ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 7, January 2015

Comparative Analysis of Methods of Banks Bankruptcy Risk Forecasting under Uncertainty

Ovi Nafas Aghaei Agh Ghamish (Iran), Yuriy Zaychenko
Kiev Polytechnic Institute, Kiev, Ukraine

Abstract- The problem of banks bankruptcy risk forecasting under uncertainty is considered. For its solution novel methods: fuzzy neural networks TSK, ANFIS and Fuzzy group method of data handling (FGMDH) are suggested. For estimation of their efficiency the numerous experiments were carried out using as input data financial indices of 256 leading European banks according to International standard IFRS. For estimation of the efficiency of the suggested methods, they were compared with conventional methods: ARMA, logit and probit models. During experiments, the most informative financial indices for banks bankruptcy forecasting were detected.

***Index terms*- banks bankruptcy risk, forecasting, fuzzy neural networks, fuzzy GMDH, logit-model, probit-model.**

I. INTRODUCTION

The problem of bank ruptcy fore casting is of great importance in modern economy. As bank system is a fundament of any economy timely determination of possible bankruptcy risk is extremely significant.

Timely detection of features of coming bank bankruptcy enables the top managers to adopt urgent measures to prevent possible bankruptcy and improve the financial bank state. Now there are a number of techniques and methods for banks financial state analysis and determination of proper bank rating: Web Money, CAMEL, and Moody's S&P [1]. But their common drawback is that they work only with complete and reliable data. At the same time many bank managers intend to hide real financial state that to attract shareholders and to receive loans and credits. The main goal of this paper is to develop and investigate novel methods of banks bankruptcy risk forecasting which may work under uncertainty with incomplete and unreliable data.

Besides, the other goal of this investigation is to determine which factors (indices) are to be used in forecasting models that to obtain results close to real data. Therefore, we used a set of financial indices (factors) of European banks according to the International accountant standard IFRS. We collected annual financial indices of about 300 European banks in 2004-2008 years, preceding the start of crisis of bank system in Europe in 2009 year. The data source is the information system Bloomberg. The resulting sample included the reports only from the largest European banks as system Bloomberg contains the financial reports only from such banks. So the following indices out of European banks annual reports were taken for analysis (which is presented in system Bloomberg):

- Assets;
- Liabilities;
- Cost;
- Income;
- Net Financial Result (Profit/Loss);
- Loan to Deposit Ratio;
- Liquid Asset Ratio;
- Asset Interest Yield ;
- Break Even Yield;
- Net Interest Margin;
- Return on Assets (ROA)
- Leverage Multiplier
- Return on Equity (ROE);
- Capital Adequacy Ratio;
- Loan Loss Provision;
- Total Loans;
- Risk Adjusted Margin;
- Overhead Burden Ratio;
- Productivity Ratio;
- Cost/Income Ratio;
- Asset Yield;
- Profit Margin;
- Capital Assets Ratio;
- NPL to Total Loans.

The output of fore casting models for European bank swere two values:

1, if the considerable aggravation of financial bank state is not expected in the nearest future;

0, ifnegativefinancialresultsinbank (losses) are expected in the nearest future.

The above mentioned output data concerning banks state at the forecasting moment (after 2008 year) were also taken out of Bloomberg system. In particularly, the list of banks was collected which had losses in 2009 year

(Net Financial Result <0). For correct utilization of input data they were normalized in interval [0,1].

II. APPLICATION OF FUZZY NEURAL NETWORKS FOR EUROPEAN BANKS BANKRUPTCY RISK FFORECASTING

Input data were the financial indices of the largest European banks data obtained from system Bloomberg. The period for which the data were collected was 2004-2008 years. The possible bank ruptcy was analyzed in 2009 year. The indices of 165 banks were considered

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 7, January 2015

among which more than 20 banks displayed the worsening of the financial state in that year. Fuzzy neural networks and Fuzzy Group Method of Data Handling (FGMDH) were used for bank financial state forecasting.

In accordance with the goal the investigations were carried out for detecting the most informative indices (factors) for financial state analysis and bankruptcy forecasting. Taking into account incompleteness and unreliability of initial data fuzzy neural networks (FNN) ANFIS and TSK were suggested for bankruptcy risk forecasting. [3]

The main advantages of FNN are the following: the possibility of work under fuzzy and incomplete data; the possibility to use expert knowledge in a form of fuzzy rules [3, 4].

After performing a number of experiments the data set of financial indices was found using which FNN gave the best forecast. These indices are the following:

- Debt/Assets = (Short-term debt + Long-term debt)/Total Assets
- Loans to Deposits
- Net Interest Margin (NIM) = Net Interest income / Earning Assets
- Return on Equity (ROE) = Net Income/Stockholder Equity
- Return on Assets (ROA) = Net Income/Assets Equity
- Cost/Income = Operating expenses / Operating Income
- Equity/Assets = Total Equity/Total Assets

The series of experiments were carried out for determining the influence of the number of rules and period of data collection on forecasting results.

In the first series of experiments FNN TSK was used for forecasting.

Experiment №1.

- Training sample = 115 banks of Europe ;
- Testing sample= 50 banks; number of rules = 5
- Input data period = 2004 year.

The results of FNN TSKapplication are presented in Table1.

Table1.Forecasting results of TSKusing data 2005year

Results	
Total number of errors	8
%% errors	16%
First type of errors	0
Second type of errors	8

Experiment №2

- Training sample = 115 banks of Europe;
- Testing sample= 50 banks; number of rules = 5
- Input data period= 2005

The results of FNN TSK application are presented in Table2

Table 2.Forecasting results of TSK using data 2005year.

Results	
Total number of errors	7
%% errors	14%
first type of errors	0
Second type of errors	7

Experiment №3.

- Training sample = 115 banks of Europe;
- Testing sample= 50 banks; number of rules = 5
- Input data period = 2006

The results of FNN TSK application are presented in Table3

Table 3. Forecasting results of TSK using data 2006 year

Results:	
Total number of errors	5
%% errors	10%
first type of errors	0
Second type of errors	5

Experiment №4

- Training sample = 115 banks of Europe;
- Testing sample= 50 banks; number of rules = 5
- Input data period = 2007

The results of FNN TSK application are presented in Table4.

Table 4. Forecasting results of TSK using data 2007 year

Results:	
Total number of errors	1
%% errors	2%
first type of errors	0
Second type of errors	1

Further, the similar experiments were performed with FNN ANFIS while the period of data collection varied since 2004 to 2007 year.

The experiments for detection of influence of rules number on forecasting results were also carried out. The corresponding results are presented in Table 5 for FNN TSK. The results for FNN ANFIS are presented in table 6

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 7, January 2015

which display the influence of data collection period on forecasting accuracy.

Table 5.Forecasting results for FNNTSK versus number of rules and data period

Experiment/number of rules	Total errors number	%% of errors	Number of the 1-st type errors	Number of the 2-nd type errors
2004 - 5	8	16%	0	8
2005 – 5	7	14%	0	7
2006 – 5	5	10%	0	5
2007 – 5	1	2%	0	1
2004 - 10	8	16%	0	8
2005 – 10	8	16%	1	7
2006 – 10	11	22%	7	4
2007 – 10	4	8%	0	4

Table 6 Forecasting results for FNN ANFIS versus number of rules and data period

Experiment/ number of rules	Total errors number	%% of errors	Number of the 1-st type errors	Number of the 2-nd type errors
2004 - 5	8	16%	0	8
2005 – 5	8	16%	1	7
2006 – 5	8	16%	4	4
2007 – 5	4	8%	0	4

After analysis of these results the *following conclusions* were made

1. FNN TSK gives better results than ANFIS while forecasting the bankruptcy risk for European banks.
2. The best input variables (indices) for European banks bank ruptcy risk forecasting are the following:
   a. Debt/Assets = (Short-term debt + Long-term debt) / Total Assets
   b. Loans to Deposits
   c. Net Interest Margin (NIM) = Net Interest income / Earning Assets
   d. Return on Equity (ROE) = Net Income / Stockholder Equity
   e. Return on Assets (ROA) = Net Income / Assets Equity
   f. Cost/Income = Operating expenses / Operating Income
   g. Equity/Assets = Total Equity / Total Assets
3. Input data collection period (forecasting interval) makes influence on forecasting results.
4. The increase of rules number does not influence on results.

III. THE APPLICATION OF FUZZY GMDH FOR BANK FINANCIAL STATE FORECASTING

While performing experiments fuzzy Group Method of Data Handling (FGMDH) was also applied for financial state of European banks forecasting. As it known fuzzy GMDH enables to construct forecasting models using experimental data automatically without expert [3, 4]. The additional advantage of FGMDH is possibility to work with fuzzy information [3].

As the input data in these experiments the same indices were used like experiments with FNN TSK. The result of method application is an output value which should drop in a certain interval. In our case the output value should be transformed to value 1 (where 1 = is bank with good financial state) or -1 (-1 = bank with bad financial state). As a threshold between two classes's we chose the middle of the output interval.

In Table 7 forecasting results are presented independence on input data period for FGMDH

Table7. Comparative analysis of forecasting results for FGMDH

Input data period	Total number of errors	%% of errors	Number of the first type of errors	Number of the second types of errors
2004	7	14%	0	7
2005	6	12%	1	5
2006	4	8%	1	3
2007	2	4%	0	2

If we compare the results of FGMDH application with theresultsofFNNTSKonecanseethatneuralnetworkgivesbetterresultsontheshortforecastinginterval (one year) while fuzzy GMDH gives better results on greater intervals (2 or more years). So FGMDH has undoubted advantages in long-term forecasting as compared with FNN. In table 8 thecomparativeresultsofapplicationofdifferentmethodsfor bankruptcyriskforecastingarepresented

Table 8. Forecasting results of different fuzzy methods

Method (period)	Total number of errors	%% of errors	Number of the first type errors	Number of the second type of errors
ANFIS (1 year)	4	8%	0	4
TSK (1 year)	1	2%	0	1

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 7, January 2015

FGMDH (1 year)	2	4%	0	2
ANFIS(2 years)	8	16%	4	4
TSK(2 years)	5	10%	0	5
FGMDH (2 years)	4	8%	1	3

IV. APPLICATION OF LINEAR REGESSION AND PROBABILISTIC MODELS

Regression models. For analysis of fuzzy methods efficiency at the problem of bankruptcy risk forecasting the comparison with crisp methods: the regression analysis of linear models was performed. As input data the same data were used which were recognized optimal for FNN? Additionally, the index Net Financial Result was also included in the input set. This index makes great impact on forecasting results. Thus, input data in this experiments were 8 financial indices of 256 European banks according to their reports based on International standards

- Debt/Assets – X1
- Loans/Deposits – X2
- NetInterestMargin – X3
- ROE (ReturnonEquity) – X4
- ROA (ReturnonAssets) – X5
- Cost/Income – X6
- Equity/Assets – X7
- Net Financial Result – X8

The input data were normalized before the application. The experiments were carried out with full regression ARMA model, which used 8 variables and shortened models with6 and 4 variables.

Each obtained model was checked on testing sample consisting of 50 banks. The comparative forecasting results for all ARMA models are presented in Table 9.

Table 9.Comparative analysis of ARMA models

Input data	Testing sample	I type errors	II type errors	Total number of errors	% of errors
All variables (8)	50	5	4	9	18%
6 variables	50	5	4	9	18%
4 variables	50	5	4	9	18%

As one may see in Table 9, the application of all types of simple line regression models gives the same error18%, that is much worse than application of fuzzy neural networks.

Logit-models. Further the experiments were performed using logit-models for bankruptcy forecasting. The training sample consisted of 206 banks and the testing sample – of 50 banks.

The first one was constructed linear logit-model using all the input variables. It has the following form (estimating and forecasting equations):

I_Y = C(1) + C(2)*X1 + C(3)*X2 + C(4)*X3 + C(5)*X4 + C(6)*X5 + C(7)*X6 + C(8)*X7 + C(9)*X8

Y = 1-@CLOGISTIC(-(C(1) + C(2)*X1 + C(3)*X2 + C(4)*X3 + C(5)*X4 + C(6)*X5 + C(7)*X6 + C(8)*X7 + C(9)*X8))

The next constructed model was a linear probabilistic logit-model with 6 independent variables. It takes the form

I_Y = C(1) + C(2)*X1 + C(3)*X3 + C(4)*X4 + C(5)*X5 + C(6)*X6 + C(7)*X7

Y = 1-@CLOGISTIC(-(C(1) + C(2)*X1 + C(3)*X3 + C(4)*X4 + C(5)*X5 + C(6)*X6 + C(7)*X7))

As the investigations have shown, the indices Debt / Assets (X1) and Net Interest Margin (X3) insignificantly influence on the forecasting quality so they were excluded from the list of inputs in the next experiment. The next constructed logit-model was based as the previous one on 206 banks data and consisted of 4 variables. It has the form: (estimating and forecasting equations)

I_Y = C (1) + C (2)*X4 + C (3)*X5 + C (4)*X6 + C (5)*X7

Y = 1-@CLOGISTIC(-(C(1) + C (2)*X4 + C (3)*X5 + C (4)*X6 + C (5)*X7))

The final table including the forecasting results of all the logit-models is presented below (Table 10)

Table 10.Comparative analysis of logit-models

Input data	Testing sample	I type errors	II type errors	Total number of errors	% of errors
All variables	50	6	2	8	16%

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 7, January 2015

(8)					
6 variables	50	6	2	8	16%

Probit-models.The next experiments were carried out with probit- models. The first constructed model was the linear prod bit –model based on 206 banks using all the input variables. It had the following form

I_Y = C(1) + C(2)*X1 + C(3)*X2 + C(4)*X3 + C(5)*X4 + C(6)*X5 + C(7)*X6 + C(8)*X7 + C(9)*X8
Y = 1-@CNORM(-(C(1) + C(2)*X1 + C(3)*X2 + C(4)*X3 + C(5)*X4 + C(6)*X5 + C(7)*X6 + C(8)*X7 + C(9)*X8))

As the experiments had shown the inputs Net Interest Margin (X3) and Net Financial Result (X8) very weakly influence on the results and they were excluded in the next experiments. The next constructed probit-model included 6 variables and had the following form:
I_Y = C(1) + C(2)*X1 + C(3)*X2 + C(4)*X4 + C(5)*X5 + C(6)*X6 + C(7)*X7
Y = 1-@CNORM(-(C(1) + C(2)*X1 + C(3)*X2 + C(4)*X4 + C(5)*X5 + C(6)*X6 + C(7)*X7))

Further in this model were excluded insignificant variables Debt / Assets (X1) and Loans / Deposits (X2) and as result linear probit-model with 4 variables was obtained.

Each of the constructed probit- models was checked on the test sample of 50 banks. The comparative results of application of all probit- models are presented in the Table 11

Table 11. Forecasting results of probit-models

Input data	Testing sample	I type errors	II type errors	Total number of errors	% of errors
All variables (8)	50	5	2	7	14%
6 variables	50	5	2	7	14%
4 variables	50	6	3	9	18%

As one may see from the table 11, the application of all the probit-models gives relative error14-18% that is much worse than results obtained by fuzzy neural networks. It is worth to mention the decrease of model forecasting quality after exclusion of insignificant variables. In particularly, after exclusion of variables Debt / Assets and Loans / Deposits error value has increased from 14% to 18%. The low quality of constructed logit and probit-models testify of the absence of direct dependence between input and output model variables.

V. CONCLUDING EXPERIMENTS

In the final series of experiments investigations and detailed analysis of various methods for forecasting bankruptcy risk were performed. The following methods were investigated:

FNN ANFIS; FNN TSK; FGMDH; Regression models; Logit-models; •Probit-models.

Period of input data was - 2007 year (for 1 year before possible bankruptcy).

Comparative analysis of all the considered forecasting methods is presented in the Table 12.

Table 12. Comparative analysis of methods for banks bankruptcy forecasting

Method (period))	Total number of errors	%% of errors	I type errors	II type errors
ANFIS	4	8%	0	4
TSK	1	2%	0	1
FGMDH	2	4%	0	2
ARMA	9	18%	4	5
LOGIT	8	16%	2	6
PROBIT	7	14%	2	5

As one may readily see from this table fuzzy methods and models have much better results than crisp methods: ARMA, logit-models and probit-models. When forecasting by one year ahead fuzzy neural networks give better results than FGMDH. But when forecasting for longer intervals (several years) FGMDH is the best method.

1. Different methods and models were considered and investigated for financial state forecasting of European banks. The following methods were considered:

- Fuzzy neural network ANFIS;
- Fuzzy neural network TSK;
- Fuzzy GMDH;
- Regression models;
- Logit - model;
- Probit-model.

As input data the financial indices of European banks were used

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 7, January 2015

2. As the experiments have shown the fuzzy methods FNN TSK, ANFIS and FGMDH give much better forecasting results than crisp methods: regression models; logit - model; probit-model.

3. It was detected that among FNN network TSK gives better results than FNN ANFIS.The increase of fuzzy rules doesn't improve the forecasting results.

4. While comparing different fuzzy methods it was established that fuzzy neural networks give better forecasting results using more fresh data, that is, by short-term forecasting while FGMDH is better by long-term forecasting (2 or more years**)**.

5. While investigating crisp methods it was established: the worst results (maximal error rate) was obtained by conventional linear regression (18%) while the application of probabilistic logit- and probit –models allows to decrease the error rate to 14-16%.

REFERENCES

[1] Rating estimate of commercial bank. http:// www. nbuv. gov.ua/Soc_Gum/ Ekpr/2009_25/ mescheryakov/htm (ukr).

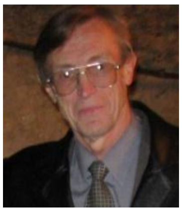
[2] The bank rating construction using the technique of calculation of aggregated multi level indices of bank financial state/ (rus) http:// www. Credit-rating.ua/ ru/analytics/ analytical-articles/.

[3] Zgurovsky M.Z., Zaychenko Yu.P. The fundamentals of computational intelligence. – Kiev: Publ. House « Naukova Dumka»,. 2013.-406 p. (rus).

[4] Zaychenko Yu. P.Fuzzy models and Methods in Intelligent Systems. Kiev. Publishing House"Slovo", 2008.-344p. (Rus).

AUTHOR BIOGRAPHY

Ovi Nafas Aghaei Agh Ghamish (Iran)- master of sciences., post-graduate student, National Polytechnic Institute 03056, Kiev-56, Ukraine. e-mail: ovinafas@yahoo.com field of scientific interests: artificial intelligence, fuzzy neural networks and their applications Author of 10 works.



Yuri Zaychenko – Professor, doctor of technical sciences, Institute for applied system analysis, NTUU "KPI", 03056, Kyiv, Peremogy pr. 37, Corpus 35; e-mail:baskervil@voliacable.com . Major Fields of Scientific Research: Information systems, Fuzzy logic Decision-making theory, Computationalintelligence, author of 15 books and more than 100 scientific works.