# MODELS FOR CREDIT RISK ASSESSMENT WITH REGARD TO THE FINANSIAL MARKET OF MACEDONIA

**Seadin Xhaferi, PhD. SUT, Faculty of Economics, Republic of Macedonia,**
**Luljeta Sadiku, MBA. SUT, Faculty of Economics, Republic of Macedonia,**
**Murat Sadiku, PhD, SIEU, Republic of Macedonia**

**SUMMARY**
*Estimation of credit risk it's an important topic discussed among academics and policymakers because of the recent crisis which made financial institutions to face increasing risk of crediting. Researchers have projected and studied a large range of models based on expertise, mathematical algorithms or artifice techniques of intelligence. Taking under the consideration this fact and the different circumstances those countries face there is not a unique model of estimation to follow up. In order to have a result which will have a high level of trust we have to take under consideration characteristics of recent levels that will lead us to techniques for risk assessment of models of evaluation. Some important features of empirical performance are described: collecting data for the analysis; discrimination and stability. The study presents performances of metrics, timing, efficiency, overloading and costs. We discuss the importance, advantages and disadvantages of each feature in the evaluation of the disposable methodology.*

**Keywords***: crisis, credit, model, impact, risk*

## 1. INTRODUCTION

The global economic crisis has increasingly infected more and more countries, producing this way unbelievable cases of bankruptcy, and to provoke this way a metaphoric "quake" for the traditional theory. This crisis in comparison with the initial consequences noticed in the US case has a little time-lag, crisis has affected the economy of Macedonia too, beginning by the end of 2008. Its consequences are considerable; such a conclusion comes through every step of our everyday. Most of the researchers are concentrated on the effects of crisis over the loan market. Numerous studies reflect the hypothesis according to which the loan market is the initial point of current crisis. Large amounts of loans granted, that at the end turned out default resulted in lack of liquidity created by current economic crisis. As consequence, banks are trying to find different methods that insure their protection against default cases and the minimization of calculated risk. An alternative could be price on risk basis: clients which are thought most likely to be chosen upon their loans are required to pay higher interest rates, to compensate banks for the increasing risk of default. The defenders of this system believed that it would be unfair in the global aspect to increase the interest rates, because this would also lead to punishment of low risk borrowers. The opponents give social arguments: they take into consideration that the practice has a tendency to generate capital gains for the rich, and repress of the poor class with modest financial resources. Banks in the region carry out conservative policy (loan granting of population and SMEs, with property are closed and do not have any important impact on stock exchange [5]. Lenders argue that interest rates are established correctly having in mind the risk that lander takes into consideration, and that

competition between landers will ensure availability of suitable loans with price for high risk clients. Others think that while their interest rates can be justified with the risks, this is unaccountable for landers to encourage or to allow borrowers with default problems to get loans with high price. Although it is applied in many fields, such as bill market, insurance industry, and stock market, or in many other places of open market, this practice is considered questionable in the case of retail loans. Other ways of credit risk for improvement of evaluation models are represented by taking of new models, means that are capable of proving their efficiency in conditions of crisis. The main question being asked is whether these investments are very important for the case of Macedonia. This study will show that traditional econometric models being used by most of banks in Macedonia could easily fail when economic characteristics change the environment.

Macedonia after 2008 (5.0%), in two pursuant years, will have fluctuation of economic growth (-0,9% in 2009 and 0.7 in 2010), while projection of the Ministry of Finances estimates; for 2011 (3,5%), 2012 (4,5%) and 2013 (5,5%). The rate of unemployment in 2008 was 33,1% while in pursuant years 2009 and 2010 it will be; 31.4% respectively 31,2%.

In general it is accepted that Macedonia is affected by crisis and our goal is to analyze how crisis affects upon the credit market.

## 2. LITERATURE OVERVIEW

Many researchers of credit risk estimation have invented and tried to apply different kinds of models of risk estimation. Each of which is characterized by a much or less complex structure, and we can say that in all studies it's true that the same model gives different results and continually will change with the change of environmental conditions. An example of the model of credit risk estimation is *Data Envelopment Analysis (DEA)* which, unlike the traditional statistical methods (discriminant analysis, logit models and neural networks) doesn't need a priori information [9]. The advantage of this model is clear if we take into account the fact that one of main problems for efficient estimation of credit risk is presented by lack of historical data for the analyzed customers.

One of the oldest methods used is *Linear Discriminant Analysis (LDA),* which was presented for the first time in 1936 by Fisher (in the article *The Use of Multiple Measures in Taxonomic Problems*). The method brings up as main instrument the Linear Discriminant Function (LDF). This is a method used in statistics for calculating the linear combination with characteristics which separate the best two or more layers of objects or events. This resulted combination can be used as linear classifier, more often, for reduction of dimension before the later classification. In the cases of models for credit risk assessment, this instrument can be used for classification of clients into those successful and those unsuccessful. This classification is based upon the characteristics of each profile and process, and is closely related with ANOVA (analysis of variance) and analysis of regression, which also tries to express a dependent variable as a linear combination of other characteristics or measures. However, in other two methods, the dependent variable is a numerical quantity, while for LDA it is categorical variable (good client or bad client).

The main disadvantage of this method consists on the presumption that there exists a linear dependence between variables in input and output, a fact that for the most of the cases as a matter of fact is not true. At the same time i.e. class and covariance are necessary, but unknown. In most cases they are estimated through installation of training. It should be noted that although the estimations of covariance could be considered optimal, this doesn't mean that the calculated discriminant with these values results as optimal in any sense, even if the assumption of normally allocated classes is correct. Another field where LDA is applied on a large scale is in bankruptcy forecast. This field is very similar to assessment risk, the goal is to determine which firms go to bankruptcy and which will survive. However limitations

included in unknown consistencies of accounting reports for normal distribution assumptions of LDA, Edward Altman model of 1968 still applies in many applications.

Researchers have also implied the use of special tools of artificial intelligence: *Artificial Neural Networks (ANN) - Probabilistic Neural Networks* or *Multi-layer feed-forward nets*. It has been noticed that registered assessments of good results when it's about customers that have failed and correctly distinguished (this is very important point: in order to have a high performance in clear classification of bad clients rather in clear classification of good clients (risk of not granting a loan to a client is evaluated with higher risk than real one)). The hardest aspect of implementation of neural networks in credit assessment process is the difficulty of explaining the algorithm that gives the decision for accepting or refusing of certain customers. DEA requires only input and output data in order to determine the calculation of the credit result.

Another method which was introduced for the first time in 1984 is CART (*Classification and Regression Tree*) method and it consists of building a maximum tree that contains all the units of training determined and could be divided into several trees of which one is chosen most efficient through the method of inter-validity.

In 1991 MARS (*Multivariate Adaptive Regression Splines*) model was introduced for the first time, bringing the advantage of being combined with neural networks. The algorithm has two steps: first, there are build numerous basic functions, of which some will be eliminated in a manner of smaller contributions (the cross validation method). *Case Based Reasoning* (CBR) is a model with a goal of learning from actual failures and tries to build a model upon which should be based the classification of new clients. According to Agnar Aamodt and Enric Plaza [2], the argumentation of the case with base has been formalized for purposes of computer reasoning as a process in four steps:

1. Having into consideration a real problem (decision whether to grant or no a loan to a potential client), we take the relevant cases from the memory for its solution (of all client profiles that are similar to applicant). A particular case makes one problem, its solution and typically, the description how the solution run.
2. The map of solution from previous case for the actual problem. (If the previous customer was granted a loan and he was successful in its repayment, than new client should also be accepted). This might include adjustment of solution according to the new situation, as the new client has few characteristics which are different from the other case. For instance, the environment conditions could change: even though client could be considered to be in a situation of evidently financial stability, the economic crisis lowers his power and should be taken into consideration.
3. Having entire solution for the previously assumed situation, test new solution in real world terms (or a simulation) and, if its necessary review. Suppose the client didn't accomplish his commitments, the old solution should be modified in a manner to prevention future mistakes.
4. After the solution the problem was successfully adjusted into the planned problem, save the experience of the result into the memory as a new.

The RCSM (*Reassigning Credit Scoring Model*) model is an algorithm made in two steps: in classification of «good» applicants or the «bad» cases and redistribution of «good» cases that have been previously mistakably classified as «bad», in a way to prove the efficiency of the model, researchers have attempted to apply this algorithm for the data of ID cards [7]. During the first phase the algorithm has used MARS in a way to reduce the number of knots given by ANN (Artificial Neural Networks) and to simplify the networks. The simplified neural network (backpropagation network characterized by a unique layer input, a hidden layer and an output layer) is used for the classification of applicants into good or bad customers. In order for customers to be considered "bad" will be determined through CBR method, through

identification of existing similarities between them, from one side, and good practical cases, on the other hand. If the new applicants profile is more closely to successful cases rather than those of failures, the person is classified as "good one" and the loan is granted. The similarity is measured through "distance" between two cases, using the nearest neighbor access. Other examples of risk evaluation of credit models determined in scientific literature are:

- Multi Criteria Decision Making
- Recursive Partitioning Algorithm
- Mathematical Programming Approaches
- Logistic Regression Analysis (LRA)
- fuzzy algorithms
- Multivariate Conditional Probability Model
- k-N

## 3. METHODOLOGY OF CREDIT RESULT

The process of result for a loan applicant that should be applied for choosing a risk assessment method concludes of the following steps:

1. The particular observation data are selected (selection of those loan applicants for which is likely to find the values for all the characteristics that could impact into the risk)
2. Identification of possible characteristics for each candidate (factors such as: income, age, shelter, marital status, etc.)
3. Selection of later characteristics (instruments: regressions – econometric tools used for testing if selected variables are important for the result – and scope experts). The result is a group of data that contains the most representative indicators. It should be noted that indicators can be input indicators (former loans, liquidities, buildings, etc.), or output indicators (in the case of companies, can be taken into consideration whether the company is financed by its own resources or its ability to pay interest from own revenues).
4. Calculation of credit result as a result of the chosen model.
5. Validity of the results with the aid of regression (the selected indicators represent independent variables and the result is a dependent variable), the discriminant analysis (in order to determine the quantity of classification by this method) and the cases of real world failures. It is interesting to analyze whether results correspond with the results given by other methods of classification.
6. Selection of final method of credit result.

## 4. THE POSITION OF CREDIT MARKET IN MACEDONIA

According to statements of National Bank of Macedonia (NBRM), loan application has become more restricted during the end of 2008 and in the beginning of 2009. For example, the last quarter of 2008 is marked by increasingly rigidity of loan institutions, when it comes about granting loans to their applicants. This behavior can be explained in a way that the economic environment has been impacted by recession, so loan institutions and other organizations of financial system have adjusted their supply for new economic conditions. Statistical analysis shows the fact that credit terms have perceived their tendency in sense that they become increasingly more rigid, while trend has proven to be more vigilant by the end of 2008 and the beginning of 2009. The risk related to companies is perceived as growing, including corporations. This evolution has to do with reactions noticed in the "Eurozone", where this reaction is negative related to loan risk that has begun to develop by the end of 2007.

If we will analyze the evolution of interest rated for loan and deposits, we could easily encounter the fact that interest rates for current loans in euro have considerable downfall during the first months of 2009, while interest rates for current deposits have been increased

in general, so the bank policy in general has encouraged people to bring liquidity in bank. (Figure 1)
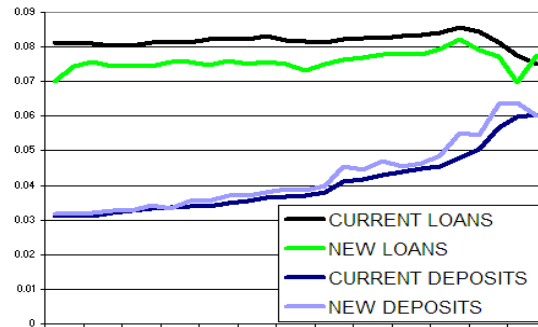


Figure 1. Evolution of interest rates for loans and deposits denominated in euros
Source: Analysis of the Yearly statement of NBM BPM 2002-2009

On the other side, the same analysis can be conducted for interest rates of loans and deposits denominated in denars. (Figure 2). It can be easily noticed that new loans and deposits have interest rates higher than current ones. This can be viewed as a monetary policy with aim to encourage customers trust national currency and to increase its use in their financial transactions, in order to determine depreciation of euro in relation to national currency.
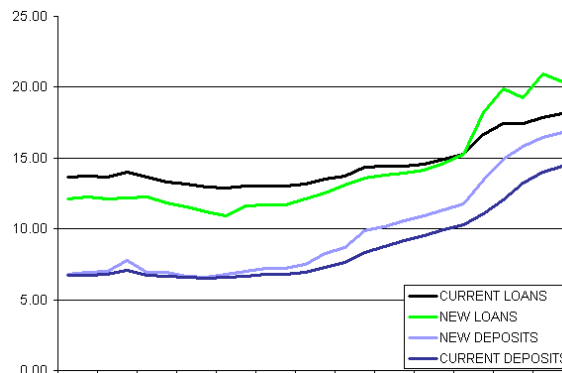


*Figure 2. Evolution of interest rates for loans and deposits in denars*
Source: Analysis of Yearly Statement of NBM 2002-2009

At the same time, one of the most important economic events is approving the loan that Macedonia has requested from International Monetary Fund. Many analysts think that this amount of money will bring couple of changes in credit market credit policies. A huge part of loan has its last destination in National Banks reserves. Therefore, National Bank will gain from a "tightening belt" that would give the opportunity to rest to the minimal required reserves that it demands from other banks.

The minimal required reserves represent an instrument that National Bank of Macedonia uses to control the number of granted loans. Reducing the minimal required reserves, an amount of liquidities are allowed to return into the financial market.

A conclusion that can be withdrew after the analysis of different types of statements is that the economic environment has changes as a consequence of recession, and financial institutions must redesign their loans, assessment models for finding more efficient tools for their application. The remaining part consists of a simplified econometric model with the intention of proving the effects of crisis over the forecasted accuracy of instruments.

## 5. MODEL

The banking sector in Macedonia is characterized by the use of traditional statistical methods rather systems based on experts of neural networks or case based reasoning, we have used traditional econometric tools (multiple regression) in order to build our research model. The main source of data is presented by financial statements published by different companies, statements that available in the Ministry of Finances and some of information from the stock market [10].

Both data contain same characteristics of the same customer potential registered in two different moments: in the end of second quarter, when the economic impact of recession in Macedonia was not felt, and in the fourth quarter of 2008, when the recession was admitted. Therefore, the data reflect status in the middle of 2008 and in the end of 2008. Initial data include many variables from which those more representative will be selected by a suitable technique: total debt, return, net income, debt rate (total debt / capital), liquidity index (current assets / current debt), economic return (net income / return or total income), and PER (closing price of bills / earning per share). Authors suppose that net income is a determinant which could be used for estimating the financial stability of company and its ability to repay its debts, including loans. Therefore the goal is to determine which variables to have significant impact over the net income and which are their coefficients. This is the first step toward building a credit model, marking as different results that are related with the loan variables, and marking assessed models by the regression.

Prior to regressions, it is made a descriptive analyze for comparison of values of selected variables in the middle and in the end of 2008.

A general overview would be that overall debts are increased during the second half of 2008. This is explained through the effects of economic crisis that are understood during this period and can be considered as indicator of financial instability. Some former studies show the impact of process of selection over the quality of results gained through application of risk evaluation model. Each time an attempt for building a new assessment model is made, one addresses the problem chosen by a numerous potential variables (financial performance of client, environment attributes, etc.). The goal is to select the optimal number (necessary and sufficient) from characteristics that are used for building a client's profile.

Risk of including unimportant and too many variables altogether with relevant is quiet high in the field of credit risk assessment, as the quantity of data that a financial institution gets from a potential client is quiet big and it's difficult to create linear relationship between variables. Therefore, including many attributes that can give false results as well as some attributes will not bring new information and their excess may lead to unimportant assessments, while for some variables may bring insufficient information for a good decision.

There exist three regular approaches for selecting attribute – *filter-based, wrapper-based and embedded-based*. The filter bases approach selects attributes before the algorithm classification is applied. Therefore, this approach is general attribute of the method for independent selection of classification algorithms. On the other hand, wrapper-based approach includes an objective classifier as a black box for performance evaluation. In other words, an intensive assessment of calculations is committed many times in function of candidate for selecting important attributes. Embed approach is the undivided ability of algorithm classification, i.e., attribute selection is made naturally as a part of learning. Many modern algorithms that combine different classical methods are implemented lately. For example[6], building a base of selections bases upon evolving approach with new criteria for data mining and applying it into the loan granting data.

Authors have attempted to build different regression models and variables with elimination and inputting in accordance with their coefficient importance. In order to conduct regressions, the data are evaluated under different scenarios, applying less usual squares.

The most important regression model is with net income as dependent variable and independent variable selected capital, total debt and return. Regression model is applied for the both groups of data. Here an observation should be made: it can be seen that p-value are still quiet high and their values might have been reduced. This is because of the fact that regression model includes only three independent variables, while there are other that have impact, also the opportunity for loan repayment. Adding other attributes p-value will decrease and the model efficiency will increase.

*Table 1. Results for half of 2008 and end of 2008 regressions of net income*

| Variable | Coefficient half of 2008 | p-value | Coefficient end of 2008 | p-value |
|---|---|---|---|---|
| CAPITAL | 0.000751 | 0.0000 | 0.001783 | 0.0000 |
| TURNOVER | 0.051040 | 0.0157 | -0.141287 | 0.0195 |
| TOTAL DEBTS | -0.096196 | 0.0010 | -0.043378 | 0.0110 |

*Source: Author's estimation*

Both regression models have high $R^2$ values (0,93 and 0,91) and low p-values of F statistics (0,0000), facts that prove that models are built accordingly. The most important overview that could be made analyzing the results is that relatedness of selected variables decreases during the second half of the year. Capital remains the most important attribute for both groups of data, while p-value for return increases from 0,0157 to 0,0195 and the one for total debt increases from 0,0010 (2008) to 0,0110 (2008). In concluding of factors with inclination to losing their importance exists necessity for reanalyzing and potentially redesign the named loan models.

## 6. CONCLUSION

This paper analyses current situation of the credit market in Macedonia and presents different kinds of results that prove manifestimin of economic crisis and its impact over credit market. The study was realized in order to prove actually the efficiency downfall that characterizes results of traditional credit methodology. At the same time the paper presents a summary of bibliography naming some of most important advantages and disadvantages of results of the existing credit model.

Main conclusion is that the quality of gained results is directly impacted by environmental factors. As a consequence, models should consider variables like quantity and stability, and the power of economic environment, and the main goal for building such a model. Another interesting attribute is dynamic character of variable values. In time, the customer profile could change and the initial evaluation of risk might lose its stability. Therefore, we should also find a method for measuring this dynamic aspect and to include its value in risk model.

## 7. REFERENCES

[1] Atish P. Sinha and Huimin Zhao, "Incorporating domain knowledge into data mining classifiers: An application in indirect lending", Decision Support Systems, Vol. 46, No. 1, 2008, pg. 287-299

[2] Agnar Aamodt and Enric Plaza, "Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches", Artificial Intelligence Communications, Vol. 7, No. 1, 1994, pg. 39-52.

[3] B. Griffiths and M. J. Beynon, "Expositing stages of VPRS analysis in an expert system: Application with bank credit ratings", Expert Systems with Applications, No. 29, 2005, pp. 879–888

[4] Ben-David Arie and Frank Eibe, "Accuracy of machine learning models versus ''hand crafted" expert systems – A credit scoring case study", Expert Systems with Applications, Vol. 36, No. 3, 2009, pg. 5264-5271

[5] Besa Xhaferi, Seadin Xhaferi „ Financial crises in Albanian Area" 4th Annual metting of Alb-Shkenca Institute, August 31-2 September 2009, pg.237-244

[6]   Chia-Ming Wang and Yin-Fu Huang, "Evolutionary-based feature selection approaches with new criteria for data mining: A case study of credit approval data", Expert Systems with Applications,  Vol. 36, 2009, pg. 5900–5908

[7]   Chun-Ling Chuang and Rong-Ho Lin, "Constructing a reassigning credit scoring model", Expert Systems with Applications , Vol. 36, No. 2, 2009, pg. 1685-1694

[8]   Hussein Abdou, John Pointon and Ahmed El-Masry, "Neural nets versus conventional techniques in credit scoring in Egyptian banking", Expert Systems with Applications, Vol. 35, 2008, pg. 1275–1292

[9]   Jae H. Min and Young-Chan Lee, "A practical approach to credit scoring", Expert Systems with Applications, Vol. 35, 2008, pg. 1762–1770

[10] Makedonska berza na akcii, http://www.mse.org.mk

[11] Ministerstvo za finansii na RM www.finance.gov.mk

[12] National Bank of Macedonia, http://www.nbrm.gov.mk

[13] S. Piramuthu, "On preprocessing data for financial credit risk evaluation", Expert Systems with Applications, Vol 30, No. 3, 2006, pg. 489-497

[14] S. H. Huang, "Dimensionality reduction in automatic knowledge acquisition: A simple greedy search approach", IEEE Transactions on Knowledge and Data Engineering, Vol. 15, No. 6, 2003, pg. 1364–1373.