UNIVERSITY OF MISKOLC
FACULTY OF ECONOMICS

KITTI FODOR

The modelling of household loan defaults using multivariate statistical methods

Theses of the PhD dissertation

Miskolc
2023.

UNIVERSITY OF MISKOLC
FACULTY OF ECONOMICS
HANTOS ELEMÉR BUSINESS AND REGIONAL SCIENCES DOCTORAL SCHOOL

KITTI FODOR

The modelling of household loan defaults using multivariate statistical methods

Theses of the PhD dissertation

HEAD OF DOCTORAL SCHOOL:
Prof. Dr. Tamás Sikos T.

SUPERVISOR:
Dr. Beatrix Margit Varga

Miskolc
2023.

Table of contents

List of figures

List of tables

1 Justification for the choice of topic

1.1. The relevance of the research and research problem

For both corporate and retail lending, it is important for financial institutions to lend to customers with a low risk of non-repayment. Although institutions have their own credit assessment process, they may still not properly select the customers for whom it is appropriate to provide credit.

While there is a large literature on corporate bankruptcy prediction, there is little literature on the classification of retail borrowers. In the case of the corporate sector, it has been observed that the range of analytical tools used has steadily increased with the development of methodological possibilities and information technology, from the initial univariate analyses to the present day models using neural networks.

The latter statement is particularly true in our country, where there is very little scientific work on this topic. This is true despite the fact that the last crisis was not so long ago, and lending has been booming recently, with many people taking out loans as if there were no tomorrow. Recognising who is becoming a bad debtor is not easy. There are several ways to analyse the data, which may even show different results. From an analytical point of view, it may be an interesting question what data to include in the database, what to consider as affecting factors, which variables are most likely to determine the probability of default.

The aim of my research is to investigate the default of household loans using methods based on multivariate statistical techniques. The different methods and models can help to identify the factors that contribute to someone becoming a defaulting debtor and to find out which multivariate statistical methods produce the best results. The method is expected to produce a model that filters out bad debtors as efficiently as possible.

However, the research is complicated by the limited range of data available, which I faced during the research, and it took me a long time to find a partner for the database.

The novelty of the research, as described above, is that in Hungary this research area focuses on the corporate sector, while my research focuses on the household sector.

1.1 Structure and hypotheses of the dissertation

After the introduction, the dissertation includes a literature review, in the course of which I discuss the brief history of lending, when the origins of credit can be traced back to, and what important events have taken place in recent years. Particular attention is given to analysing the lending situation in Hungary and the possible dangers of indebtedness. Furthermore, I summarise briefly what types of credit are currently available and what are the basic concepts that are important to be aware of when taking out a loan. At the end of the second chapter, I look at other factors that may play an important role in lending, highlighting the importance of financial inclusion, culture and resilience, and briefly present the results of my previous research.

In the literature review, I examine the methods used in similar research. In answering this question, I draw on the history of bankruptcy prediction, as a large amount of research has been and is still being focused on this area. For a complete picture, I examine both international and domestic bankruptcy model research. This part of the literature will help to identify which methods I can apply in my analysis, and then I present each method.

The next chapter starts with an introduction of BISZ Zrt. and the database. I have carried out the necessary analyses on the database provided by them. To support my hypotheses, I build several global and several specific models. These models will help me to examine which variables contribute to the increase in the probability of default and to determine which methodology provides the most accurate results.

During the research I formed the following hypotheses:

H1: The information recorded by the KHR can be used to predict the risk of default with a high degree of confidence using the classification procedures used in the bankruptcy prediction.

H2: In corporate bankruptcy prediction, analyses run using neural networks have had the highest classification accuracy (Odom-Sharda, 1990; Olmeda-Fernandez, 1997). I hypothesize that this statement is also true for household loans, so neural networks will also provide the best results for categorizing household loans, outperforming all the methods I have used.

H3: The range of information recorded by the KHR is less than the information requested by banks when applying for a loan (e.g. income data). The inclusion of new variables not recorded by the KHR can improve classification accuracy.

I used the following methods to test hypotheses H1, H2, H3:

- Logistic regression,
- Decision tree,
- Neural network,
- ROC curve,
- Gini coefficient,
- Correlation

H4: The different social and economic situation in different parts of Hungary can be linked to loan defaults.

H5: Different demographics (gender and age) affect non-payment.

I used the following methods to test hypotheses H4, H5:

- Logistic regression,
- Decision tree,
- Neural network,
- Crosstabs

H6: For different types of credit, banks may require different conditions from the customer, which may even affect default. My hypothesis is that if we build the predictive model by loan type, better classification accuracy can be achieved.

I used the following methods to test hypotheses 6:

- Neural network,
- ROC curve,
- Gini coefficient,

2 Theoretical background of the research

The first step of my research was to study the Hungarian and international literature. This is briefly summarised in this chapter. I believe this will contribute to the understanding of my research and its results.

Lending has been part of human culture for thousands of years. The first rules date back to Hammurabi, who laid down rules for lenders and borrowers on stone tablets. (Fekete-Tatay, 2012) In ancient times, non-payment of loans was punishable by serious consequences, including death. In the Middle Ages, the Council of Nicea imposed an interest ban, and this resistance persisted until the Reformation. It was in the 15th century that the idea that money should be circulated for the benefit of the economy was born. In the 1600s and 1800s, loans were mainly granted by wealthy landowners and citizens. In the continental countries, we can speak of lending from the 19th century onwards (Vértesy, 2008).

The Great Depression also hit Hungary hard. After the First World War, many landowners became so indebted that they could not pay their loans, and the land on which their mortgages were based was at risk (Szabó, 2021).By the 1940s, the impersonalisation of this industry was noticeable, but one question remained constant: would the bank grant the applicant a loan or not? In the early 1990s, several banks were affected by bankruptcy, but the beginning of the banking system's stability was delayed until the mid-1990s (Vértesy, 2008).

For a long time, the focus of lending was mainly on large corporate lending, with a significant change only in the late 1990s.By 2004, the proportion of the Hungarian population with credit was in the millions, and banks were competing for lenders, with a steady easing of conditions. Savings have developed in the opposite direction. Even before the crisis, there were indications that the Hungarian population had a poor financial awareness and financial literacy, which could be a dangerous factor in terms of lending, and that the lack of this knowledge could even lead to a debt trap.

This was evidenced by the dramatic 50% increase in the number of non-performing loans in 2004. However, there were some institutions where previous defaults were not a negative factor, specifically targeting clients on the BAR list, pushing them further into debt (Dobák-Sági, 2005). The 2008 crisis caused the franc exchange rate to rise sharply, leading to an increase in repayments to such an extent that creditors could no longer meet them. Finally, foreign currency loans were phased out in 2014.

In 2015, a new product was added to the range of products offered by credit institutions: the Family Housing Allowance (CSOK). Customers and couples with children or who are expecting children can still apply for the non-refundable subsidy and the preferential loan. In mid-2019, the range of preferential loans was extended to include the „Baby Waiting Loan", which can be applied for by married couples planning to have child. The introduction of these loans has also had an impact on the steady growth in domestic lending in recent years.

However, the outbreak of the coronavirus at the end of 2019 opened a new chapter. One of the measures I would highlight is the moratorium on repayments, which could have an impact on defaults. The moratorium was introduced in March 2020 and was initially open to all, but has been tightened over time. In Hungary, more than one and a half million customers have taken advantage of this option.

By the last quarter of 2022, a significant decline in retail credit issuance was observed. In the case of personal loans, 18% fewer loans were issued in the last quarter of 2022 than in the same period of the previous year, a decline of more than HUF 20 billion. And for housing loans, more

than HUF 175 billion fewer loans were issued in the last quarter of 2022 than in the same period of the previous year, a decrease of 54% (MNB, 2023a).

2.1 Hungarian credit situation, indebtedness

In the case of consumer credit, the amount of credit taken out fell steadily after the peak in 2009, but this was followed by a rapid increase after 2016, passing the 2009 peak in 2019 and doubling it by 2021. This development is partly due to the change in the regulation.

Figure 1: Development of consumer credit of the Hungarian population 2001-2021 (billion HUF)
Source: Trading Economics

The credit crunch has left people with many bad memories, which may also be behind the fall in credit up to 2016. These bad memories have led to a growing lack of trust in banks and the financial sector (Lautenschläger, 2015).

One of the biggest risks of taking out a loan is if the borrower cannot repay the amount borrowed. This can result in being charged late interest, being charged debt management fees, or being placed on the KHR list. The most serious consequence, however, is the loss of the property registered as collateral for the loan, which can lead to housing problems and, if the collateral is not sufficient to cover the full amount of the loan, it may be necessary to continue making repayments.

A number of measures have been introduced in recent years to control over-indebtedness.The JTM and HFM, contained in MNB Regulation 32/2014, entered into force on 1 January 2015. In setting these limits, the MNB continuously monitors developments in the credit market and, where necessary, changes the regulations, as has been the case on several occasions. The latest amendment enters into force on 1 July 2023. The background to each change is the inflation environment and wage dynamics in the recent period. Currently, a higher JTM limit is available for net incomes above 500,000 HUF, which will increase to 600,000 HUF after the new rules come into force. This means that people with this income can use up to 60% of their income for loan repayments (MNB, 2023).

2.2 Other factors related to lending that are difficult to measure

In my dissertation, my research in this area has covered financial inclusion, culture and resilience, but here I will briefly focus on financial culture only.

Béres-Huzdik said that lending is an important driver of the economy, which needs to function well for the economy to perform well. However, lending itself is not enough, it is important that borrowers make informed financial decisions. Lending decisions can be influenced by the financial knowledge of the customer, as well as by various subconscious patterns. Financial culture can be a good tool to examine financial literacy, and it can also be important to understand consumer habits, because almost half of our decisions are not conscious (Béres-Huzdik, 2022).

I fully agree with the opinion of the authors, financial literacy is an important factor in lending, and therefore I would like to dedicate some thoughts to this topic. I believe that with better financial literacy we can make better decisions. However, nowadays there is a very wide and complex range of offers on the financial market, and in order to be able to navigate confidently through the wealth of information, it is necessary to have the right financial knowledge (Kovács et al., 2014).

Béres and his co-authors believed that financial personality consists of two parts, and that financial personality "can be used to infer the debtor's personality, behavioural patterns and, from the bank's point of view, whether the loan will be repaid." (Béres et al., 2015, p27.)

Potóczki's article suggests that the level of financial literacy in Hungary is below average. This conclusion was drawn on the basis of international research initiated by the OECD. Hungary ranked 22nd out of 30 countries participating in the research (Potóczki, 2017).

Another international survey by S&P found that 46% of the Hungarian population is financially illiterate. Hungary is in the middle of the European ranking. The graph below shows how the level of financial literacy varies across European countries. Romania had the worst score (Lusardi et al., 2015).

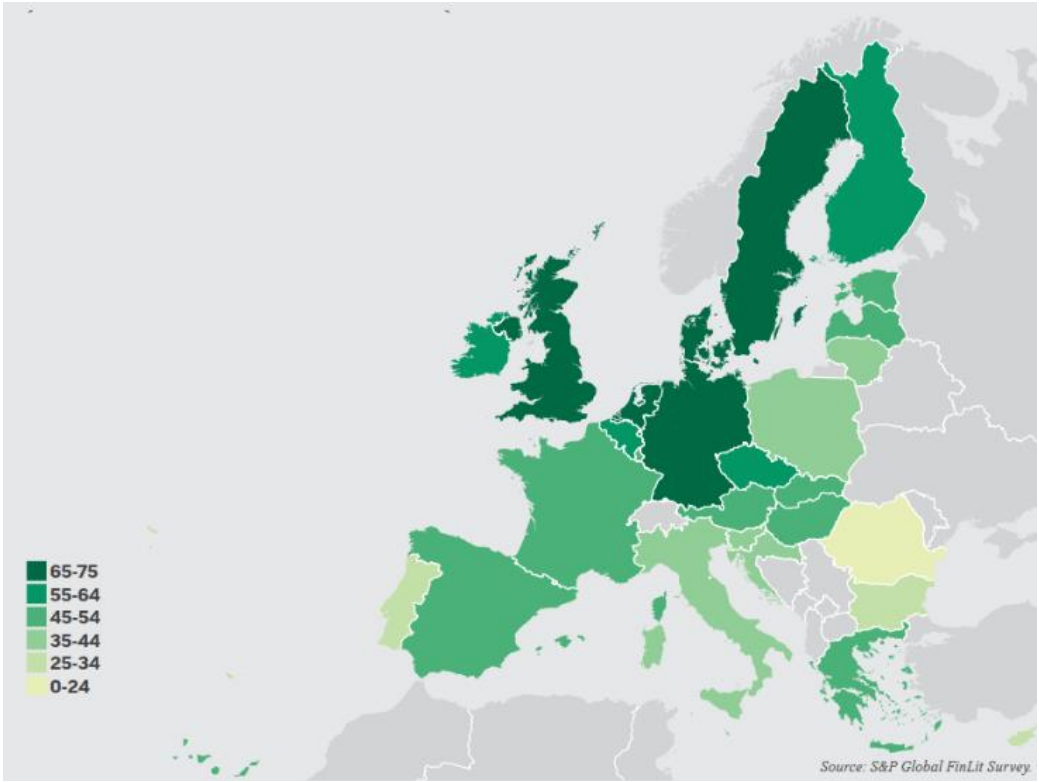

Figure 2: Levels of financial culture in European countries
Source: Financial literacy around the World, Klapper et al., 7., 2015

2.3 International bankruptcy research

Research into bankruptcy forecasting is less than 100 years old, with the first attempts dating back to the 1930s, but the models we use today only appeared in the 1960s. However, models and methods have changed a lot in the intervening period, thanks to advances in analytical capabilities and information technology.

In fact, the initial "models" were not really models yet, the researchers looked for indicators for which bankrupt and surviving companies were significantly different, compared these indicators and tried to find different correlations.

The first univariate analysis was performed by Beaver. He included 158 companies in his analysis, with equal proportions of bankrupt and surviving companies. His method was able to categorise companies with 90% classification accuracy. The disadvantage of the method is that it is a univariate model, so the categorisation is based on a single indicator, so if different indicators result in different classifications, the method cannot handle this. This is one of the reasons why this method has not been widely used (Beaver, 1966; Virág, 2004)

The first real model was created by Altman, who built his model on 5 financial indicators that could predict insolvency with 95% confidence. A few years later, an extended seven-variable model was developed based on this model (Altman, 1968; Virág, 2004). Deakin also used discriminant analysis to predict bankruptcy and tested his results on a 34-item sample. The model had a classification accuracy of 97% (Deakin, 1972) Blum's 1974 model also had a classification accuracy of over 90% (Blum, 1974) Altman's extended version of his five-variable model was developed in 1977, with the new model using a larger sample of 111 items, including 58 surviving firms.

Altman's models were not representative, and the sample included approximately equal proportions of solvent and insolvent companies. The first survey that could be considered representative was conducted by Ohlson. Ohlson was also the first to use logistic regression in bankruptcy prediction models. The sample he studied included 2 163 companies, of which 4.85% were bankrupt. If the P-value calculated in the model exceeds 0.038, the company is considered to be at risk of bankruptcy. The model has an accuracy of approximately 83% (Ohlson, 1980)

The next novelty was the emergence of recursive partitioning algorithms, which dates back to the mid-1980s. The first adopters of this method include Altman, Frydman and Kao. The classification accuracy of the model was 94%, but there was a significant difference in the correct categorisation between surviving and failed firms (Frydman et al., 1985)

The next big step was the emergence of neural networks in the 1990s. The first application of neural networks was by Odom and Sharda. Their model was based on the variables used by Altman in 1968. The sample consisted of 129 companies. For the training sample, the classification was perfect, thus outperforming the results achieved by discriminant analysis. For the test sample, the classification accuracy of 82% significantly exceeded the results achieved by discriminant analysis (Odom-Sharda, 1990)

Tam-Kiang conducted their analyses on banks, with the neural network performing best over a one-year horizon, but logistic regression performing best over a two-year horizon (Virág-Kristóf, 2005).The Coats-Fant pair compared the performance of discriminant analysis with that of the neural network, and came to similar conclusions. In the second half of the 1990s, Olmeda and Fernandez analysed data from Spanish banks. Their research was carried out using all the models mentioned above, of which the neural network proved to be the best, with a classification accuracy of 82.4% (Olmeda-Fernandez, 1997) Zhang-Hu-Patuwo (1999)

compared the neural network with logistic regression. The former achieved a classification accuracy of 88.2% and the latter 78.6% (Zhang et al., 1999)

Overall, of the different methods of analysis, neural networks basically produced the best results.

2.4 Hungarian bankruptcy studies

In Hungary, Law XLIX of 1991 provided the legal framework for the regulation of bankruptcy, thus giving the opportunity to start Hungarian research. The first published model was by Miklós Virág and Ottó Hajdu. For their analysis they used discriminant analysis and logistic regression. Their database consisted of 154 manufacturing companies. There was a half and half ratio of bankrupt and surviving firms in the sample. Discriminant analysis correctly categorized with a classification accuracy of 77.92%. The performance of the logistic regression exceeds that of the discriminant model, reaching 81.8% (Virág-Kristóf, 2005).

In 1996, the authors presented a new model family of bankruptcy models called the Early Warning System. A special feature of these models was that they were designed specifically for each industry and sector. Their analyses were carried out using discriminant analysis on a database of around 10.000. The large global model is accompanied by sectoral models, 30 in number. However, the details of the model and its equations have not been published for reasons of commercial confidentiality.

In later years, the Virág-Kristóf pair experimented with neural networks, using a technique called "backpropagation". The results of other researchers have shown that neural networks give the best results when the sample contains (approximately) the same proportion of surviving and failing companies. Their analysis was carried out on a sample of 156 items. ¾ of the sample represented the training sample. The final model had two hidden layers, the first hidden layer consisted of 6 neurons and the second hidden layer consisted of 5 neurons. The results of their analysis were similar to international experience, which shows that a better performing model can be created using a neural network. The neural network outperformed the discriminant analysis by 8.6 percentage points and the logistic regression by 4.7 percentage points (Virág-Kristóf, 2005).

2.5 Methods of bankruptcy prediction

Among the methods of bankruptcy prediction, I have examined four analysis options in more detail. One type of classification procedure is **discriminant analysis**, which is an analysis method that classifies cases into existing categories of the dependent variable. The analysis can be used to identify those variables for which the existing groups are significantly different from each other. One disadvantage of the method is that it is not suitable for analysing cases where a non-linear relationship is assumed and there are a number of important conditions that are not easy to satisfy together, but there is a nowadays very popular alternative, logistic regression, which requires much fewer conditions to be satisfied (Malhotra, 2008; Sajtos-Mitev, 2007)

In **logistic regression,** the aim is the same as in discriminant analysis, but in contrast to discriminant analysis, the list of conditions is much shorter in logistic regression (Hajdú, 2003; Malhotra, 2008)

The **recursive partitioning algorithm** is known by several names, many people refer to it simply as the decision tree. The resulting subgroups are called nodes. The basis for prediction is the leaves, which are the part of the tree that is not redistributed. (Hajdú, 2008) The goal is

to minimize the variance within the group, so that the variance between groups is as large as possible. The algorithm is an iterative process designed specifically for computers.

There are several types of decision trees, of which CHAID has been widely used in bankruptcy forecasting. A major advantage of the analysis is that there is no restriction on the variables included, both metric and non-metric variables can be included. An argument in favour of the method is that the conditions do not include a normal distribution of variables. As a result, the corresponding solvency classification is achieved at a high rate, and the exact classification data can be found in the classification matrix in this case as well.

The disadvantage of this methodology is that it cannot be used for forecasting applications, as it is mostly specialised on the training database. However, the problem can be solved by using the method developed to control over-learning in artificial intelligence models, i.e. by dividing the data into a training and a testing part and examining whether similar results are obtained in both cases (Hámori, 2001)

The last method is the **neural network**. The best performing computer today is still the human brain. Neurons, information-processing units, help to perform tasks. "Neural networks, or more specifically artificial neural networks, are information processing paradigms inspired by the highly interconnected parallel processing structures and processes of the mammalian brain. In essence, neural networks are mathematical models that operate on the basis of certain information processing principles of biological nervous systems and are therefore capable of adaptive learning." (Ketskeméty et al., 2011, p394.)

In my analyses, I used the MLP method, which extends the simple perceptron with hidden layers between the input and output layers, improving the learning performance. Information can flow between the layers with and without feedback. The best known is the back propagation network, where the error propagates backwards, continuously changing the weights (Ketskeméty et al., 2011)

In my analyses, I used the MLP[1] method, which extends the simple perceptron with hidden layers between the input and output layers, improving the learning performance. Information can flow between the layers with and without feedback. The best known is the back propagation network, where the error propagates backwards, continuously changing the weights (Ketskeméty et al., 2011)

Neural networks have many advantages:

- it handles nonlinearity,
- no problem with missing data
- it can handle a large number of variables and elements (Kristóf, 2002).

I used the classification matrix, ROC curve and Gini coefficient to evaluate and compare the models. For the AUC value calculated from the ROC curve, if the value is around 80-90%, it is considered to be outstanding. Also for the Gini coefficient, a value between 80-90% indicates a very good predictive model.

[1] Multi-Layer Perceptron

3 Database of the research

3.1 Central Credit Information System

In Hungary, information on household creditors is recorded by the Central Credit Information System, or KHR for short, which helps banks to share information on creditors, assists in credit assessment and reduces the risk of over-indebtedness. The KHR records a so-called complete list, i.e. customers who meet their obligations on time are also included in the system.

Looking at the trend in defaults for the year 2021, the number of defaults has steadily decreased over the year, with the number of outstanding debts falling by 13.9% from January to December. The trend in outstanding debts has also been influenced by legislative changes, such as the gradual reduction of the moratorium on repayments.

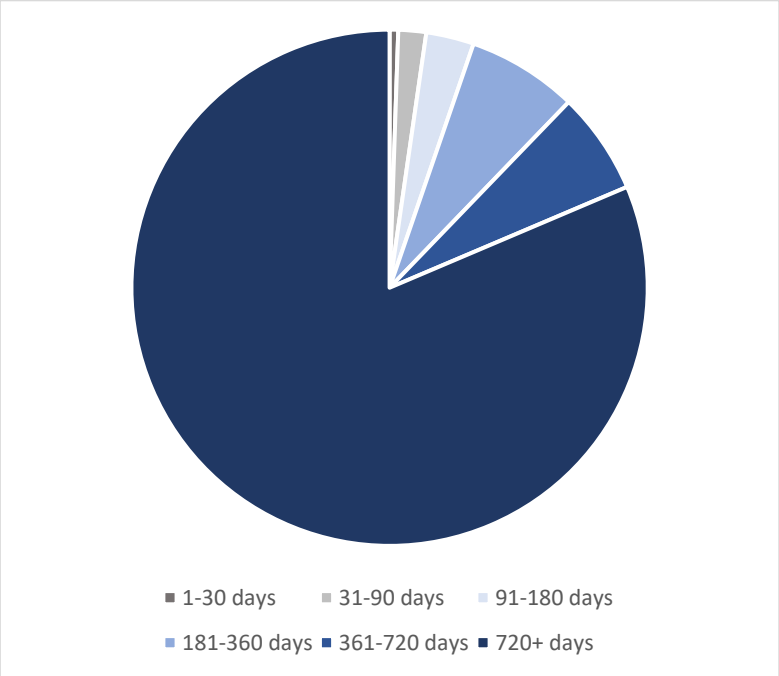

Figure 3: Distribution of outstanding defaults by duration of default (%; 2021)
Soure: Own edited, based on KHR informations

In terms of duration of defaults, 12.21% of the outstanding defaults have been outstanding for up to one year, 6.4% for less than 720 days and a significant proportion, 81.39%, for more than almost 2 years (KHR Annual Information, 2021)

3.2 Database

The database for the analyses was provided by BISZ Zrt. The data were extracted on 30 September 2021, so the database contains the persons registered on that date. A unit in the database represents one loan transaction, so there may be persons in the database who are listed more than once with different loan transactions. Overall, on that date, the register contained 10 767 452 credit transactions and the following variables:

- Anonymous identifier of the consumer
- Type of the consumer
    - debtor
    - co-debtor
- Age os the consumer (in 2021)
- Gender of the consumer: less than 1% of the database is classified in other categories. This is due to the fact that the variable is defined by examining the correspondence between male and female first names given by the MTA. Names from foreign countries and names with data errors are included in the "other" category.
- Distorted agreement identifier
- Type of agreement
- Status of agreement
    - closed
    - outstanding
- Date of the agreement
- Expiry date of the agreement
- Amount and currency of the agreement
- Amount and currency of principal debt outstanding
- Information on regular repayments (amount and currency)
- Amount and currency of the default
- Status of the default
    - closed
    - outstanding
- Date the default occurred
- Date the default was terminated
- Residence of consumer (on the district level of the country): Aggregated district classification from the postcode of the customer's address. Unclassifiable/incorrect postcodes have been placed in the "other" category.

In addition to the above variables, I added the following to the database:

- Loan maturity (difference between the date and expiry date of the agreement; in months)
- Default (yes or no)
- Repayment amount as a percentage of agreement amount (repayment amount/amount of agreement
- Age of the consumer at the time of borrowing the loan
- County: defined on the basis of the district using information from the KSH
- Region: defined on the basis of the county using information from the KSH
- Average salary (eFt): average salary in the county in the year of borrowing
- Unemployment rate: the unemployment rate in the county in the year of borrowing
- Employment rate: the employment rate in the county in the year of borrowing
- GDP per capita: value of the year of borrowing in the county
- Graduation rate: value of the year of borrowing in the county
- Ratio of average salary to contract amount
- Ratio of average salary to repayment amount

As the analysis is based on a smaller base, I will not further characterise the raw database, but some important features of the database will be identified in the process of filtering.

Before starting the analysis, the first step was to clean the database and reduce it to the size needed for the research objectives, which is summarised in the flowchart below:

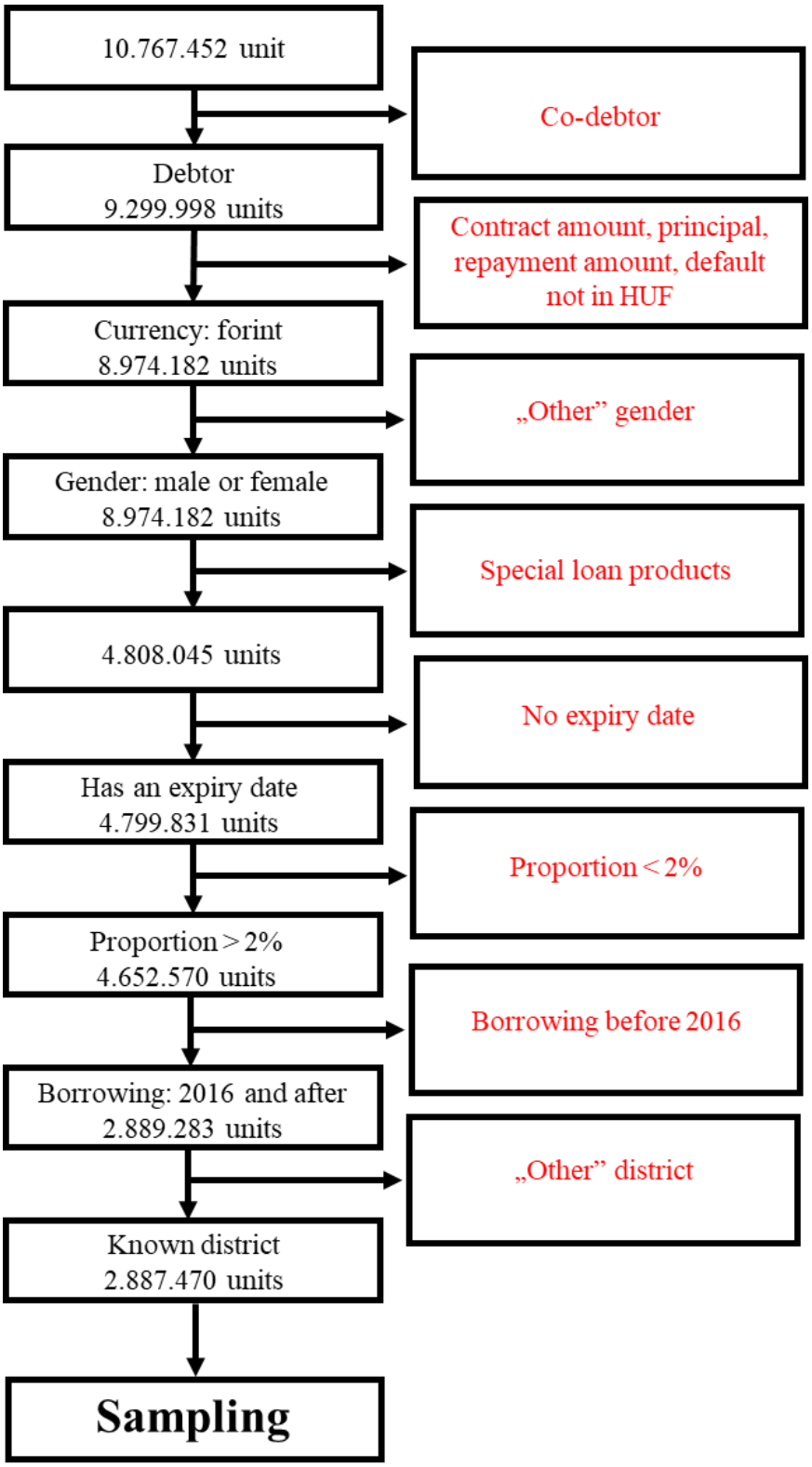

Figure 4: Steps to reduce the database

Source: Own editing

4 Results of the research

In my dissertation I formulated six hypotheses. The new and novel findings of my research are presented in the following chapter.

In order to meet the research objectives and questions, I worked with two types of samples when creating the full models, and extended them later. In both cases the sample was 500 units. The first sample was created using simple random sampling.

H1: The information recorded by the KHR can be used to predict the risk of default with a high degree of confidence using the classification procedures used in the bankruptcy prediction.

Recent methods used for bankruptcy prediction include logistic regression, recursive partitioning algorithms and neural networks. I hypothesize that these methods are capable of predicting with high accuracy which customer or loan transaction will default.

I classified as non-performing the loan transaction that had a default amount.

To support this claim, I constructed a classification model using all three methods. To perform the analysis, I used the database provided by the KHR and, to validate the results, I divided the sample into a training and a testing part. The training sample included 70% of the items.

Logistic regression I.

First, I performed a logistic regression analysis. Of the available explanatory variables, only the ratio of the repayment to the contract amount was found to be significant. The Omnibus test ($p<0.001$) and the Hosmer and Lemeshow goodness-of-fit test ($p=0.212$) resulted in a reliable model with a good fit. The generated model has medium explanatory power (Nailkerke R2=38.8%).

Table 1: Significant variables in the Logistic Regression I model

Sample		B	S.E.	Wald	df	Sig.	Exp(B)
Training	Rep. ratio	,029	,004	49,181	1	<,001	1,030
	Constant	-3,448	,308	125,008	1	<,001	,032

Source: Own editing

The resulting model equation can be written in the following form:

$$P_{(default)} = \frac{e^{0,032+1,030x_1}}{1+e^{0,032+1,030x_1}}$$

where,

x_1: repayment amount as a percentage of agreement amount.

Table 2: Classification matrix for the Logistic Regression I model

Sample	Observed		Predicted		
			Default		Percentage Correct
			0	1	
Training	Default	0	261	62	80,8
		1	4	23	85,2
	Overall Percentage				81,1
a. The cut value is ,039					

Source: Own editing

Overall, the model categorised the credit transactions correctly with 81.1% accuracy, 66 items were misclassified. Similar classification accuracy is achieved for the training and testing samples, but a much higher percentage of well performing loans are correctly categorised. This may be due to the predominance of performing loans in the sample, i.e. the sample composition is unfavourable for analysis. To improve this, I have performed the analysis on another sample later. The models created on the training and testing samples have similar classification accuracy, so the resulting model was accepted.

Decision tree I.

It is also important to mention at the beginning that one of the disadvantages of the decision tree is that it is prone to over-learning, which is also a risk in this case, as the sample is predominantly composed of well performing loans (93%). To construct the decision tree, I used the same four explanatory variables as for the Logistic Regression I model, and in this case the ratio of the repayment to the contract amount proved to be a good discriminating variable based on the algorithm.

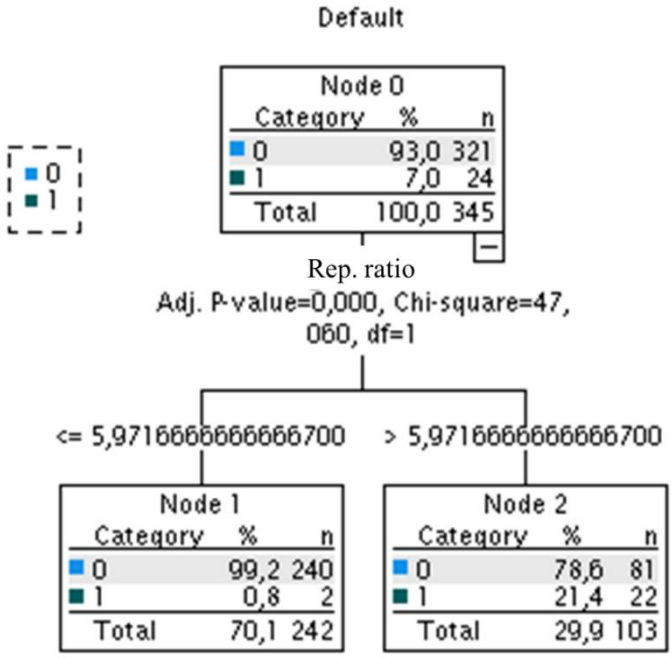

Figure 5: Decision tree I. for the training sample

Source: Spss output, own editing

The resulting decision tree consists of one level, which indicates that if the value of the variable is less than 5.9717, the number of non-performing loans is negligible.

Information on the accuracy of the classifications is provided by the classification matrix.

Table 3: Classification matrix for the Decision Tree I model

Sample	Observed	Predicted		
		0	1	Percent Correct
Training	0	312	0	100,0%
	1	24	0	0,0%
	Overall Percentage	100,0%	0,0%	93,0%

Source: Own editing

For the training dataset, the model achieved a classification accuracy of 93.0%, i.e. 93.0% of the transactions were correctly categorised, but none of the non-performing loans were correctly categorised. This is because the number of non-performing loans was too low in the sample, so the algorithm overestimated the classification of performing loans. A solution to this problem could be to create a sample with (approximately) equal proportions of performing and non-performing loans.

Neural network I.

In the case of neural networks, I chose the "Multi-Layer Perceptron" option, which is widely used in bankruptcy prediction. The neural network does not have any assumptions, but like the decision tree, there is also a risk of over-learning. All of the available variables can be included in the analysis, resulting in the following network:

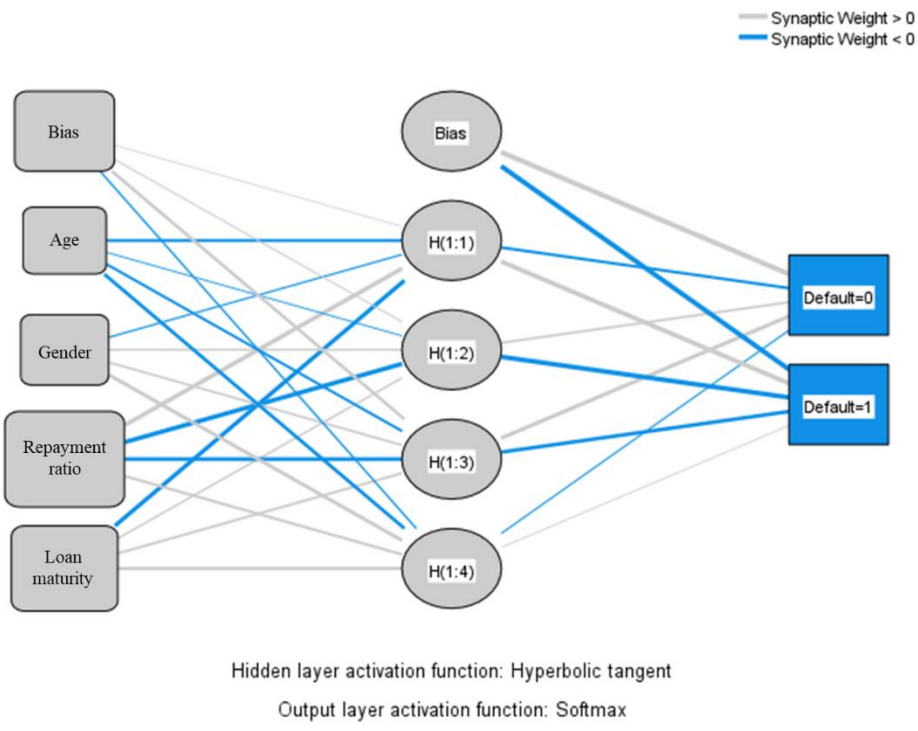

Figura 6: Neural network for the Neural Network I model

Source: Spss output, own editing

The algorihm chose the ratio of the repayment to the contract amount as the most important explanatory variable. The resulting model has a high classification accuracy, with 6.7% of loan transactions misclassified in the case of the training database and 3.2% in the case of the testing database.

Table 4: Classification matrix for the Neural Network I model

Sample	Observed	Predicted		
		0	1	Percent Correct
Training	0	307	9	97,2%
	1	14	12	46,2%
	Overall Percent	93,9%	6,1%	93,3%

Source: Own editing

In the classification matrix, we can see that the sensitivity value is lower than the specificity value, so this analysis method, like the decision tree, achieves higher accuracy for the classification of well performing loans. In this case, too, the use of a sample with (approximately) equal proportions of performing and non-performing loans could be a solution.

Creating a new sample

Comparing the results of the first three models, it can be concluded that the models constructed with logistic regression and neural networks are more favourable, but there is room for further improvement. The models do not benefit from a significantly higher proportion of either group in the sample. This problem has been faced by bankruptcy prediction researchers, who have carried out their analyses on samples with the same proportion of each group, e.g. Beaver (1966), Altman (1968), Deakin (1972), Odom and Sharda (1990). Accordingly, I have developed a new sample, also with 500 units, with a 50%-50% proportion of performing and non-performing loans.

Logistic regression II.

The model created on the new sample was found to be a reliable model based on the Omnibus test ($p<0.001$), and the Hosmer and Lemeshow goodness-of-fit test showed that the model created was a good fit ($p=0.105$). A significant difference in the explanatory power of the model was observed between the training (Nagelkerke R2=69.3%) and test (Nagelkerke R2=55.5%) samples. The variable that was found to be significant in the previous case was also found to be significant in this case ($p<0.001$), and the model was extended to include the variable loan maturity.

Table 5: Significant variables in the Logistic Regression II model

Sample		B	S.E.	Wald	df	Sig.	Exp(B)
Training	Rep. ratio	,050	,009	29,528	1	<,001	1,051
	Loan maturity	-,019	,006	9,418	1	,002	,982
	Constant	-,662	,287	5,309	1	,021	,516

Forrás: Saját szerkesztés

The resulting model equation can be written in the following form:

$$P_{(default)} = \frac{e^{0,516+1,051x_1+0,982x_2}}{1+e^{0,516+1,051x_1+0,982x_2}}$$

where,

$x_{1:}$ repayment amount as a percentage of agreement amount.

$x_{2:}$ loan maturity.

Finally, I examined the classification accuracy, the detailed values of which are given in the classification matrix below.

Table 6: Classification matrix for the Logistic Regression II model

Sample	Observed		Predicted		
			Default		Percentage Correct
			0	1	
Training	Default	0	139	24	85,3
		1	32	142	81,6
	Overall Percentage				83,4
a. The cut value is ,390					

Source: Own editing

Overall, the model categorised the credit transactions correctly with 83.4% accuracy, 56 items were misclassified. Based on the classification matrix, it can be concluded that the model achieved a higher classification accuracy for the training sample, which is due to the fact that the model identified the performing loans with a higher accuracy for the training sample, however, as this is not the final model, I consider the validation to be accepted in this case.

Decision tree II.

Again, the algorithm was based on the same four explanatory variables, and in this case the same variable was found to be significant as in the first case, namely the ratio of the repayment to the contract amount

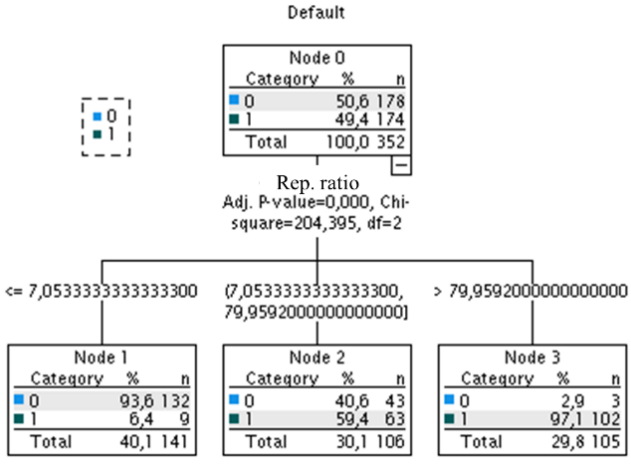

Figure 7: Decision tree II. for the training sample

Source: Spss output, own editing

This is the decision tree also has a level, but in this case the previous 2 leaves, there are now 3. Based on the resulting tree it can be concluded that if the value of the variable does not exceed 7.053, then the default rate is low, but if it exceeds 79,959, it is almost certain that the loan is a non-performing loan.

The classification matrix provides information on the accuracy of the classifications.

Table 7: Classification matrix for the Decision Tree II model

Sample	Observed	Predicted		
		0	1	Percent Correct
Training	0	132	46	74,2%
	1	9	165	94,8%
	Overall Percentage	40,1%	59,9%	84,4%

Source: Own editing

Although the accuracy of the first model (93.0%) decreased in this case, the fact that the first model could not categorise any non-performing loan transaction was not good for the purpose of the analysis. In this case, however, a significant proportion of non-performing loans were categorised in the correct group, and the results achieved on the training and test (86.5%) samples are not significantly different, so I consider the results obtained to be valid.

Neural network II

The set of selected independent variables is unchanged from the previous model, but the resulting neural network is significantly different, with more neurons in the new network.

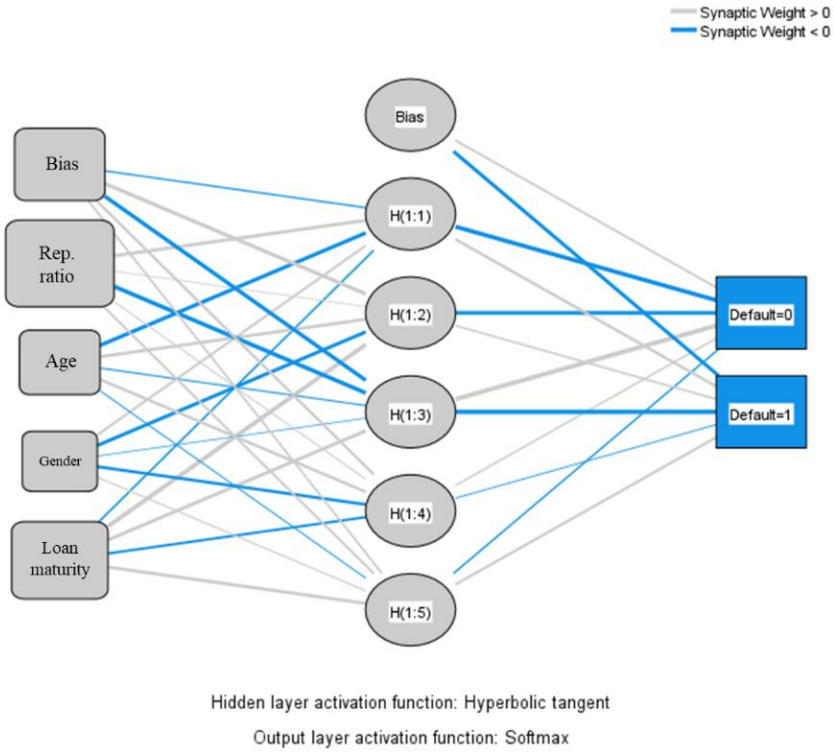

Figure 8: N eural network for the Neural Network II model

Source: Spss output, own editing

If we look at the importance of the different independent variables, we can still conclude that the most significant variable is the ratio of the repayment to the contract amount.

Although the overall accuracy of the classification has decreased, the specificity and sensitivity are approaching each other and the classification of non-performing loans has improved significantly.

Table 8: Classification matrix for the Neural Network II model

Sample	Observed	Predicted		
		0	1	Percent Correct
Training	0	163	14	92,1%
	1	53	132	71,4%
	Overall Percent	59,7%	40,3%	81,5%

Source: Own editing

The predictive ability of the model is similar for the test (86.2%) and the training sample, so the model can be validated.

Comparison of results

I based the initial models on four explanatory variables, and the table below summarises which explanatory variables were found to be significant by the different methods

Table 9: Summary of variables used by classification models

	Analysis I			Analysis II		
	Log. regr. I	Decision tree I	Neural network I	Log. regr. II.	Decision tree II	Neural network II
Rep. ratio	X	X	X	X	X	X
Loan maturity			X	X		X
Age			X			X
Gender			X			X

Source: Own editing

Based on the above, it can be concluded that the most significant of the data recorded by the KHR in terms of loan defaults is the ratio of the repayment to the contract amount.

There is also a significant difference in the classification accuracy of the different sample-based methods, as summarised in the table below.

Table 10: Classification accuracy of the used methods

	Percent correct					
Observed	Log. regr. I	Decision tree I	Neural network I	Log. regr. II.	Decision tree II	Neural network II
0	80,8%	100,0%	97,2%	85,3	74,2%	92,1%
1	85,2%	0,0%	46,2%	81,6	94,8%	71,4%
∑	81,1%	93,0%	93,3%	83,4	84,4%	81,5%

Source: Own editing

Based on the above, it can be concluded that all the chosen methods can be successfully applied to predict defaults with high accuracy. Although the classification accuracy of the initial models was higher in almost all cases, a significant difference in the categorisation of each group can be observed. For the newly developed sample, although the prediction ability decreased, the correct categorisation of defaulted loans improved significantly and the difference between sensitivity and specificity decreased, so that these models can be considered as better.

Based on the AUC and Gini coefficients in Table 20, all models can be rated as outstanding, strong models.

Based on the analysis, I **accept hypothesis H1** and formulate the following theses.

T1: The analyses confirm that the information recorded by the KHR can be used to construct a set of variables that can be used to predict the risk of default with a high accuracy.

T2: In the case of the methods used, the phenomenon of overlearning has been confirmed. The use of a new sample with an equal proportion of good and bad performing loans is a solution to this problem. It can be concluded that, although the classification accuracy decreases, the sensitivity increases, which is a positive characteristic for the aim of the analysis.

Expanding the database

When applying for a loan, the lending institution will ask for a wide range of information about the potential customer to decide whether or not to grant the applicant a loan, such as the customer's documented income. The KHR records do not include this information, so the information I use is narrower than the information on which banks have to base their decision. For this reason, I decided that it was necessary to add new information to the database. As these are sensitive data, the actual values were not available to me, so I used data from the KSH. I determined the county or region of residence of the credit applicant. I used the following tables to extend the database:

- 20.1.2.8 Unemployment rate by county and region [%]
- 20.1.2.6 Employment rate, by county and region [%]

- 20.8.2.1.1 Average monthly gross earnings of full-time employees by location of employer's headquarters, by county and region
- 21.1.2.2 Gross domestic product per capita by county and region
- 23.1.2.4 Percentage of graduates and tertiary degree holders by county and region [%]

I also used this new data to create ratios. A detailed description of the database is given in section 3.2.

H2: In corporate bankruptcy prediction, analyses run using neural networks have had the highest classification accuracy (Odom-Sharda, 1990; Olmeda-Fernandez, 1997). I hypothesize that this statement is also true for household loans, so neural networks will also provide the best results for categorizing household loans, outperforming all the methods I have used.

H3: The range of information recorded by the KHR is less than the information requested by banks when applying for a loan (e.g. income data). The inclusion of new variables not recorded by the KHR can improve classification accuracy.

I also performed all analyses on both samples with the new variables included.

Logistic regression III

This time, I carried out the analysis on the extended database and the first sample. The newly created model is a reliable (p<0.001), well-fitting (p=0.855) model with an explanatory power of 54.2%. No significant difference was observed between the training and testing samples. However, the set of variables found to be significant was extended compared to the original model. The explanatory variables in the table below were included in the final model:

Table 11: Significant variables in the Logistic Regression III model

Sample		B	S.E.	Wald	df	Sig.	Exp(B)
Training	Rep. ratio	,039	,007	32,258	1	<,001	1,040
	Sal. to contract	-,016	,004	14,017	1	<,001	,984
	Loan maturity	-,050	,016	10,065	1	,002	,951
	Constant	-,089	,758	,014	1	,907	,915

Forrás: Saját szerkesztés

A létrejött modell egyenlete a következő formában írható fel:

$$P_{(default)} = \frac{e^{0,915+1,040x_1+0,984x_2+0,951x_3}}{1+e^{0,915+1,040x_1+0,984x_2+0,951x_3}}$$

ahol,

x_1: repayment amount as a percentage of agreement amount,

x_2: ratio of average salary to contract amount,

x_3: loan maturity.

The inclusion of the new variables had a positive impact on the classification of non-performing loans and the resulting model correctly categorised 83.1% of the transactions.

Table 12: Classification matrix for the Logistic Regression III model

Sample	Observed		Pedicted			Log. regr. I.
			Default		Percentage Correct	
			0	1		
Training	Default	0	267	56	82,7	80,8
		1	3	24	88,9	85,2
	Overall Percentage				83,1	81,1
a. The cut value is ,068[2]						

Source: Own editing

As shown in Table 12, the new variables have led to a significant increase in the correct categorisation rate for non-performing loans.

Decision tree III

In the case of the decision tree, the algorithm did not add any new variables to the original one, so almost the same model (Decision Tree I) was created in this case. The only difference is that I chose the built-in validation option for the analysis, in which case the software always recreates the training and testing sample. The explanatory variable included was the same as in the first case, but in this sample split there are three leaves in the tree.

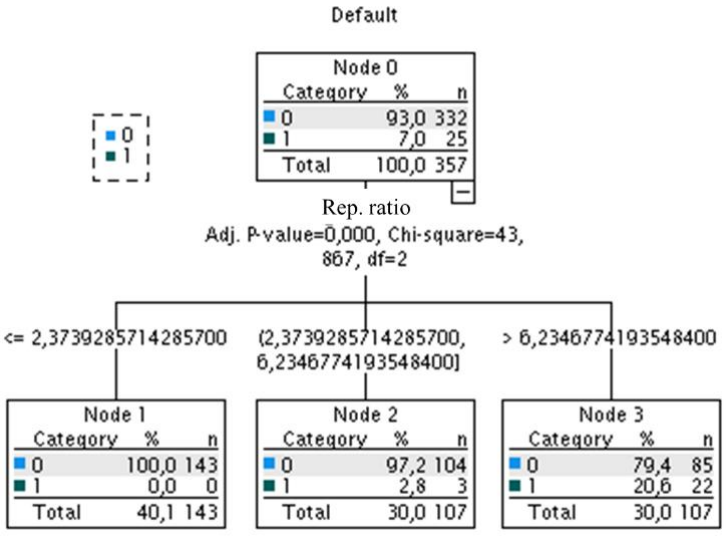

Figure 9: Decision tree III. for the training sample

Source: Spss output, own editing

[2] Alacsony vágási pont, azonban a csődelőrejelzések során is volt erre példa, Ohlson (1989) esetében p=0,038, az általa használt mintában a csődbe jutott vállalatok aránya 4,85% volt.

Based on the decision rules, it can be said that if the value of the variable exceeds 6.235, the probability of defaulting is the highest. The creation of the three letters does not affect the classification accuracy of the model. None of the non-performing loans could be correctly categorized by the algorithm.

Table 13: Classification matrix for the Decision Tree III model

Sample	Observed	Predicted		
		0	1	Percent Correct
Training	0	332	0	100,0%
	1	25	0	0,0%
	Overall Percentage	100,0%	0,0%	93,0%

Source: Own editing

Neural network III

In the case of the neural network, each new variable was included in the model, creating a much larger neural network with 4 neurons.

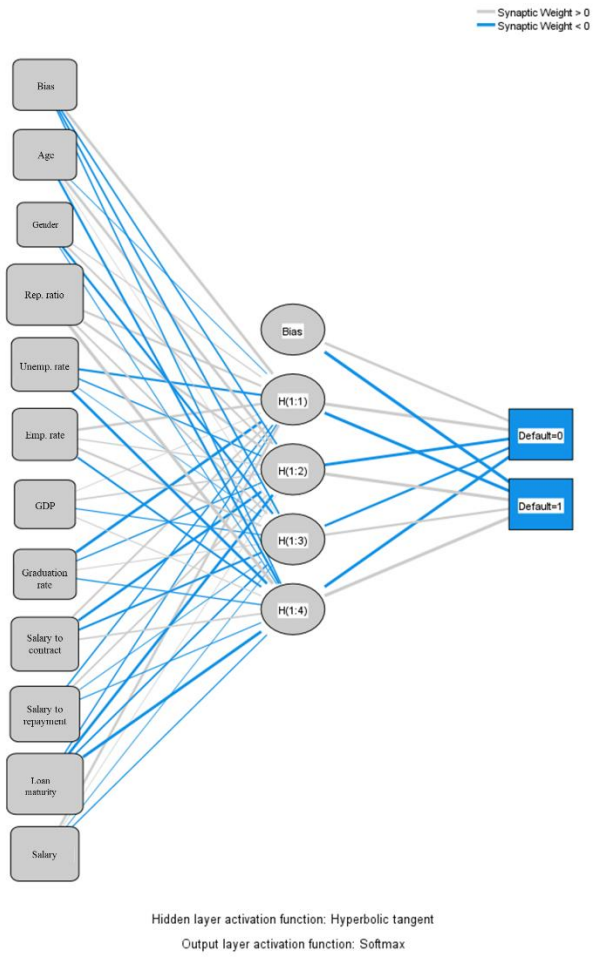

Figure 10: Neural network for the Neural Network III model

Source: Spss output, own editing

Examining the importance of the independent variables, the algorithm chose the ratio of repayment to contract amount, the duration and the ratio of salary to contract amount as the 3

most significant variables, which were also found to be significant variables in the logistic regression.

Table 14: Classification matrix for the Neural Network III model

Sample	Observed	Predicted			Neural network I.
		0	1	Percent Correct	
Training	0	318	1	99,7%	97,2%
	1	7	20	74,1%	46,2%
	Overall Percent	93,9%	6,1%	97,7%	93,3%

Source: Own editing

The classification accuracy of the new model increased compared to Neural Network I due to the new variables, the new model correctly categorised 97.7% of the transactions, but there is still a significant difference between specificity and sensitivity.

Logistic regression IV.

The latest logistic regression model has been extended with new variables compared to the previous ones. The model is reliable (Omnibus test p<0.001) and has a good fit (Hosmer and Lemeshow goodness-of-fit test p=0.218). In addition, it is found to have a high explanatory power (Nagelkerke R2= 74.0%). In the Logistic Regression II model, the ratio of the repayment to the contract amount and the maturity date were included as significant explanatory variables. These variables were also included in this model, but two additional variables were added, namely average salary and the ratio of salary to repayment amount.

Table 15: Significant variables in the Logistic Regression IV model

Sample		B	S.E.	Wald	df	Sig.	Exp(B)
Training	Rep. ratio	,038	,008	21,074	1	<,001	1,039
	Sal. to rep.	-,001	,000	7,369	1	,007	,999
	Loan maturity	-,024	,007	12,388	1	<,001	,976
	Average sal(eft)	-,008	,003	7,041	1	,008	,992
	Constant	3,027	1,015	8,885	1	,003	20,627

Source: Own editing

The resulting model equation can be written in the following form:

$$P_{(default)} = \frac{e^{20,627+1,039x_1+0,999x_2+0,976x_3+0,992x_4}}{1+e^{20,627+1,039x_1+0,999x_2+0,976x_3+0,992x_4}}$$

where,

x_1: repayment amount as a percentage of agreement amount,

x_2: ratio of average salary to repayment amount,

x_3: loan maturity,

x_4: average salary (eFt).

It is the last of the logistic regression models and has the largest number of independent variables.

The new variables also had a positive impact on classification accuracy.

Table 16: Classification matrix for the Logistic Regression IV model

Sample	Observed		Predicted			Logistic regression II
			Default		Percentage Correct	
			0	1		
Training	Default	0	131	32	80,4	85,3
		1	11	163	93,7	81,6
	Overall Percentage				87,2	83,4
a. The cut value is ,338						

Source: Own editing

For the model, classification accuracy has improved, with 87.2% of transactions now correctly categorised. In addition, a higher proportion of non-performing loans were correctly categorised by the model, so the inclusion of the new variables had a positive impact on the model.

Decision tree IV

For the last decision tree, the algorithm has chosen the ratio of the repayment to the contract amount as the first decomposing variable, i.e. it has classified the loan transactions into three groups according to this ratio at the first level of the tree.

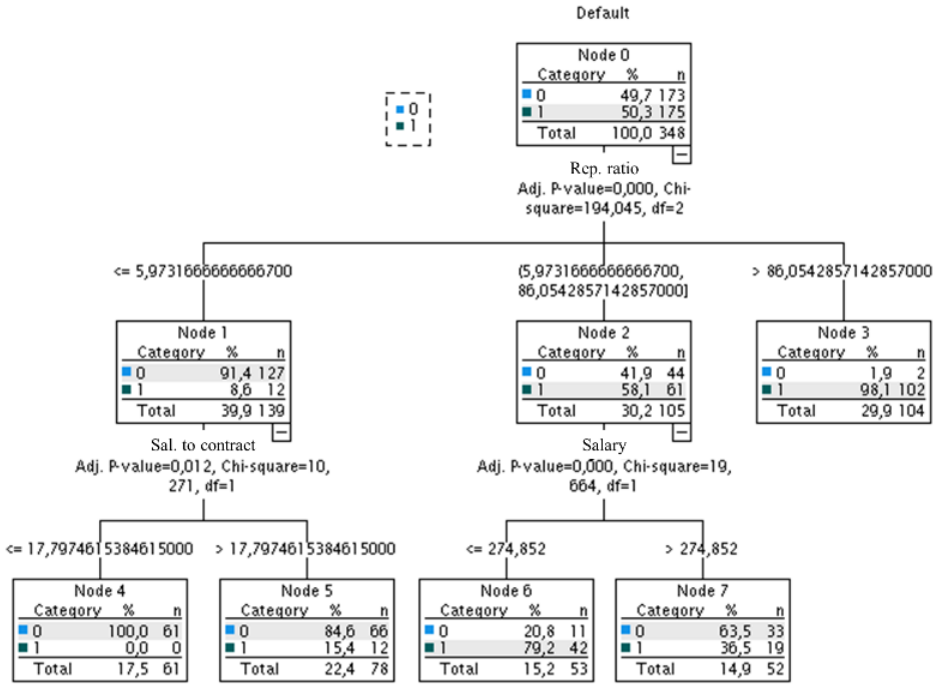

Figura 11: Decision tree IV for the training sample

Source: Spss output, own editing

For the Node 1 the next separating variable is the ratio of the salary to the contract amount, while for the Node 2, the two groups differed most in the average salary.

The decision tree shows that if

- the ratio of the repayment to the contract amount is less than 5.973 and the ratio of the salary to the contract amount is less than 17.798, the loan is a performing loan. If the ratio of the salary to the contract amount exceeds the above value, there is a chance that the loan will become non-performing;
- the ratio of the repayment to the contract amount is greater than 86.054, the loan is almost certainly a non-performing loan;
- the ratio of the repayment to the contract amount is between the above two values and the average income of the customer is less than 274,852 HUF the loan is more likely to be a non-performing loan.

Table 17: Classification matrix for the Decision Tree IV model

Sample	Observed	Predicted			Decision tree II
		0	1	Percent Correct	
Training	0	160	13	92,5%	74,2%
	1	31	144	82,3%	94,8%
	Overall Percentage	54,9%	45,1%	87,4%	84,4%

Source: Own editing

Compared to the Decision Tree II model, the classification accuracy has improved. The specificity and sensitivity values are closer to each other, which makes the latter model more favourable. The results obtained on the training and test (83.6%) samples do not differ significantly, so I consider the results obtained valid.

Neural network IV

The most complex neural network was found to be the last one, with 6 neurons.

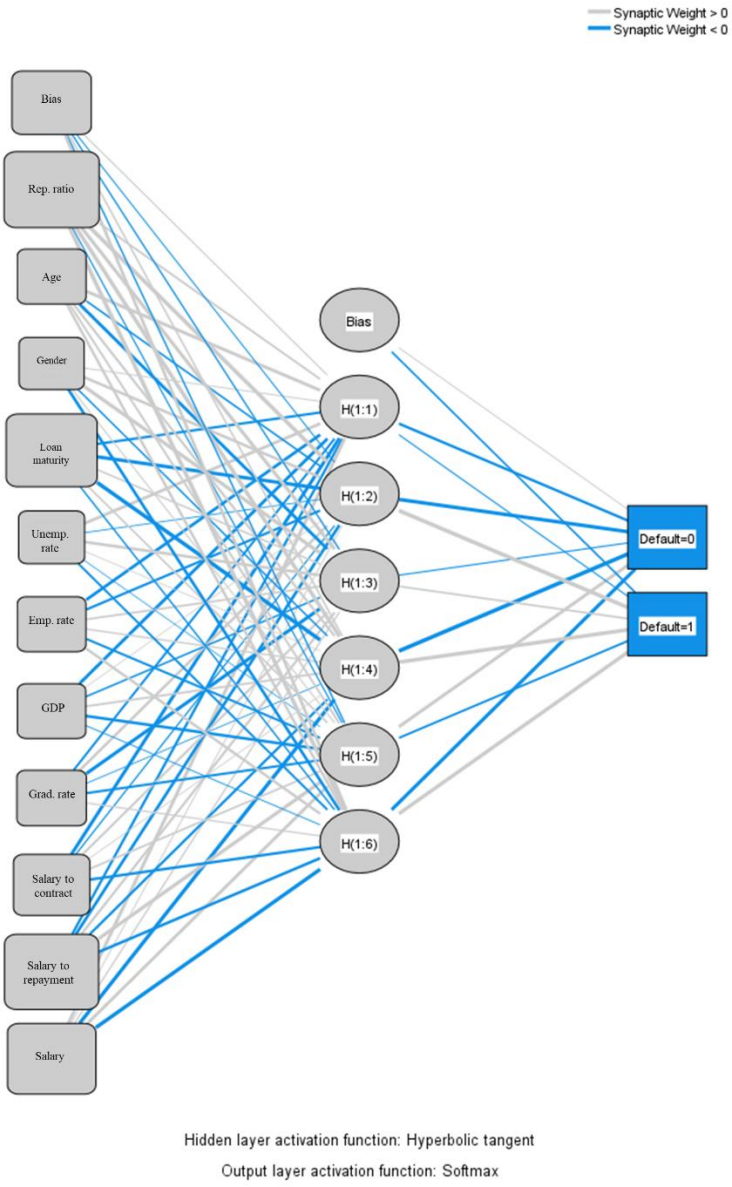

Figure 12: Neural network for the Neural Network IV model

Source: Spss output, own editing

When examining the importance of the explanatory variables, the following 5 explanatory variables were found to be the most important:

- the ratio of the repayment amount to the contractual amount
- salary as a percentage of the repayment amount
- loan maturity
- average salary
- ratio of salary to contract amount.

However, the other variables are also important for the algorithm, as leaving them out reduces the accuracy of the model classification. The least important variable considered by the algorithm is gender.

The final model has a high (88.3%) classification accuracy.

Table 18: Classification matrix for the Neural Network IV model

Sample	Observed	Predicted		
		0	1	Percent Correct
Training	0	161	14	92,0%
	1	26	142	84,5%
	Overall Percent	54,5%	45,5%	88,3%

Source: Own editing

The classification accuracy is almost the same for the training and the test sample (83.4%), so the resulting model can be finalised and 88.3% of the transactions are correctly categorised by the model.

Comparison of results II.

In total, 12 models were created, of which the first six (Logistic Regression I-II; Decision Tree I-II; Neural Network I-II) contained only the data recorded by the KHR, and then I extended the models with the county-level data recorded by the KSH and the ratios calculated from them.

In the table below I summarise which explanatory variables were found to be significant in the models created.

Table 19: Summary of the explanatory variables of the models created

	Logistic regression				Decision tree				Neural network			
	I.	II.	III.	IV.	I.	II.	III.	IV.	I.	II.	III.	IV.
Rep. ratio	X	X	X	X	X	X	X	X	X	X	X	X
Sal. to cont.			X					X			X	X
Sal. to rep.				X							X	X
Loan mat.		X	X	X					X	X	X	X
Av. salary				X				X			X	X
Age									X	X	X	X
Gender									X	X	X	X
Unemp. rate											X	X
Emp. rate											X	X
GDP											X	X
Graduates											X	X

Source: Own editing

For the analyses where the new variables were already available, they were found to be significant in 5 out of 6 models, with only Decision Tree III not finding a significant variable, presumably due to the unfavourable sample composition.

In Table 19, it can be observed that the ratio of the repayment amount to the contract amount is included as an explanatory variable in each model. In addition, the loan maturity, average salary, salary to repayment and contract amount were found to be significant in several cases.

However, the inclusion of new variables does not necessarily guarantee that the model will actually be improved. Thus, I considered it important to compare the indicators used to evaluate the classification methods.

First, I examined the ROC curves, which are shown here divided into two groups.

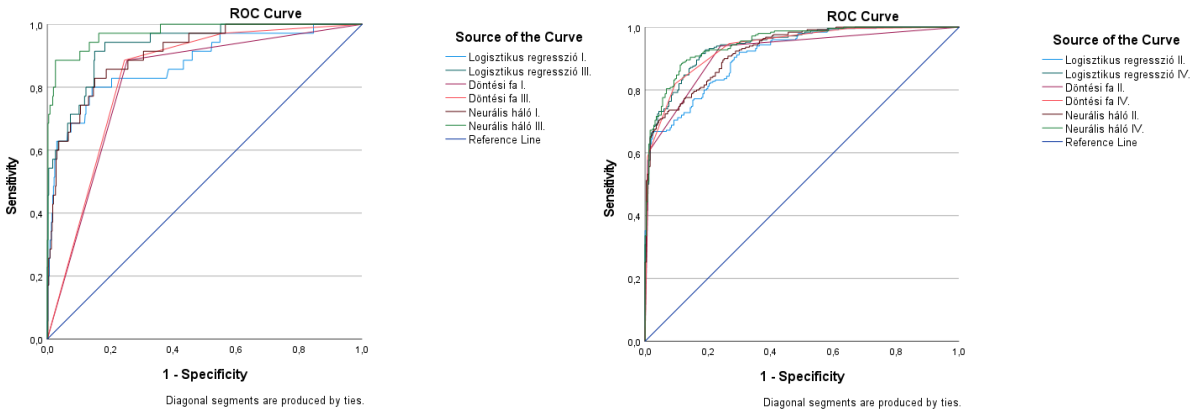

Figure 13: ROC curve of the created models

Source: Spss output, own editing

A more significant difference is observed for models I and III. The weakest models in the figure are Decision Tree I and Decision Tree III. In the figure on the right, such strong differences are not observed anymore. To make it easier and clearer to rank the different models, I have compiled the results obtained from the different evaluation techniques in a table.

Table 20: Performance of the models developed using different evaluation techniques

		Accuracy			AUC (%)	Gini (%)
		0	1	∑		
Logistic regression	I.	80,8	85,2	81,1	87,7	75,4
	II.	85,3	81,6	83,4	91,2	82,4
	III.	82,7	88,9	83,1	93,4	86,8
	IV.	80,4	93,7	87,2	94,1	88,2
Decision tree	I.	100	0	93	81,6	63,2
	II.	74,2	94,8	84,4	91,7	83,4
	III.	100	0	93	83,5	67,0
	IV.	74,2	94,8	84,4	93,7	87,4
Neural network	I.	97,2	46,2	93,3	90,8	81,6
	II.	92,1	71,4	81,5	92,5	85,0
	III.	97,2	46,2	93,3	97,4	94,8
	IV.	92	84,5	88,3	94,9	89,8

Source: Own editing

For the AUC value[3], a value of around 80-90% is considered to be outstanding. The AUC of all the models created is significantly higher than 80% and in most cases even 90%.

A similar conclusion can be drawn for the Gini coefficient, where a value above 70% is indicative of a very strong model. With the exception of Decision Tree I and Decision Tree III, all models have values above this level.

According to all three criteria considered, the best model would be Neural Network III, but there is a significant difference in the categorisation of good and bad performing loans, so this model is not the most ideal. Based on the above values, the classification accuracy of the Neural Network IV model is 88.3%, the AUC is 94.9% and the Gini coefficient is 89.8%. Based on these results, this model is the most ideal for predicting defaults on household loans. The second best model is Logistic Regression IV, with values that are just below those of the neural network. Both models achieved this performance on the extended database, so it can be concluded that the inclusion of new variables had a positive effect.

Based on the analysis, I therefore **accept hypotheses H2 and H3,** and my findings are summarised in the following theses:

T3: The new variables had a positive effect on the performance of the analyses for all models (classification accuracy, AUC and Gini coefficient). Despite an overall decrease in classification accuracy, the difference between specificity and sensitivity was significantly reduced and together these led to an improvement in model performance.

T4: Of the methods used, neural networks have the highest classification accuracy, with Neural Network IV being the best-ranked model, achieving classification accuracy of nearly 90%.

T5: The following variables were the main determinants of default:

- **repayment amount as a percentage of agreement amount**
- **loan maturity**
- **average salary**
- **ratio of average salary to contract amount**
- **ratio of average salary to repayment amount**

Similar classification accuracy can be achieved for retail loans using the same methods used in bankruptcy forecasting.

[3] Area under the ROC curve

Impact of economic, social, demographic factors on default

H4: The different social and economic situation in different parts of Hungary can be linked to loan defaults.

H5: Different demographics (gender and age) affect non-payment.

To test the assumptions, I first take the classification models that have been carried out. In both cases, the last of the logistic regression and decision tree models was the one where average salary was found to be a significant explanatory variable. In addition, in these models, no data collected from the KSH was included as a significant variable. The difficulty with neural networks is that the algorithm unconditionally includes all variables in the analysis, i.e. the different economic, social and demographic variables.

I then used non-parametric tests to confirm or reject the hypothesis. The question is interesting in terms of the form in which I would like to include non-performance in the analysis. In the same form as I have used previously (0= performing well, 1= defaulted), or would I like to look at the sum of the defaults. Since the focus of the research was only on non-performance and not on its amount, I chose the former and used a cross tabulation analysis to see if a relationship could be observed for region and gender. For both gender and region, I concluded that there was no significant relationship ($p>0.05$).

Based on this, I **reject hypotheses H4 and H5**.

T6: Based on the classification analyses and non-parametric tests carried out, I conclude that none of the social and economic characteristics can be associated with the fact of default at the territorial aggregate level.

T7: Based on the classification analyses and non-parametric tests carried out, I conclude that none of the available demographic data can be associated with the fact of default at the territorial aggregate level.

Is there a difference between the different types of credit?

For different types of credit, banks may require different conditions from the customer, which may even affect default. Thus, I would like to investigate whether a better model is available when analysed by type of loan and what differences can be observed between the importance of the explanatory variables included.

H6: For different types of credit, banks may require different conditions from the customer, which may even affect default. My hypothesis is that if we build the predictive model by loan type, better classification accuracy can be achieved.

Based on the analyses carried out earlier, I decided to use the sample with equal proportions of good and bad performing loans to answer the question, and also to use the explanatory variables from the KSH. And of the methods, I chose the neural network as it produced the best results in all aspects. That is, I will compare the new models to the Neural Network IV model.

Other credit transaction

As with the global model, 30% of the sample was the testing sample. The neural network created has 6 neurons. Similar classification accuracy was observed for the training and testing samples. If we consider the other credit transactions independently, no better results are obtained for classification accuracy, AUC value and Gini coefficient.

Since all variables are part of the model in the case of neural networks, I examined the importance of each variable. Among the explanatory variables that were considered important, I highlighted the five most important variables. For other credit transactions, age also appeared as an important explanatory variable, which was not included among the five most important variables in the global model.

Consumer credit

For consumer credit and goods credit, the largest neural network ever created, with 9 neurons. The classification accuracy of the training and test samples is identical and the model achieved the highest classification accuracy ever. So in this case, a stand-alone model may be recommended.

The algorithm has identified variables that are not considered important by any previous model, namely GDP and unemployment rate. The reason for this may be that this type of credit is easy to apply for and requires fewer conditions to be met for disbursement. Basically, a certain segment of the society is more likely to apply for these loans.

Vehicle leasing

In the case of vehicle leasing, I found that the classification of non-performing loans is better than in the global model and the model performance is also better, so it may be appropriate to use a separate model for this loan as well.

For the variables, the maturity has been replaced by the value of the GDP per capita of the county, which can provide information on the economic situation of the are

Home mortgage loan

For housing loans, a smaller neural network with two neurons has been created. The performance of this model is similar to that of the base model, with only a marginal difference. However, the sensitivity value has improved significantly, which is positive.

The model basically chose the same variables, but age is included among the five most important variables instead of average salary.

Free use mortgage loan

In the case of a free-use loan, the performance of the stand-alone model is almost identical to that of the global model, so there is no advantage or disadvantage to using the stand-alone model. However, for the stand-alone model, other variables have been found to be important, which mainly characterise the economic situation of the county, namely employment rate, unemployment rate and GDP per capita.

Personal loan

The last group is personal loans. With personal loans, the same can be said as with free-use loans. Nearly the same classification accuracy was obtained, with the classification of non-performing loans slightly exceeding the classification of the global model. The set of explanatory variables is fully consistent with the explanatory variables of the global model.

The table below summarises and compares the performance of each model and the global model.

Table 21: Comparing the performance of the models

	Other credit trans.	Consumer credit	Vehicle leasing	Home mortgage loan	Free use mortgage loan	Personal loan	Global model
0	88,60%	97,20%	90,50%	90,10%	92,10%	92,00%	92,00%
1	86,10%	95,70%	91,20%	91,20%	83,30%	86,60%	84,50%
∑	87,30%	96,50%	90,80%	90,60%	88,00%	89,40%	88,30%
AUC	93,30%	99,80%	97,20%	93,20%	95,70%	95,40%	94,90%
Gini	86,60%	99,60%	94,40%	86,40%	91,40%	90,80%	89,80%

Source: Own editing

It can be seen that none of the individual models fall significantly short of the global model performance, with performance being close to or better than the global model. In the case of consumer credit, vehicle leasing and home mortgages, the sensitivity value increased significantly, so that the resulting individual models were able to identify non-performing loans with greater accuracy in this case.

For most credit types, at least 2 of the indicators used to measure the models exceed the global model and can therefore be considered better.

Based on the above analysis, I **accept hypothesis H6.**

T8: For consumer credit, vehicle leasing, free use mortgages and personal loans, better performance can be achieved using neural networks when analysed separately for each loan type.

4.1 Usability of the results, further research directions

My findings are based on data for the years 2016-2021.

A limitation of the research is that the analyses were carried out for the six most common types of loans in the reduced database, and there were some loans that had to be excluded from the analyses because of their characteristics (e.g. "Baby waiting loans"). An interesting area for future work could be the analysis of these loans.

The results obtained may be distorted by the fact that due to the COVID-19 virus, the state introduced a moratorium on repayments on 19 March 2020, which was still in place when the database was retrieved. For this reason, a future study could examine whether different conclusions could be drawn at a later date, when the moratorium was no longer in place and a longer period of time had passed since its lifting. The OECD has also mentioned in its 2021 survey that there is a likelihood of an increase in the number of non-performing loans following after the end of the moratorium (OECD, 2021b).

In terms of the research and the usefulness of the results, I think that this is a missing topic, as few researchers and few academic works deal with the characterisation and prediction of household loan defaults using statistical methods.

5 Summary

My PhD research investigated the default and prediction of default on household loans. My aim was to develop a model that can be used to categorise defaulted loans with high accuracy.

In the first half of the research, I reviewed the literature on the topic. I gave a brief overview of the history of lending, the current lending situation in Hungary and the current level of indebtedness. I have explained how the loans available to the public can be categorised, and what are the basic concepts that are of great importance when it comes to lending. I have also briefly examined the factors that may play an important role in lending behaviour, despite the fact that these are factors that are difficult to measure.

Since there is a scarcity of literature on the classification of retail loans, the research was based on bankruptcy prediction models. To get to know this discipline, I conducted a chronological and methodological review. Starting from the first univariate bankruptcy prediction analyses, I examined Altman's famous model and the evolution of this discipline up to the neural network used today.

After having identified the most commonly used methods, a short methodological overview followed, in which I discussed the limitations, advantages and disadvantages of each of the analysis options, and finally I came to the conclusion that I would like to use logistic regression, decision tree and neural network in my research. Also in this chapter I summarised the techniques that can be used to evaluate and compare the classification procedures.

The database on which the research was based was provided by BISZ Ltd.., so I briefly introduced the KHR and its activities. The initial database provided by BISZ Ltd. contained more than 10.000.000 loan transactions, so it was necessary to narrow down the database, which I did according to the research objectives. Finally, the sampling was based on a database of 2,887,470 loan transactions, from which I took a sample using two different techniques.

Based on the analyses, I conclude that the methods used work optimally when the sample composition has (approximately) the same proportion of performing and non-performing loans. The most accurate categorisation was achieved by using the neural network method, which not only used the data recorded by the KHR, but also used information collected from the KSH.

I also found that various economic, social and demographic indicators are not related to non-payment.

And finally, I have found that there are some types of loans for which it may be advisable to carry out the analyses independently. This may be due to the disbursement requirements of each type, as there are some loans for which providers have lower restrictions.

Úgy gondolom, hogy a kutatás elején kitűzött célokat teljesíteni tudtam, és bízom benne, hogy a kapott eredmények mások számára is hasznosnak bizonyulnak.

I believe that I was able to meet the objectives I set at the beginning of the research and I am confident that the results will be useful to others.

6 References

1. Altman, E. I. (1968): Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, The Journal of Finance, Vol. 23. No. 4. 589-609. old.
2. Altman, E. I., Haldeman, R., Narayanan, P. (1977): Zeta Analysis: A New Model to Identify Bankruptcy Risk of Corporations. Journal of Banking & Finance, 1, 1977.
3. Beaver, W. H. (1966): Financial ratios as predictors of failure, Empirical Research in Accounting: Selected Studies, Journal of Accounting Research, Supplement to Vol. 5. 71-111. old.
4. Béres, D., Huzdik, K., Németh, E., Zsótér, B. (2015). Pénzügyi személyiség - A magyar lakosság pénzügyi magatartása, szokásai és attitűdjei. Forrás: Pénziránytű Alapítvány: https://www.penziranytu.hu/sites/default/files/csatolmany/penzugyi_szemelyiseg_kutatasi_jelentes_2015_11_02.pdf
5. Béres, D., Huzdik K. (2022): A fiatal felnőttek és a hitelek (kockázatkerülési csapda) Pénzügyi szemle, 2022/1 33-55
6. Blum, M. (1974): Failing company discriminant analysis. Journal of Accounting Research, Vol. Spring pp.1-21.
7. Deakin, E.B. (1972): A Discriminant analysis of predictors of business failure. Journal of Accounting Research, 10-1 (Spring 1972) pp. 167-179
8. Dobák, P., Sági, J. (2005). Fogyasztási hitelek: növekvő eladósodottság? Forrás: Hitelintézeti Szemle: http://www.bankszovetseg.hu/Content/Hitelintezeti/51Dobak_Sagi.pdf
9. Fekete, O., Tatay, T. (2012): Hitelezők és adósok kapcsolatának szabályozási kérdései https://kgk.sze.hu/images/dokumentumok/kautzkiadvany2012/penzugy/fekete_tatay.pdf
10. Frydman, H., Altman, E. I., Kao, D. L. (1985): Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress, The Journal of Finance, Vol. 40. No. 1. 303-320. old.
11. GFLEC. (n.d.) S&P Global Finlit Survey. Global Financial Literacy Excellence Center, The George Washington University. https://gflec.org///initiatives/sp-global-finlit-survey/
12. Hajdu, O. (2003): Többváltozós statisztikai számítások, Központi Statisztikai Hivatal, Budapest
13. Hámori, G. (2001): A CHAID alapú döntési fák jellemzői, Statisztikai Szemle, 79. évf. 8. sz. 703-710. old. http://www.ksh.hu/statszemle_archive/2001/2001_08/2001_08_703.pdf
14. KHR éves tájékoztatók. (2011-2022). Forrás: https://www.bisz.hu/dokumentumtar
15. Ketskeméty, L., Izsó, L., Könyves Tóth, E. (2011): Bevezetés az IBM SPSS Statistics programrendszerbe, Artéria Stúdió Kft, Budapest
16. Kovács, P., Révész, B., Ország, G. (2014). A pénzügyi kultúra és attitűd mérése. Forrás: www.eco.u-szeged.hu/download.php?docID=40014
17. Kristóf, T. (2002): A mesterséges neurális hálók a jövőkutatás szolgálatában. Jövőelméletek, 9. BKÁE Jövőkutatási Kutatóközpont, Budapest.
18. KSH táblázatok:
    - 20.1.2.8. Munkanélküliségi ráta, vármegye és régió szerint [%]
    - 20.1.2.6. Foglalkoztatási ráta, vármegye és régió szerint [%]
    - 20.8.2.1. A teljes munkaidőben alkalmazásban állók havi bruttó átlagkeresete a munkáltató székhelyének elhelyezkedése alapján, vármegye és régió szerint
    - 21.1.2.2. Egy főre jutó bruttó hazai termék vármegye és régió szerint

- 23.1.2.4. Érettségizettek és felsőfokú oklevelet szerzettek aránya vármegye és régió szerint [%]

19. Lautenschläger, S. (2015). Reintegrating the banking sector into society: earning and reestablishing trust. Forrás: ECB: https://www.ecb.europa.eu/press/key/date/2015/html/sp150928.en.html
20. Lusardi, A., Klapper, L., Oudheusden, P. (2015). Financial literacy around the World (S&P). Forrás: http://gflec.org/wp-content/uploads/2015/11/Finlit_paper_16_F2_singles.pdf
21. Malhotra, N. K. (2008): Marketingkutatás (Akadémiai Kiadó, Budapest)
22. MNB (2023): Hitelezési folyamatok 2023. március
23. MNB (2023): Tovább erősíti a hitelpiac egészséges hosszú távú működését az MNB https://www.mnb.hu/sajtoszoba/sajtokozlemenyek/2023-evi-sajtokozlemenyek/tovabb-erositi-a-hitelpiac-egeszseges-hosszu-tavu-mukodeset-az-mnb
24. Odom, M.D., Sharda, R. (1990): A Neural Network Model for Bankruptcy Prediction. IJCNN International Joint Conference on Neural Networks, Vol. 2, San Diego, CA, 1990. pp. 163-1
25. Ohlson, J. (1980): Financial ratios and the probabilistic prediction of bankruptcy, Journal of Accounting Research, Vol. 18. No. 1. 109-131. old.
26. Olmeda, I., Fernandez, E. (1997): Hybrid Classifiers for Financial Multicriteria Decision Making: The Case of Bankruptcy Prediction. Computational Economics, Vol. 10. Issue 4. pp. 317–352.
27. Potóczki, J. (2017). A magyar lakosság pénzügyi kultúrájának szintje az öngondoskodás tükrében - nemzetközi és hazai kutatási eredmények. Forrás: JATEPress, Szeged: http://acta.bibl.u-szeged.hu/49707/1/gtk_2017_157-170.pdf
28. Sajtos L., Mitev A. (2007): SPSS kutatási és adatelemzési kézikönyv (Alinea Kiadó, Budapest)
29. Szabó, D. M. (2021): A lakossági hitelezés fejlődés és a pénzügyi fogyasztóvédelem
30. Vértesy, L. (2008): A pénzügyi intézmények finanszírozási tevékenységének jogi szabályozása Magyarországon Ph. D értekezés
31. Virág, M. (2004): A csődmodellek jellegzetességei és története, Vezetéstudomány, 35. évf. 10. sz. 24-32. old.
32. Virág, M., Kristóf, T. (2005): Az első hazai csődmodell újraszámítása neurális hálók segítségével. Közgazdasági Szemle, 52. évf. 2. sz. 144-162. old.
33. Zhang, G., Hu, M., Patuwo, B. (1999): Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. European Journal of Operational Research, Vol. 116. pp. 16–32.

7 Publications

1. Fodor, K., & Varga, B. (2023). A pénzügyi kultúra és egyes társadalmi tényezők kapcsolata. In What will our Future be Like? 2 essays in German, 7 in English, 30 in Hungarian language (Német, angol és magyar nyelvű esszék) (pp. 191–199).
2. Kitti, F., & Beatrix, V. (2023). CAN HOMOGENEOUS GROUPS OF STUDENTS BE FORMED IN THE CASE OF FINANCIAL ATTITUDES? In Building Up Financial Literacy (pp. 35–48).
3. Fodor, K. (2023). Logistic Regression or Neural Network? Which Provides Better Results for Retail Loans? THEORY METHODOLOGY PRACTICE: CLUB OF ECONOMICS IN MISKOLC, 19(1), 53–62. http://doi.org/10.18096/TMP.2023.01.05
4. Beatrix, V., & Kitti, F. (2022). HAS THE FINANCIAL LITERACY OF STUDENTS CHANGED? In *Emerging Markets Economics and Business. Contributions of Young Researchers.* (pp. 105–108).
5. Varga, B., Fodor, K., Szilágyi, R., & S., G. A. (2022). A hallgatói lemorzsolódás jelensége és okai, tapasztalatok a Miskolci Egyetemen (The Phenomenon and Causes of Student Dropout, Experiences at the University of Miskolc). In *"Mérleg és Kihívások - Fenntarthatóság" XII. Nemzetközi Tudományos Konferencia* (pp. 604–613).
6. Varga, B., & Fodor, K. (2022). Változott-e az egyetemisták pénzügyi attitűdje? (Has the financial attitude of students changed?). In *"Mérleg és Kihívások - Fenntarthatóság" XII. Nemzetközi Tudományos Konferencia* (pp. 594–603).
7. Fodor, K., & Varga, B. (2022). Kritikus fontosságú nyersanyagok a történelemben és napjainkban. In *Doktoranduszok Fóruma, MIskolc, 2020. november 19.* (pp. 73–79).
8. Varga, B., & Fodor, K. (2021). Technological Innovation and the Critical Raw Material Stock. In *Data Analysis and Rationality in a Complex World* (pp. 321–329). http://doi.org/10.1007/978-3-030-60104-1_35
9. Beatrix, V., & Kitti, F. (2020). ECONOMIC STRENGTH – CRITICAL RAW MATERIAL STOCK. In *Proceedings of the Miskolc IPW- IV. Sustainable raw materials international project week* (pp. 268–276).
10. Szilágyi, R., Molnár, L., Lengyel, L., Fodor, K., & Tóthné, K. A. (2020). Munkaerő-piaci kompetencia-igény prognózis. In *Társadalmi és gazdasági folyamatok elemzésének kérdései a XXI. században* (pp. 61–79). http://doi.org/10.14232/tgfek21sz.5
11. Varga, B., & Fodor, K. (2020b). ANALYSIS OF CRITICAL RAW MATERIALS THROUGH THE EYES OF ECONOMISTS. *GEOSCIENCES AND ENGINEERING: A PUBLICATION OF THE UNIVERSITY OF MISKOLC*, *8*(12), 94–104.
12. Varga, B., & Fodor, K. (2020a). A TECHNOLÓGIAI INNOVÁCIÓ ÉS A KRITIKUS NYERSANYAGOK. In *TÁRSADALMI, TECHNOLÓGIAI, INNOVÁCIÓS HÁLÓZATOK ASPEKTUSAI* (pp. 146–162).
13. Kitti, F. (2020). Using Multivariate Statistical Methods for Analysing Financial Literacy, as a Possible Appearance of Social Innovation. *THEORY METHODOLOGY PRACTICE: CLUB OF ECONOMICS IN MISKOLC*, *16*(1), 11–18. http://doi.org/10.18096/TMP.2020.01.02
14. Fodor, K., & Dr. Varga, B. (2019b). The technology innovation and the critical raw material stock. In *16th Conference of the International Federation of Classification Societies, IFCS2019* (pp. 142–142).
15. Fodor, K., & Dr. Varga, B. (2019a). Kann die logistische Regression zur Klassifizierung kritischer Rohstoffe verwender werden? In *Society, Economics, Law Conference, Program, Abstracts* (pp. 23–23).

16. Kitti, F. (2019). MONEY OR LIFE? – THE FINANCIAL LITERACY OF STUDENTS. In *„MÉRLEG ÉS KIHÍVÁSOK" XI. NEMZETKÖZI TUDOMÁNYOS KONFERENCIA Ph.D. szekciók tanulmánykötete* (pp. 65–74).
17. Beatrix, V., & Kitti, F. (2019a). Development of Transport and Critical Raw Materials. In *Emerging Markets Economics and Business. Contributions of Young Researchers. Proceedings of the 10th International Conference of Doctoral Students and Young Researchers* (pp. 135–139).
18. Fodor, K., & S., G. A. (2019). Eladósodottság, mint a pénzügyi kultúra állapotának mutatója. In *Doktoranduszok Fóruma 2018: A Gazdaságtudományi Kar szekciókiadványa* (pp. 29–36).
19. Varga, B., & Fodor, K. (2019a). A FENNTARTHATÓ FEJLŐDÉS ÉS A KRITIKUS ELEMEK. In *"Mérleg és Kihívások" XI. Nemzetközi Tudományos Konferencia* (pp. 393–403).
20. Lengyel, L., & Kitti, F. (2019). ANALYSIS OF EFFECTS ON LIFE EXPECTANCY AT BIRTH. In *"Mérleg és Kihívások" XI. Nemzetközi Tudományos Konferencia* (pp. 225–233).
21. Varga, B., & Fodor, K. (2019b). Electric Vehicles and Critical Raw Materials. *ANNALS OF THE UNIVERSITY OF ORADEA ECONOMIC SCIENCE*, *28*(2), 444–455.
22. Beatrix, V., & Kitti, F. (2019b). Kann die logistische Regression zur Klassifizierung kritischer Rohstoffe verwendet werden? In *People and their values in the society* (pp. 15–25).
23. Kitti, F., Andrea, S. G., & Andrea, K. K. (2019). Financial Literacy and Indebtedness. In *MultiScience - XXXIII. microCAD International Multidisciplinary Scientific Conference*. http://doi.org/10.26649/musci.2019.104
24. Fodor, K., Lengyel, L., Molnár, L., Szilágyi, R., & Tóthné, K. A. (2018). Duális kompetencia igényfelmérés és prognózis.
25. Lengyel, L., & Fodor, K. (2018). Analysis of effects on life expectancy at birth. In *Multiscience XXXII. MicroCAD International Multidisciplinary Scientific Conference*.
26. Fodor, K. (2017). STATISTICAL TESTS OFMARKET EFFICIENCY. In *„Mérleg és Kihívások" X. Nemzetközi Tudományos Konferencia = „Balance and Challenges" X. International Scientific Conference* (pp. 778–787).