pairs to decide whether the bins and the categories of the companies in them (solvent or insolvent) are independent of each other. If they are, the two bins will be combined. The process is continued until statistically not independent bins remain only. As a result of the process, the value stock of explanatory variables is broken into bins (Nyitrai ,2017).

The application of AI, particularly of neural networks started to be used in corporate bankruptcy prediction in the 1990s. Neural networks consist of interconnected neurons and can learn in contrast to the methods discussed earlier. The way the neurons are connected varies in each network. Neural networks learn through examples; professionally trained networks can be used for prediction on different data (Kristóf, 2005).

According to Du Jardin (2010), the total number of methods used for corporate bankruptcy prediction exceeds 50. As a result of the high number of applicable models, researchers strive to perfect the existing procedures rather than designing new ones (Nyitrai, 2014). It can happen from several directions, partly by increasing the group of explanatory variables and partly by offering the highest possible use of the information gained from them (Nyitrai, 2017). Using dynamic variables is an effort made to achieve the latter goal.

Researchers used to employ static models for corporate bankruptcy prediction for a long time. In other words, they tried to draw conclusions about the future of an undertaking from its status observed at a single point in time. Insolvency, however, is not a sudden occurrence, it is rather a process in time. Therefore, taking into consideration the dynamics of trends has become an important new direction of research in the field of corporate bankruptcy prediction (Nyitrai, 2017).

Chen et al (2013), criticised the static nature of corporate bankruptcy prediction. They believe that the data gained from time series fail to consider that enterprise operations are processes. They draw conclusions about the future of an enterprise on the basis of a situation observed at a given point in time, i.e., when the balance sheet is prepared, which means a lot of significant information is lost.

According to Nwogugu (2007), companies do not go bankrupt because of a sudden occurrence but at the end of a lengthy process in time, which should be considered when bankruptcies are modelled. Niklis et al (2014) are of the opinion that considering the trends of indicators will be an important area of research in future.

There are two dominant views about the dynamisation of models. According to one, the magnitude of the changes in the values of financial indicators from one year to the next should be applied as an explanatory variable. According to the other, the data of earlier business years rather than simply those of year $t - 1$ immediately preceding bankruptcy should be used to design the explanatory variables (Nyitrai–Virág, 2007).

3 THE SAMPLE ANALYSED AND THE METHOD APPLIED

The sample used in this study is a data series comprising 1527 building companies in total. Of them, 1188 are operational while 339 are under liquidation proceedings. It also illustrates that properly operating healthy companies represent a higher proportion of the economy than those in bankruptcy. The data have been taken from the Crefort data base. The building industry has been selected partly because it plays an important part in the national economy of Hungary and partly because some estimations suggest the rate of enterprises in a difficult position is over 5 percent in it. Therefore, barriers to market entry as well as business risk is higher than average in the industry (Hegedüs 2023). Another argument for analysing the building industry is that it is particularly affected by debt gridlock (Limpek et al, 2016).

3.1 The sample analysed

The research is based on the data of financial reports submitted between 2014 and 2018. The reason for that was to avoid the distorting effect caused by the Covid epidemic. Because of lockdowns during the epidemic, enterprises went bankrupt that had not seem to be in danger in terms of their finances earlier (Boratynska, 2021). To avoid such distortion, reports from the Covid period have not been included in the sample. The sample includes currently operational enterprises that employed at least 5 people in the period reviewed and have been operating for at least 5 years since 2013, i.e., the date of the study. The reason for that is that even well-operating solvent enterprises may resemble at-risk companies in the first few years of their operations in terms of their financial structure, which could have a distorting effect on the study (du Jardin, 2010). A total of 1188 operational companies were included in the sample.

The sample also included 339 enterprises currently under liquidation proceedings. The financial reports of those enterprises were available for at least 3 years preceding their bankruptcy.

Companies having the classification code of economic activities (TEÁOR) No 4120 had the highest proportion in the sample. The code denotes the economic activity of *construction of residential and non-residential buildings*. It was the main economic activity of 39 percent of the busted companies. The second place was taken up by enterprises with the management of building projects as their main economic activity, while the third involved companies engaged in plumbing, gas, heating and air conditioning assembly. Simple random sampling was used both for the learning and the test sample.

3.2 The method applied

The method of logistic regression has been selected for the study since it is still highly popular among banks as they analyse the survival chances of their customers (Rajka–Pollák, 2024). Logistic regression can be used well for explanatory variables and the probability of binary replies. The result variable is a dummy variable representing solvent or insolvent categories for corporate bankruptcy prediction. Using the method you do not need to make preliminary assumptions about the survival or bankruptcy of an enterprise (Kim et al, 2021). During the process, maximum likelihood is used to fit a logistic regression function to the observations. Using the method, coefficients best fitting to the model are searched (Ágoston, 2022). Weighting the independent variables a value “Z” is received, which expresses the probability of a company going bust (Virág–Kristóf, 2006).

The formula of the logistic regression is as follows (Virág–Kristóf, 2006):

$$\Pr(solvent) = \frac{e^z}{1+e^z} = \frac{e^{ß_0+\sum ß_j Z_j}}{1+e^{ß_0+\sum ß_j Z_j}} \tag{1}$$

where "Pr" is the probability of bankruptcy, „$ß_j$" are regression coefficients, „Z_j" are the independent variables.

A great advantage of the model is it does not require normal distribution of the variables, or matching covariance matrices in the two categories. For applying the method, the number of the variables should be reasonably reduced, therefore, the best parameters must be identified in a multi-step process (Székelyi–Barna, 2002). It is most often performed by backward elimination. The less significant variables of the model are omitted one by one. The regression coefficients and p-values are always recalculated after a variable is omitted until suitably significant variables remain only. The final model is built considering collinearity, significance and ranking accuracy at the same time. After defining the regression parameters, the cut-off value must be identified. It is the value of the dependent variable of the function. Comparing the enterprises to the value you can decide if a company should be categorised as solvent or insolvent (Virág–Kristóf, 2006).

A disadvantage of the procedure is its sensitivity to outlier values, which is one of the features of financial indicators and is quite typical of bankrupt companies. Therefore, the outliers of the database must be managed prior to the start of the study (Nyitrai, 2017). No consensus seems to exist in the professional literature on what can be considered outlier data. Quite often, the rule of thumb is applied to define outlier values, i.e., values beyond standard deviation are considered outliers. However, the problem with this approach is that after handling outliers, the standard deviation of the variables changes, so when calculating with the new standard deviation, indicators that were not previously considered outliers are also classified as outliers.

Review using the newly defined deviations must be continued until you receive no more outlying values after the standard deviation changed (Nyitrai–Virág, 2017). Because of the above, a built-in function of SPSS was used in the study to define outliers related to the sample.

However, there is no uniformly accepted model for the management of outliers, either. Two procedures are applied most frequently. Outliers can be substituted with the nearest non-outlier (Nyitrai–Virág, 2017). By another approach, observations including outliers are omitted from the sample (Nyitrai, 2017). In this study, substitution of outliers with the nearest non-outliers has been applied, since it has been found effective for the prediction of the bankruptcy of Hungarian enterprises in earlier studies (Szántó, 2023).

4 THE MODELS

Table 1 presents the indicators used and the methodology of their calculation. Financial indicators most frequently used in the professional literature have been included in the models.

Table 1
Indicators used and the methodology of their calculation

	Indicator	Methodology of calculation
X_1	Liquidity rate	Current assets / Short-term liabilities
X_2	Liquidity flash rate	(Current assets – Stocks) / Short-term labilities
X_3	Cash flow / Liabilities	(Profit after tax + Depreciation) / Liabilities
X_4	Cash flow / Short-term liabilities	(Profit after tax + Depreciation) / Short-term liabilities
X_5	Capital adequacy	(Fixed assets + Stocks) / Equity
X_6	Current assets ratio	Current assets / Balance sheet total
X_7	Assets turnover rate	Net sales / Balance sheet total
X_8	Stocks turnover rate	Net sales / Stocks
X_9	Turnover time of liabilities	Liabilities / Net sales
X_{10}	Indebtedness	Liabilities / Balance sheet total
X_{11}	Equity ratio	Equity / Balance sheet total
X_{12}	Return on equity (ROE)	Profit after tax / Equity
X_{13}	Creditworthiness	Liabilities / Equity
X_{14}	Return on sales (ROS)	Profit after tax / Nat sales
X_{15}	Return on assets (ROA)	Profit after tax / Balance sheet total
X_{16}	Receivables / Short-term liabilities	Receivables / Short-term liabilities
X_{17}	Ratio of net working capital	(Current assets – Short-term liabilities) / Balance sheet total
X_{18}	Company size	Natural logarithm of assets
X_{19}	Ratio of fixed assets covered by long-term liabilities	Long-term liabilities / Fixed assets

Source: own design

It has generally been found that variables describing borrowers' behaviour function best to define the probability of bankruptcy; they provide higher accuracy than, for instance, financial indicators. Further, it is a weakness of accounting data that balance sheet and profit-and-loss statement figures can be manipulated under certain conditions (Cziglerné, 2020). Unfortunately, there are no databases available to observe borrowers' behaviour (Mikolasek, 2018). As a result, the application of financial indicators is typical in the practice of corporate bankruptcy prediction, as the data of financial reports are public and can be assessed by all. Scoring systems are based on objective factors and they cover all areas of company operations providing in that way a comprehensive picture of the business (Zéman et al, 2018). Thus, in accordance with the experience of the professional literature, financial indicators only have been used in this study.

The application of logistic regression may be hindered by the multicollinearity of the variables. So, it must be analysed prior to building a model (Kristóf, 2005). Variance inflation factor (VIF) was used to eliminate the multicollinearity of variables. The VIF value of a variable comes from the diagonal value corresponding to the inverse of the correlation matrix. It is an estimation of how much the variance of regression coefficients increases due to multicollinearity (Vörösmarty–Dobos, 2020). There is no consensus in the professional literature as to what VIF value is the starting point denoting multicollinearity. 5 is the most frequently used limit value, thus, no variable with a VIF value over 5 has been included in the final model.

4.1 Model with static variables only

To build the model, 19 financial indicators mentioned above were used. The variables were selected using the Wald backward elimination method, the entry criterion was set at 5% and exit at 10%. The programme has found 4 variables to be significant, i.e., Liquidity rate, Capital adequacy, Turnover rate of receivables and Company size - they have been included in the final model *(Table 2)*.

Table 2
Model including static variables only

	B	S. E.	Wald	df	Sig.	Exp(B)
X_1	–0.001	0.001	1.074	1	0.003	0.999
X_5	–0.071	0.026	7.617	1	0.005	0.931
X_9	0.001	0.000	22.992	1	0.000	1.001
X_{18}	–0.712	0.049	208.291	1	0.000	0.491
Constant	11.518	0.883	170.022	1	0.000	100529.303

Source: own design

The model with static variables only can be written as the following equation:

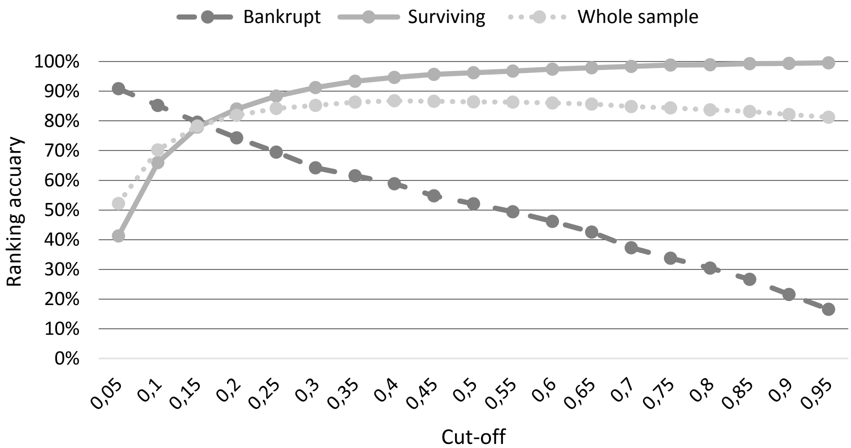
$$Z=\frac{e^{11,518+(-0,001\times X_1)+(-0,071\times X_5)+(0,001\times X_9)+(-0,712\times X_{18})}}{1+e^{11,518+(-0,001\times X_1)+(-0,071\times X_5)+(0,001\times X_9)+(-0,712\times X_{18})}} \quad (2),$$

where:
X_1= Liquidity rate
X_5= Capital adequacy
X_9= Turnover rate of receivables
X_{18}= Company size

To optimise cut-off value, it has been considered that the sample analysed includes a higher rate of solvent enterprises than insolvent ones - similarly to the economy as a total - so it is not enough to minimise the total error rate, as it could lead to a high proportion of Type 1 errors. Type 1 errors occur if a company must be categorised as a survivor by the model, but it actually goes bankrupt. Type 1 and Type 2 errors have different kinds of hidden relative costs. Banks often use corporate bankruptcy prediction models to make lending decisions (Nyitrai–Virág, 2017). In the event of a Type 1 error, a bank may categorise a bankrupt company to be solvent, i.e., if it is granted a loan, the bank may lose its capital outstanding and potential interest income. In the event of a Type 2 error, a solvent debtor is categorised insolvent erroneously. Although the bank will make a loss in both cases, a Type 1 error can cause much more damage (Zavgren, 1985).

Maximum ranking accuracy for the whole sample would be at cut-off value 0.4 *(Figure 1)*. Then 94.7% of surviving companies and 58.8% of insolvent ones would be categorised correctly. However, due to the high rate of Type 1 errors, that cut-off value cannot be considered optimal.

Figure 1
Ranking accuracy of the static model at different cut-off values



Source: own design

To define the optimal cut-off value, a value was selected the application of which can provide sufficiently high-ranking accuracy while Tape 1 errors are kept low. The value is 0.175 for a model with static variables only. Then the model categorises 81.40% of surviving enterprises and 78.17% of bankrupt enterprises accurately; ranking accuracy for the whole sample is 80.68 *(Table 3)*. The ratio of Type 1 errors, i.e., when the model erroneously categorised a bankrupt company to be a survivor is 4.85%, the ratio of Type 2 errors is 14.47%. You can see the ranking accuracy of the model built after managing the outliers is higher than that of the model received without managing the outliers.

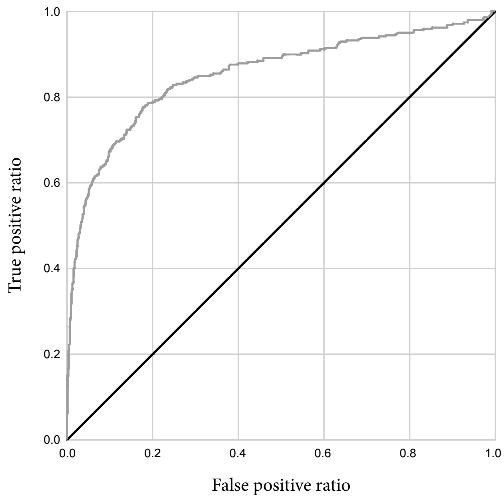
Table 3
Ranking accuracy of the static model

	Categorised accurately, pc	Categorised inaccurately, pc	Accuracy %
Survivors	967	221	81.40%
Bankrupt	265	74	78.17%
Total	1232	295	80.68%

Source: own design

Figure 2 is the cumulative ranking accuracy (ROC curve); in this case the area below the ROC curve is 85.2 percent.

Figure 2
ROC curve of the static model



Source: own design

4.2 Model including both static and dynamic variables

The model discussed above was static, i.e., it failed to consider company operations as processes; it tried to draw conclusions about the future of an enterprise on its position as observed at a given point in time. Bankruptcy prediction models are most often prepared based on the financial reports of the preceding business year. Insolvency, however, is not a sudden occurrence, it is rather a process in time. Therefore, taking into consideration the dynamics of financial indicators has become an important new direction of research in the field of bankruptcy prediction (Nyitrai, 2017). There are two dominant schools of thought around the dynamisation of models. According to one, the magnitude of the changes in the values of financial indicators from one year to the next should be applied as explanatory variable. According to the other, the data of earlier business years rather than simply those of year t–1 immediately preceding bankruptcy should be used to design the explanatory variables (Nyitrai–Virág, 2006). Series of financial indicators can be the source of variables generated in several ways to capture how a company's latest financial indicator relates to the values of the same indicator in earlier years. In the study, the formula proposed by Nyitrai (2017) has been applied, which is a combination of the two methods discussed above (Nyitrai, 2017):

$$\frac{X_{i,t-1} - X_{i,\min_{[t-2;t-n]}}}{X_{i,\max_{[t-2;t-n]}} - X_{i,\min_{[t-2;t-n]}}} \tag{3}$$

The resulting value will demonstrate how the company's i–th financial indicator observed in the last business year relates to the values of the same indicator observed until the penultimate year (Nyitrai, 2017).

In this case, too, outliers were substituted with the nearest non-outlier values. In total, 38 indicators were analysed for building the model. The indicators comprised the static and the dynamic versions of the 19 financial indicators presented earlier, the dynamic versions resulted from applying the formula proposed by Nyitrai (2017). The variables were selected using the Wald backward elimination method, the entry criterion was set at 5% and exit at 10%. 7 variables have proved to be significant following selection. As for the dynamic version, the Cash flow / Liabilities ratio and the dynamic version of Equity ratio have been included in the final model. Out of static indicators, the variables Cash flow / Short-term liabilities, Capital adequacy, the Current assets ratio, Return-on-equity and Company size have been included in the final model *(Table 4)*.

Figure 4
Final dynamic model

	B	S.E.	Wald	df	Sig.	Exp(B)
D_3	–0.121	0.024	25.052	1	0.000	0.886
D_{11}	–0.047	0.009	25.245	1	0.000	0.954
X_4	–0.008	0.005	2.861	1	0.001	0.992
X_5	–0.038	0.021	3.298	1	0.009	0.963
X_6	0.841	0.362	5.401	1	0.020	2.319
X_{12}	0.040	0.023	3.211	1	0.013	1.041
X_{18}	–0.697	0.056	155.208	1	0.000	0.498
Constant	10.484	1.081	94.130	1	0.000	35748.067

Source: own design

You can find a strong correlation between static versions of the cash-flow / liabilities ratio and the cash flow / short-term liabilities ratio as a result of the low rate of the long-term liabilities in the sample. Since the final model includes both the dynamic version of the cash flow / liabilities ratio and the static form of the variable cash flow / short-term liabilities, it must be emphasised that no multicollinearity based on the variance inflation factor (VIF) values has been found among the variables in the model *(Table 5)*. The VIF value is an indicator that is suitable to detect multicollinearity, as it is an estimation of how much the variance of regression coefficients increases due to multicollinearity. However, there is no consensus in the professional literature as to what VIF value is the starting point denoting multicollinearity. 5 is the limit value applied most frequently, so VIF values were compared to it in the study (Vörösmarty–Dobos, 2020).

Table 5
Multicollinearity between the variables of the dynamic model

	Collinearity Statistics VIF
D_3	1.071
D_{11}	1.475
X_4	1.098
X_5	1.000
X_6	1.664
X_{12}	1.001
X_{18}	1.153

Source: own design

Accordingly, the final dynamic model can be written as follows:

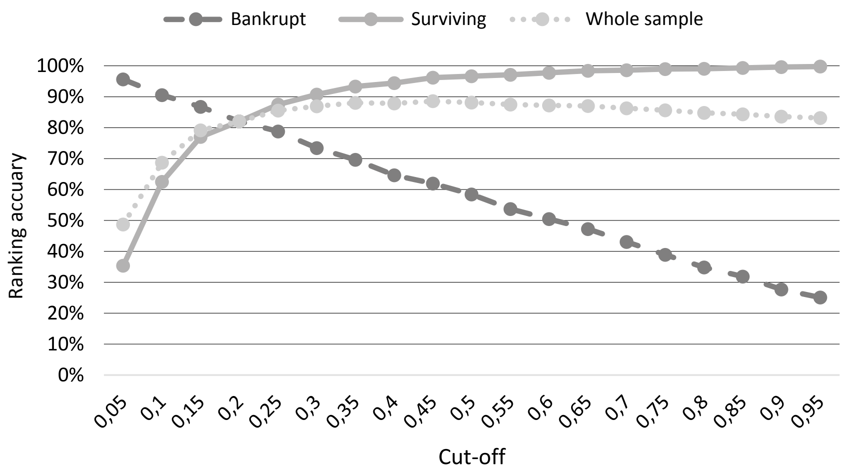
$$Z=\frac{e^{10,484+(-0,121\times D_3)+(-0,047\times D_{11})+(-0,008\times X_4)+(-0,038\times X_5)+0,841\times X_6+0,04\times X_{12}+(-0,697\times X_{18})}}{1+e^{10,484+(-0,121\times D_3)+(-0,047\times D_{11})+(-0,008\times X_4)+(-0,038\times X_5)+0,841\times X_6+0,04\times X_{12}+(-0,697\times X_{18})}} \quad (4)$$

D_3= Dynamic version of Cash flow / Liabilities
D_{11}= Dynamic version of Capital adequacy ratio
X_4= Cash flow / Short-term liabilities
X_5= Capital adequacy
X_6= Current assets ratio
X_{12}= Return on Equity
X_{18}= Company size

Company size has proved to be significant in the case of both models and it had a negative sign in both. It shows the higher a company's balance sheet is, the lower is the probability of it going bankrupt.

To optimise the cut-off value, it has been considered that the sample analysed includes a higher rate of solvent enterprises than insolvent ones - similarly to the economy as a total – so it is not enough to minimise the total error rate, as it could result in a high proportion of Type 1 errors. Maximum ranking accuracy for the whole sample would be achieved if cut-off value were 0.45. In that case, the model would run on the whole sample with a hit rate of 88.54%, but the ratio of companies correctly categorised as bankrupt would only be 61.95% *(Figure 3).*

Figure 3
Ranking accuracy of the model also including dynamic variables at different cut-off values



Source: own design

The optimum value of cut-off is 0.205. Then the model has ranking accuracy of 81.71% for bankrupt companies and categorises surviving ones at a hit rate of 82.91%, so the total ranking accuracy is 82.65 *(Table 6)*. The ratio of Type 1 errors, i.e., when the model erroneously categorised a bankrupt company to be a survivor was 4.06%, the ratio of Type 2 errors was 13.29%. The ranking accuracy of the model also including dynamic variables is higher than that of the model comprising static variables only received from managing outliers with substitution. The respective ranking accuracy figures were 81.4% for survivors and 78.11% for bankrupt companies.

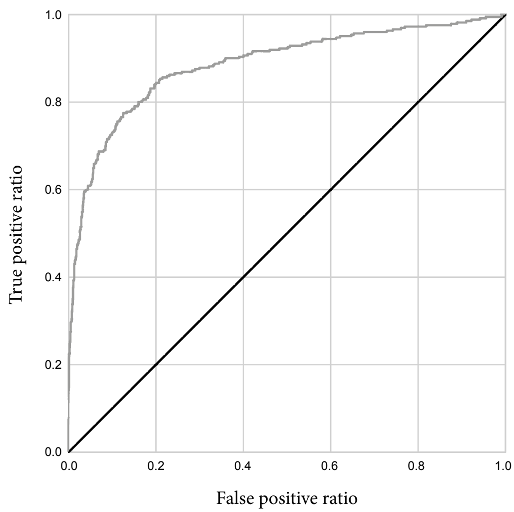
Table 6
Ranking accuracy of the dynamic model

	Categorised accurately, pc	Categorised inaccurately, pc	Accuracy %
Survivors	985	203	82.91%
Bankrupt	277	62	81.71%
Total	1262	265	82.65%

Source: own design

The dynamic model also looks more advantageous by comparing the areas below the ROC curve. The area below the ROC curve is 88.7 percent for the dynamic model, which is higher than the values received from the static model only *(Figure 4)*.

Figure 4
ROC curve of the dynamic model



Source: own design

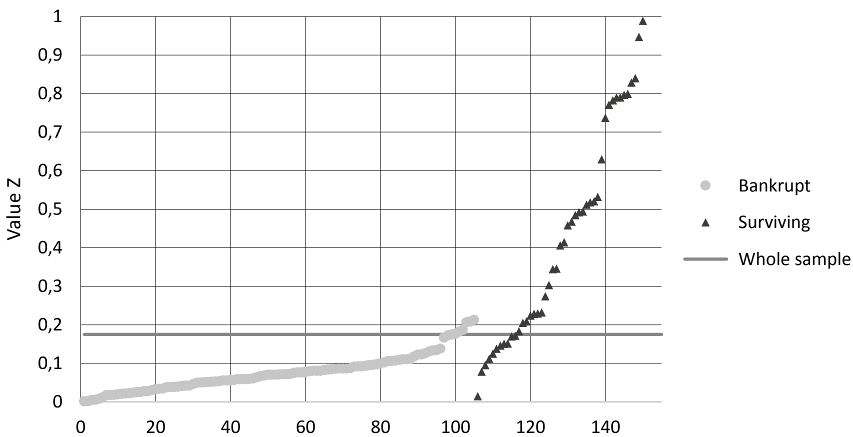
5 COMPARISON OF THE MODELS ON A TEST SAMPLE

The accuracy of the final corporate bankruptcy prediction models has been checked on an independent test sample. The sample comprised 150 building companies, 45 of which are under liquidation proceedings while 105 of them are operational.

5.1 Ranking accuracy of the model with static variables only on the test sample

Figure 5 illustrates the ranking accuracy of the model built with static variables only in the year preceding their going bankrupt. Its total ranking accuracy is 88.6%, 94.29% of surviving companies and 75.56% of those going bankrupt have been correctly categorised. The ratio of Type 1 errors is 7.33% while that of Type 2 ones is 4%.

Figure 5
Ranking accuracy of the model with static variables on the test sample

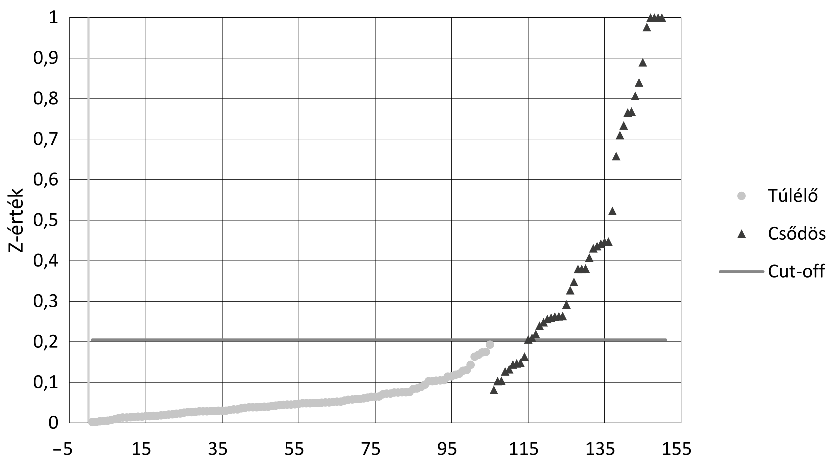


Source: own design

5.2 Ranking accuracy of the model also including dynamic variables on the test sample

One year prior to the launch of liquidation proceedings, the model also including dynamic variables categorised 75.56% of companies on the brink of bankruptcy correctly, into the insolvent category. 100% of surviving companies were correctly categorised, so the total ranking accuracy is 92.6% (Figure 6). The ratio of Type 1 errors was 7.33% for the test sample, no Type 2 errors occurred as the hit rate of surviving companies was 100%.

Figure 6
Ranking accuracy of the model also including dynamic variables 1-year preceding bankruptcy



Source: own design

The model also includes dynamic variables provided a higher hit rate on both the learning and the test samples than the model having static variables only (*Table 7*).

Table 7
Bankruptcy prediction accuracy of the models built on the learning and test samples

		Categorisation of surviving companies, %	Categorisation of companies on brink of bankruptcy, %	Accuracy of whole sample, %
Static model	Learning sample	81,4	78,17	80,68
	Test sample	94,29	75,56	88,67
Dynamic model	Learning sample	82,91	81,71	82,65
	Test sample	100	75,56	92,9

Source: own design

6 SUMMARY

In this study, corporate bankruptcy prediction models have been designed applying the method of logistic regression. One model comprised static variables only, while the other one included both static and dynamic variables to find out whether the models' ranking accuracy will improve if the trends of financial indicators are analysed during bankruptcy prediction.

The analysed sample comprised of 1527 Hungarian building companies, 1188 of which was operational, while 339 were under liquidation proceedings. The findings have been checked on an independent test sample comprising of 105 surviving companies and 45 bankrupt enterprises. As a limitation of the research, it should be noted that simple random sampling was used both for the learning and the test samples, i.e., the samples were not representative.

In the case of both models, company size had a negative sign among the explanatory variables, which suggests the probability of a company going bust is reduced if its balance sheet total increases.

The findings illustrate the area below the ROC curve is larger for models also including dynamic variables. Reviewing ranking accuracy also indicates the model also having dynamic variables has reached a higher hit rate. The findings suggest banks could be advised to include the analysis of the trends of indicators in their lending practices. There is an important difference here between lending to start-ups or to enterprises with a long operational track record. Obviously, you cannot study the earlier values of indicators with a start-up company; for this reason, the spread of the practice may further deepen (restrict) the options of start-ups to have access to resources. Therefore, the method should rather be employed for companies with a long track record.

LITERATURE

Altman, E. I. (1968): Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x.

Ágoston, N. (2022): Mesterséges intelligencia és gépi tanulási módszerek a vállalati fizetésképtelenség becslésére. *Statisztikai Szemle,* 100, 586–609. https://doi.org/10.20311/stat2022.6.hu0584.

Beaver, W. H. (1966): Financial ratios as predictors of failure. *Journal of Accounting Research*, 4, 71–111. https://doi.org/10.2307/2490171.

Boratynska, K. (2021): A New Approach for Risk of Corporate Bankruptcy Assessment during the COVID-19 Pandemic. *Journal of Risk and Financial Management,* 14, 590–604. https://doi.org/10.3390/jrfm14120590.

Chen, N. – Riberio, B. – Viera, A. – Chen, A. (2013): Clustering and visualization of bankruptcy trajectory using self-organizing map. *Expert Systems with Applications,* 40(1), 385–393. http://dx.doi.org/10.1016/j.eswa.2012.07.047.

Constand, L. R. – Yazdipour, R. (2011): Firm failure prediction models: a critique and a review of recent developments. In: Yazdipour, R. (ed.), *Advances in Enterpreneurial Finance: With Applicantions from behavioral Finance and Economics* (185–204), New York: Springer Science and Business Media. https://doi.org/10.1007/978-1-4419-7527-0_10.

Du Jardin, P. (2010): Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy. *Neurocomputing*, 73, 2047–2060. https://doi.org/10.1016/j.neucom.2009.11.034.

Cziglerné Erb, E. (2020): A reziduálisjövedelem-modell újbóli megjelenése a vállalatok és a beruházási projektek értékelésében. *Pénzügyi Szemle*, 3, 430–442. https://doi.org/10.35551/PFQ_2020_3_7.

Dun & Bradstreet (1978): The Business Failure Record. *New York: Dun & Bradstreet Inc*.

Fitzpatrick, P. (1932): A comparison of the ratios of successful industrial enterprises with those of failed companies. *Certified Public Accountant*, 6, 727–731.

Greenwald, D. (1973): The McGraw-Hill Dictionary of Modern Economics: *A Handbook of Terms and Organizations*. New York: McGraw-Hill Book Company.

Hegedűs, Sz. (2023): A nehéz helyzet kialakulásának és magyarázó változóinak vizsgálata a magyar kkv-szektorban. *Gazdaság és Pénzügy*, 10, 57–79. https://doi.org/10.33926/gp.2023.1.4.

Kim, H. – Cho, H. – Ryu, D. (2022): Corporate Bankruptcy Prediction Using Machine Learning Methodologies with a Focus on Sequential Data. *Computational Economics,* 59, 1231–1249. https://doi.org/10.1007/s10614-021-10126-5.

Központi Statisztikai Hivatal (2022): Helyzetkép az építőiparról, 2022.

Kristóf, T. (2005): A csődelőrejelzés sokváltozós statisztikai módszerei és empirikus vizsgálata. *Statisztikai Szemle*, 9, 841–863.

Kristóf, T. – Virág, M. (2019): A csődelőrejelzés fejlődéstörténete Magyarországon. *Vezetéstudomány*, 12, 62–73. https://doi.org/10.14267/veztud.2019.12.06.

Limpek, Á. – Kosztopulosz, A. – Balogh, P. (2016): Késedelmes fizetés, tartozási láncok – A Dél-Alföld régió kis- és középvállalkozásainak pénzügyi kultúrája. *Statisztikai Szemle*, 94, 365–387. https://doi.org/10.20311/stat2016.04.hu0365.

Márton, A. – Fiáth, A. – Kristóf, T. (2023): Állami energiavállalatok pénzügyi teljesítménye Magyarországon a koronavírus-járvány előtt és alatt. *Közgazdasági Szemle*, 70, 1057–1076. https://doi.org/10.18414/KSZ.2023.10.1057.

Mihalovic, M. (2016): Performance comparism of multiple discriminant analysis and logit models in bankruptcy prediction. *Economics & Sociology,* 9, 101–118 https://doi.org/10.14254/2071-789x.2016/9-4/6.

Mikolasek, A. (2018): A hitelkockázati modellek alkalmazásának néhány problémája. *Gazdaság és Pénzügy*, 3, 248–257. https://bankszovetseg.hu/Public/gep/2018/imp%20jav%20248-257Mikolasek%20Andrasuj.pdf.

Niklis, D. – Doumpos, M. – Zopounidis, C. (2014): Combining market and accounting-based models for credit scoring using a classification scheme based on support vector machines. *Applied Mathematics and Computation*, 234, 69–81. http://dx.doi.org/10.1016/j.amc.2014.02.028.

Nwogugu, M. (2007): Decision-making, risk and corporate governance: A critique of methodological issues in bankruptcy/recovery prediction models. *Applied Mathematics and Computation*, 185(1), 178–196. http://dx.doi.org/10.1016/j.amc.2005.11.178.

Nyitrai, T. (2014): Növelhető-e a csőd-előrejelző modellek előre jelző képessége az új klasszifikációs módszerek nélkül? *Közgazdasági Szemle*, 5, 566–585.

Nyitrai, T. (2017) Stock és flow típusú számviteli adatok alkalmazása a csődelőrejelző modellekben. *Vezetéstudomány*, 48, 68–77. https://doi.org/10.14267/veztud.2017.09.07 .

Nyitrai, T. – Virág, M. (2017): A pénzügyi mutatók időbeli tendenciájának figyelembevétele logisztikus regresszióra épülő csődelőrejelző modellekben. *Statisztikai Szemle,* 1, 5–28. https://doi.org/10.20311/stat2017.01.hu0005 .

Pálinkó, É. – Svoób, Á. (2016): A vállalati csőd bekövetkezésének fő okai és a csődhöz vezető folyamat. *Pénzügyi Szemle*, 4, 528–543.

Piller, Zs. (2013): A fizetésképtelenségi eljárások illeszkedési módjai nemzetközi összehasonlításban. *Pénzügyi Szemle,* 2, 151–164.

Rajka, L. – Pollák, Z.: (2024) Mesterséges intelligencia a hitelkockázati modelleknél, avagy mire képesek a gépi tanulási algoritmusok a hagyományos modellekhez képest. *Gazdaság és Pénzügy*, 11, 246–273. https://doi.org/10.33926/gp.2024.3.1.

Sharma S. – Mahajan, V. (1980): Early Warning Indicators of Business Failure. *Journal of Marketing*, 44(4), 80–89. https://doi.org/10.1177/002224298004400412.

Shi, Y. – Li, X. (2019/a): A bibliometric study on intelligent techniques of bankruptcy prediction for corporate firms. *Heliyon*, 5.

Shi, Y. – Li, X. (2019/b): An overview of bankruptcy prediction models for corporate firms: A systematic literature review. *Intangible Capital*, 15, 114–127.

Szántó, T. K. (2023): Kiugró értékek kezelése logisztikus regresszión alapuló csődelőrejelzési modellek esetén. *Pénzügyi Szemle*, 69, 91–106. https://doi.org/10.35551/PFQ_2023_3_5.

Székelyi, M. – Barna, I. (2002): Túlélőkészlet az SPSS-hez. Budapest: *Typotex Kiadó.*

Virág, M (2004): A csődmodellek jellegzetességei és története. *Vezetéstudomány*, 10, 24–32.

Virág, M. – Kristóf, T. (2006): Iparági rátákon alapuló csődelőrejelzés sokváltozós statisztikai módszerekkel. *Vezetéstudomány*, 37, 25–35. https://doi.org/10.14267/veztud.2006.01.04 .

Vörösmarty, Gy. ¬ Dobos, I. (2020): A vállalatméret hatása a zöldbeszerzési gyakorlatra. *Statisztikai Szemle*, 4, 301–323. https://doi.org/10.20311/stat2020.4.hu0301 .

Zavgren, C. V. (1985): Assessing the vulnerability to failure of American industrial firms: a logistic analysis. *Journal of Business Finance & Accounting*, 12(1), 19–45.

Zéman, Z. – Hegedűs, Sz. – Molnár, P. (2018): Az önkormányzati vállalkozások hitelképességének vizsgálata credit scoring módszerrel. *Pénzügyi Szemle*, 2, 182–200.