Data Mining Application for Anti-Crisis Management

Nafisa Yusupova
Department of Computer Science and Robotics
Ufa State Aviation Technical University
Ufa, Russia
yussupova@ugatu.ac.ru

Gyuzel Shakhmametova
Department of Computer Science and Robotics
Ufa State Aviation Technical University
Ufa, Russia
shakhgouzel@mail.ru

***Abstract*—The paper is devoted to data mining as applied to anti-crisis management, in particular to bankruptcy monitoring. The decision support system for bankruptcy monitoring, based on the intelligent information technologies (data mining, expert systems) is considered. The main stages of data mining technology applied to anti-crisis management are described in detail. The results of data mining implementation are estimated.**

Keywords-data mining; forecasting; decision support system; anti-crisis management; bankruptcy monitoring

I. INTRODUCTION

Anti-crisis management is a process of preventing or dealing with the crisis the enterprise (company) is in. This definition incorporates two components of the anti-crisis management: prevention of the crisis that has not yet come and overcoming of the crisis that has already come [1]. Interpretation of the anti-crisis management may be different depending on the state of the enterprise. In case the enterprise is in a stable position, the anti-crisis management consists in monitoring which aims at the data retrieval and processing and forecasting the enterprise performance. In case the enterprise is in an unstable position, i.e. there is a bankruptcy risk, the anti-crisis management assumes a form of regulation which is a set of measures to protect the enterprise in crisis situations and prevent bankruptcy. When an enterprise is in crisis it is necessary to resort directly to anti-crisis management as an instrument to pull the enterprise out of the crisis or to prepare it to be wound up or reorganized.

The carried-out analysis of methods of enterprises bankruptcy predicting and the analysis of possibilities of well known IT-decisions in this field showed that the development of a decision support system for bankruptcies monitoring is needed [7]. The data required for anti-crisis management are semi-structured data in the majority of cases and therefore the application of intelligent information technologies is necessary [4][6].

The researches into anti-crisis management in the field of bankruptcy monitoring have been carried on for a long time and can be found in the papers of many scientists, as well as in the IT-decisions. These problems are considered in detail in [7]. The distinguishing feature of this research is the possibility of fraudulent bankruptcy indications forecasting at its early stages when it is possible to take preventive measures.

There are two main approaches to enterprise bankruptcy forecasting in modern business and financial performance practice [1]. Quantitative methods are based on financial data and include the following coefficients: Altman Z-coefficient (the USA); Taffler coefficient (Great Britain); two-factor model (the USA); Beaver metrics system and the others. A qualitative approach to enterprise bankruptcy forecasting relies on the comparison of the financial data of the enterprise under review with the data of the bankrupt business (Argenti A-account, Scone method). Integrated points-based system used for the comprehensive evaluation of business solvency includes the characteristics of both quantitative and qualitative approaches. An apparent advantage of the methods consists in their system and complex approach to the forecasting of signs of crisis development, their weaknesses lie in the fact that the models are quite complicated in making decisions in case of a multi-criteria problem, it is also worth mentioning that the taken forecasting decision is more subjective. Modern information technologies, Data Mining in particular, provide ample opportunities for solving the problems of anti-crisis management [10].

Financial and economic application software that is available on the market nowadays is quite varied and heterogeneous. The necessity to develop such software products is dictated by the need of enterprises to promptly receive management data in due time and to forecast the signs of crisis development. To one extent or another, tools for anti-crisis management are available in a number of ready-made IT-decisions [7]. But the data analysis in many software products actually consists in providing the necessary strategic materials, while software products should meet the increasing needs such as analysis and forecasting enterprise financial performance in the next period of report.

The authors of the study aim to develop models and algorithms based on intelligent technologies for the detection of the crisis state of the enterprise while still in its early stages for the timely changes of the development strategy of the enterprise, which will increase stability and economic independence of the enterprise, as well as reduce the impact of the human (subjective) factor on making important management decisions. Data mining application in the decision support system for anti-crisis management is

 ISBN: 978-1-61208-227-1

discussed in this article on the example of monitoring bankruptcies.

The second section deals with the general scheme of the proposed decision support system. The main modules of this system are the data mining module and the expert system. The third section considers in detail the application of data mining technology to forecast the financial performance of the enterprise (company). Computer-based processing of the data on the basis of the proposed system is presented in the fourth section. Forecasting of the financial performance of the enterprise is made by the analytical platform. The stages of the implementation processes are also described. Section 5 of the article presents the results of the practical application of the system which allows assessing the effectiveness of the system.

II. Decision Support System for Monitoring Bankruptcy

The major aspect of the bankruptcy monitoring problem is the analysis and identification in good time of the signs of fraudulent bankruptcies [1].

The basis of the whole complex of techniques for the decision support system (DSS) is legally approved methodical instructions on accounting and analysis of enterprise' financial position and solvency so as to group the enterprises depending on the level of risk of bankruptcy, as well as techniques for the identification of the signs of fictitious and deliberate bankruptcy. These techniques are currently used by auditors and arbitration managers.

To develop the decision support system for monitoring bankruptcy the authors propose the following general scheme of DSS (Figure 1) and used knowledge engineering, expert system (ES) technology [4] and data mining (DM) technology [2][3].

The expert system technology underlies two modules of DSS in bankruptcy monitoring [8]:

- module for grouping companies depending on the level of risk of bankruptcy (module1);
- module for the identification of the signs of illegal bankruptcy (module 2).

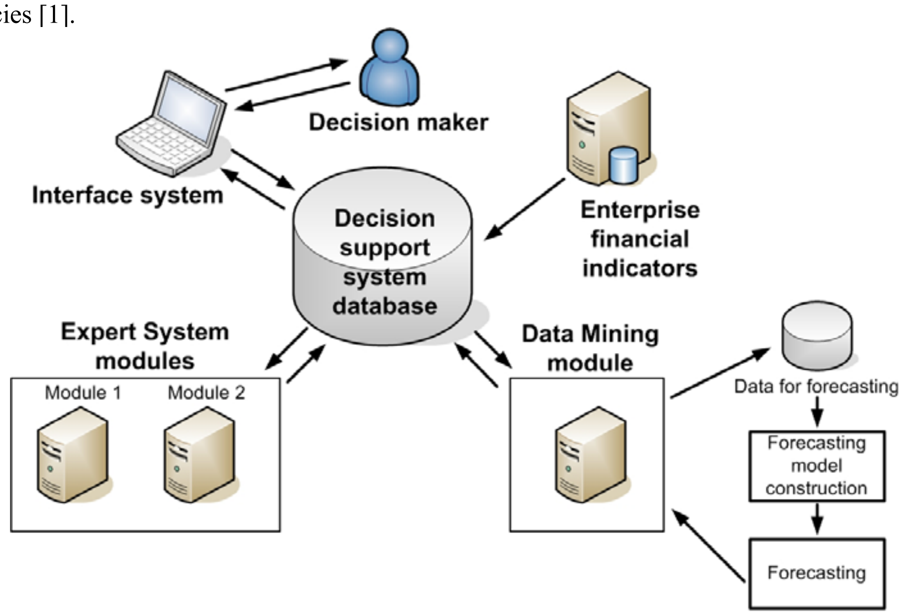

Figure 1. The general scheme of the DSS in monitoring bankruptcies

Another module of the DSS is a data mining module. This module helps to solve problems which include cleaning the data for qualitative forecasting and predicting financial indicators of the enterprise with the use of several prognostic model-building mechanisms, including self-teaching algorithms. The objective of this module is to identify negative trends in changing financial indicators as well as possible signs of fraudulent bankruptcy based on the comparative analysis of the current financial indicators and financial indicators forecasted by the data mining module.

Primary, intermediate and resulting data are stored in the main decision support database organized according to the relational model. To keep the decision support system operating the primary data on the company is imported in the system either automatically or manually. Interaction between the DSS and the user is carried out by means of an interface subsystem.

In the first phase of the DSS the enterprise is classified according to the degree of the threat of bankruptcy by means of module 1 of the expert system (Figure 2).

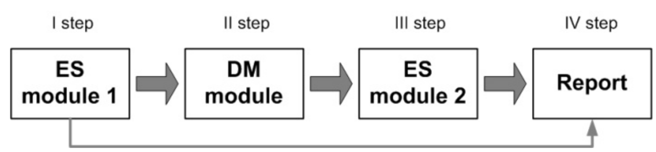

Figure 2. The steps of the DSS modules using

 ISBN: 978-1-61208-227-1

Depending on the results, the enterprise is either checked for signs of fraudulent bankruptcy (I step, module 1 of the expert system), or financial performance is forecasted using the data mining technology (II step, DM module). In the third phase on the basis of the forecasted values the signs of the deliberate bankruptcy are identified (III step, module 2 of expert system). On the IV step a report is made for the decision maker.

III. DATA MINING MODULE

The problem which is solved by the data mining module in DSS in monitoring enterprise bankruptcy is the problem of forecasting financial indicators of the enterprise (company). This problem can be seen as a problem of forecasting the time series, as the data for the prediction of financial indicators are presented in the form of measurement sequences, collated at non-random moments of time. In contrast to the analysis of random sampling the analysis of time series is based on the assumption that successive values are observed in equal periods of time. Like many other kinds of analysis the analysis of time series implies that the data contain a systematic component (generally including several components) and a random noise (error) which makes it hard to detect regular components.

A time series may be presented as decomposition of four constituents [11]:

$$X_t = f(S_t, T_t, C_t, R_t)$$

where S_t – seasonality effect; T_t – trend, or systematic movement; C_t – fluctuations around the trend with more or less regularity (cyclicality); R_t – random (unsystematic) residual component. The action of these constituents may be interdependent. The models in which the time series is presented as a sum of the given components is called additive if multiplicative models are in the form of the product of numbers. The additive model takes the form: $X_t=S_t+T_t+C_t+R_t$, the multiplicative one: $X_t=S_t*T_t*C_t*R_t$. There exist also mixed models. The main task in investigating the time series consists in detecting and determining the quantification of each component (S_t, T_t, C_t, R_t) for forecasting the future values of the series.

The dynamics of lots of financial and economic indicators has a stable fluctuation constituent. In order to obtain accurate predictive estimates it is necessary to represent correctly not only the trend but the seasonal components as well. The use of data mining methods in time series forecasting makes the solution of the given task possible. These methods have a number of benefits:

- possibility to process large volumes of data;
- possibility to discover hidden patterns;
- use of neural networks in forecasting allows obtaining the result of the required accuracy without determining the precise mathematical dependence.

There are a lot of other benefits of data mining such as basic data pre-processing, their storage and transformation, batch processing, importing and exporting of large volumes of data, availability of data pre-processing units as well as ample opportunities for data analysis and forecasting.

The algorithm for forecasting the companies' financial indicators has been developed. It works as follows (Figure 3). Let us assume that as a result of transformation by the "sliding window" method we obtain a sequence of time counts [6]:

$$X_{-n}, \ldots, X_{-2}, X_{-1}, X$$

where X is a current value. Forecasting for X_{+1} is carried out on the basis of the built model. In order to forecast the value of X_{+2} it is necessary to shift the whole of the sequence one count to the left so that the forecast of X_{+1} carried out earlier could be included in the initial values. Then once again the algorithm for computing the predicted value will be started. X_{+2} is calculated with regard for X_{+1} and further in a similar way, according to the defined forecasting horizon. To debug the prediction algorithm it is necessary to define the forecasting horizon as well as the table fields which must be filled in to carry out a forecast (to calculate the output field of the model).

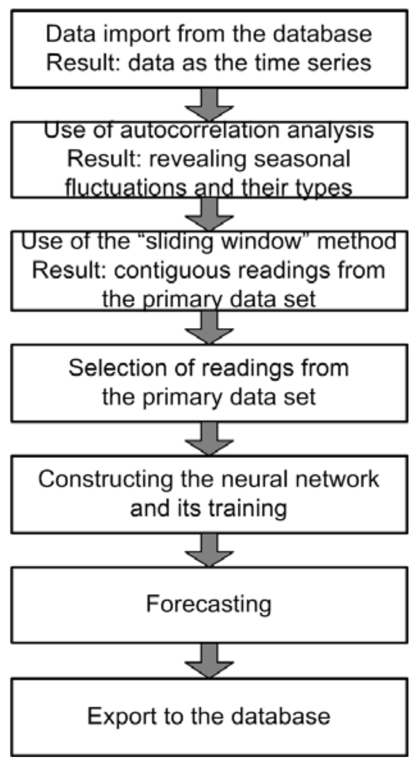

Figure 3. Main stages of the data mining module

Forecasting of enterprise financial indicators in the DSS can be performed by means of a number of DM techniques such as partial and complex data preprocessing, autocorrelation analysis, the method of "sliding window" and neural networks.

In solving the problem of forecasting the time series with the aid of a neural net it is required to input the values of several adjacent counts from the initial set of data into the analyzer. This method of data sampling is called "sliding window" (window – because only a certain area of data is highlighted, sliding – because this window "moves" across the whole data set). The efficiency of implementation considerably increases, if we do not sample the data out of a number of consecutive records, but successively locate the data related to the specific position of the window in one record. The values in one of the writing fields will be related to the current count and in other ones they will be shifted from the current count to the "future" or the "past". Thus, transformation of the sliding window has two parameters: "depth of plunging" – the number of the "past" counts in the window and "forecasting horizon" – the number of "future" counts. It should be mentioned that for the boundary positions of the window (relative to the beginning and the end of the whole sampling) incomplete records will be formed, i.e. records containing empty values for the missing past and future counts. The transformation algorithm allows either excluding such records from the sampling (in that case for several boundary counts there will be no records) or including them (in the latter case records will be made for all the counts available, but some of them will be incomplete).

With the use of the neural network the forecasting problem can be set in the following way: to find the best approach to the function defined by the final set of input values (teaching examples). In our case the neural networks help to solve the problem of recovery of the missing values as well as the forecasting of financial indicators of the enterprise being analyzed.

IV. Software Implementation of the Data Mining Module

The software implementation of the data mining module to forecast the financial indicators of the enterprise is performed by means of the analytical platform [9]. As it was mentioned in Section 3, the data mining module is realized by the following main steps:

1) primary data input (Figure 4);
2) using of "sliding window" (Figure 6);
3) neural network programming – constructing and teaching (Figures 7, 8);
4) forecasting (Figure 9).

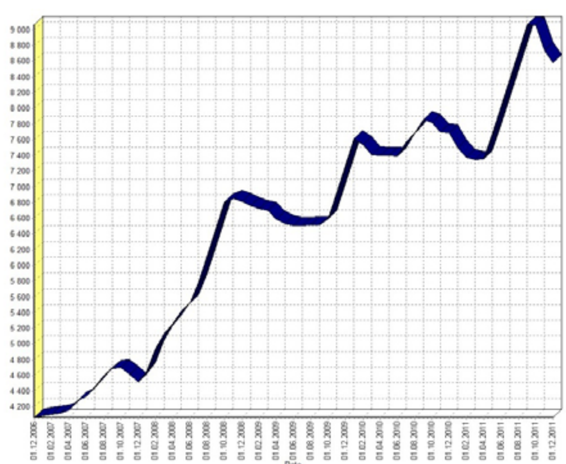

Figure 4. Primary data input (graphic representation)

Let us consider the whole of the data mining process on the example of the enterprise's financial indicator "fixed assets".

After the input of the initial information preprocessing of the raw data is required. Figure 5 shows a possible sequence of steps in preprocessing the raw data before using data mining models.

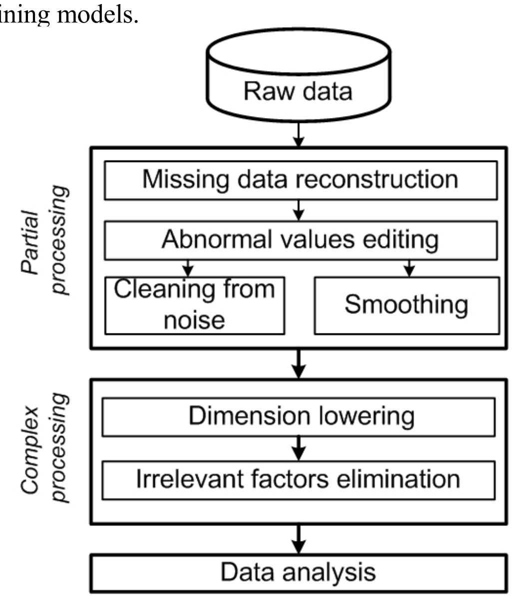

Figure 5. Data preprocessing

Partial and complex processing are distinguished in preprocessing. In partial processing the missing data are restored. Abnormal values are edited, noise is subtracted, and smoothing is carried out. For these purposes correlation analysis, factor analysis, main component method, regression analysis and other methods are used.

In complex processing the lowering of the dimension of the input data and/or elimination of irrelevant factors take place. Robust filtering, spectral and wavelet analysis are used. In practice, partial and complex preprocessing of the raw data can be performed in any sequence with any parameters at each step, any number of times, that is, the preprocessing script can be quite complex [5].

Analysis of the primary ("raw") data is made step-by-step. Cleaning, transformation and forecasting of the data is done individually with each time series of the enterprise's financial indicators. In forecasting the time series by means of the neural networks it is required to input the values of several adjacent counts from the source data set – "sliding window" (Figure 6).

The efficiency of implementation considerably increases, if we do not sample the data every time from a number of consecutive records, but successively locate the data related to the specific position of the window in one record. The experiments show that for "sliding window" the optimal value of the depth of plunging is 5, because 5 inputs is enough for neural network to forecast the fixed assets.

ISBN: 978-1-61208-227-1

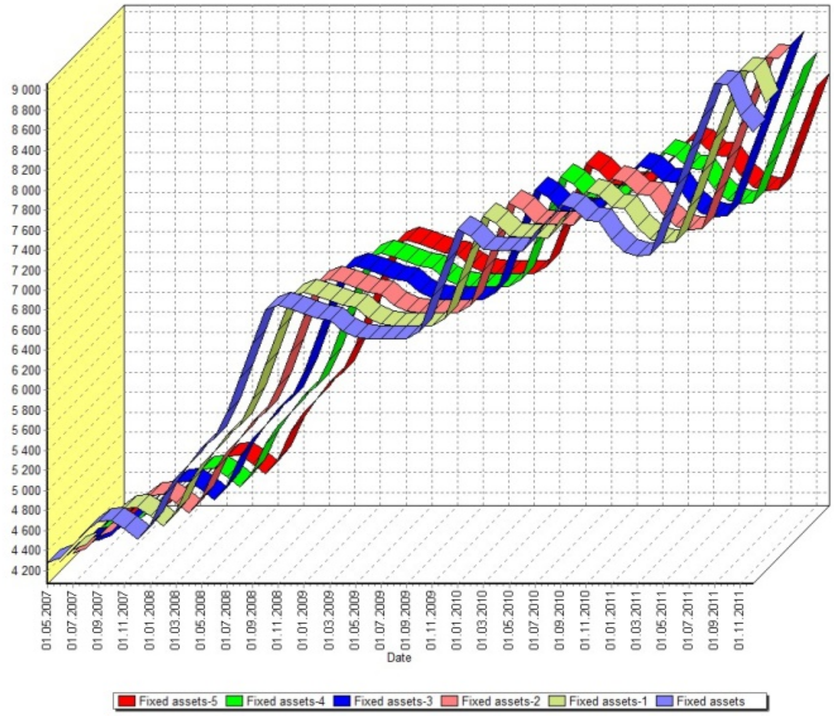

Figure 6. “Sliding window” for fixed assets

The neural network structure for forecasting the enterprise’s indicator “fixed assets” has the form 5-4-3-2-1. The graph of the neural network and dispersing diagram for the neural network quality estimation are shown in Figures 7 and 8.

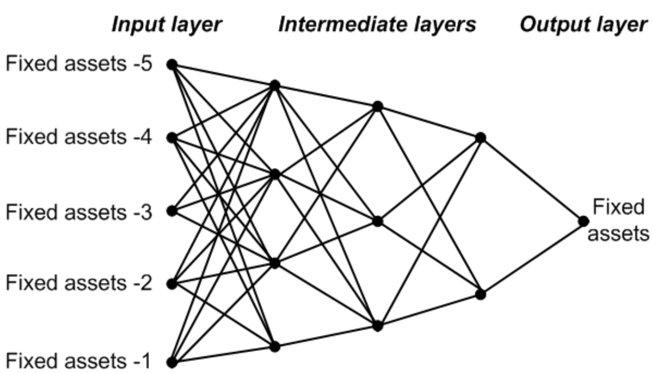

Figure 7. The neural network structure for indicator “fixed assets”

The dispersing diagram allows estimating the quality of the neural network constructing and teaching – the points are closer to the central axis the accuracy of the neural network model is higher.

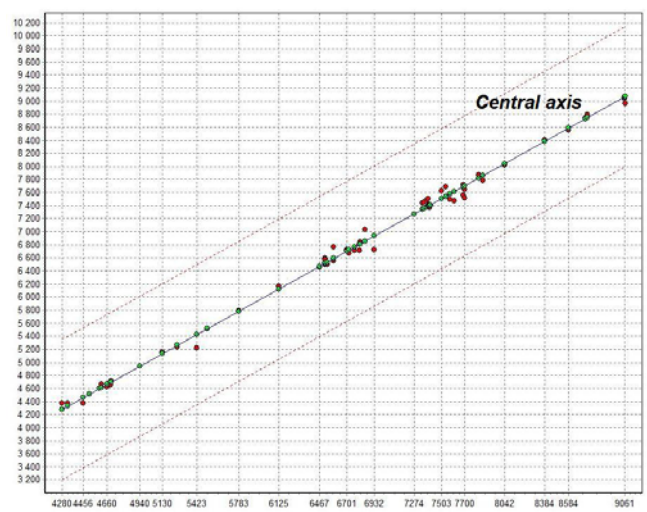

Figure 8. The dispersing diagram

The neural network performance may be evaluated also by the errors diagram (Figure 9), that shows the neural network error for each of the indices and the average error (in percentage).

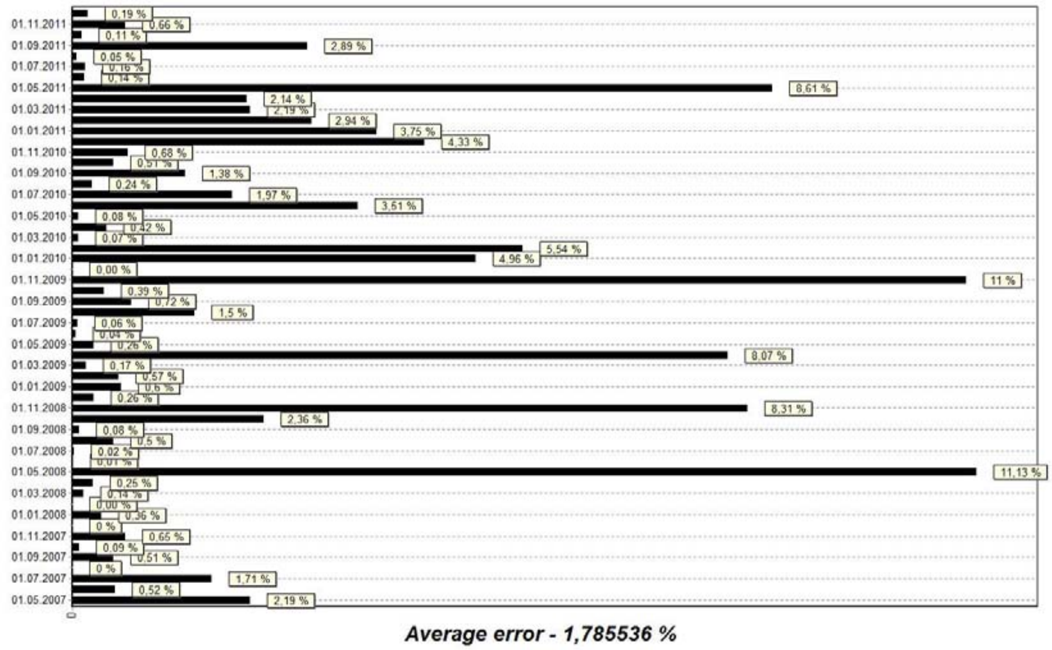

Figure 9. The errors of the neural network performance for indicator "fixed assets"(graphic representation)

The final stage in the program implementation of data mining is forecasting (Figure 10).

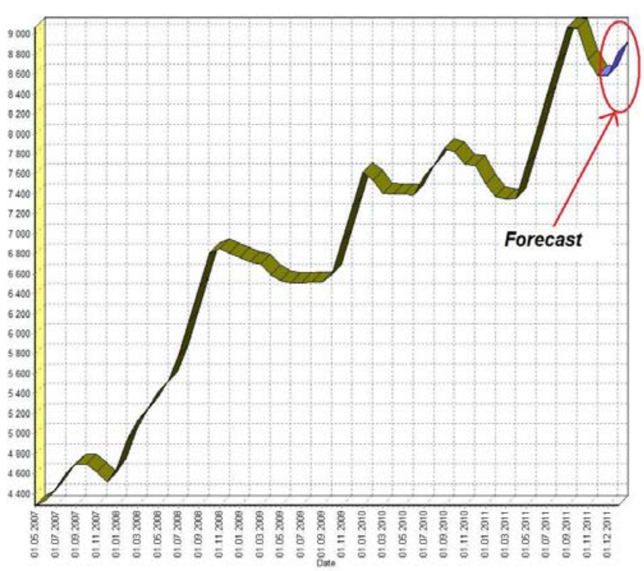

Figure 10. Results of forecasting of the fixed assets (graphic representation)

Each of the financial indicators has its own prediction algorithm that includes the size of the step of the sliding window, neural network structure, the form of the activation function and its value (Table I). These parameters are defined for each enterprise individually.

TABLE I. ALGORITHMS OF DATA MINING APPLICATION TO FORECAST ENTERPRISE'S FINANCIAL INDICATORS

Financial indicators of the enterprise	**Depth of plunging of "sliding window"**	**Form of the activation function**	**Value of the activation function**	**Neural network structure**
Fictitious assets, i.e. patents, licenses, trade marks	5	Sigma form	0.80	5-3-1
Fixed assets	***5***	***Sigma form***	***1.30***	***5-4-3-2-1***
Long-term financial investments	3	Sigma form	0.95	3-2-1
Total of non-working assets	5	Hyper tangent form	1.05	5-3-2-1
…	…	…	…	…
Reserves of forth-coming expenses and payments	5	Arctangent form	1.25	5-2-1
Total of short-term liabilities	4	Sigma form	1.10	4-2-1
Liabilities balance	5	Hyper tangent form	0.85	5-2-1

V. DATA MINING IMPLEMENTATION EFFICIENCY ANALYSIS

The analysis of the effectiveness of the data mining module is based on the comparative analysis of the financial indicators for the same period of time, obtained directly from the enterprise and forecasted through data mining.

This approach has been used in state monitoring of a number of industrial and agro-industrial enterprises of Republic Bashkortostan (Russia).

The fragment of the analysis of the effectiveness of data mining with deviation of the forecasted values of the financial indicators from the actual data is presented in Table II.

TABLE II. DEVIATIONS OF FORECASTED VALUES FROM ACTUAL DATA, IN PERCENTAGE

Financial indicators of the enterprises	Enterprise 1	Enterprise 2	Enterprise 3	Enterprise 4
Fixed assets	3.57	4.28	4.86	5.05
Long-term investments	4.62	1.45	5.61	6.63
Total non-current assets	4.89	4.22	3.77	3.25
Total current assets	6.32	4.65	5.90	7.56
All long-term liabilities	5.41	8.42	4.27	7.34
Loans and credits	5.75	2.05	7.82	8.74
Creditor indebtedness	2.56	3.41	7.68	1.35
…	…	…	…	…
Tax liabilities	7.32	1.91	6.82	5.89
Result of short-term liabilities	6.83	5.88	6.24	8.67
Balance of liabilities	2.42	4.47	8.45	3.57
Revenue for the period under review	5.87	3.55	5.62	1.52

Analysis of the effectiveness of data mining for values forecasting showed that the deviations of the forecasted values from the real data are in the range from 1.35% to 8.74%. The average deviation is about 6.5 %, which is quite a good result for forecasting.

CONCLUSION AND FUTURE WORK

The data mining application in anti-crisis management for bankruptcy monitoring is investigated. The decision support system for bankruptcy monitoring including the data mining module is developed. The decision maker using the DSS may be the top manager or supervisory authority. It is possible for users of the system to monitor the major trends in the economic processes of the enterprise. With the help of data mining means, neural networks in particular, the enterprise financial indicators can be forecasted for the definite period of time (for example for 3 months). The aim of the neural network at this stage is to catch the regularities of the financial indicators changes and detect them. Then with the help of the expert system the enterprise is classified on the basis of the forecasted indicators according to the degree of the bankruptcy threat. In other words the condition of the enterprise is defined not for present moment but for the definite time period (for example for 3 months). It gives an opportunity to take measures preventing the enterprise from fraudulent bankruptcy. The description of financial indicators forecasting based on data mining technology is given. For each of the forecasted financial indicators the separate forecasting algorithm has been developed. The efficiency analysis reveals good results of data mining implementation for forecasting financial indicators in the decision support system for bankruptcy monitoring. These results confirm the possibility of using data mining technology in the developed decision support system for bankruptcy monitoring. Future work will be connected with the increasing accuracy of financial performance forecasts.

ACKNOWLEDGMENT

This research has been supported by grants № 11-07-00687-a, № 12-07-00377-a of the Russian Foundation for Basic Research and grant “The development of tools to support decision making in different types of management activities in industry with semi-structured data based on the technology of distributed artificial intelligence” of the Ministry of Education of the Russian Federation.

REFERENCES

[1] A. Belyaev and E. Korotkov, Anti-Crisis Management. UNITI-DANA, Moscow, 2011.

[2] A. Syväjärvi and J. Stenvall, Data Mining in Public and Private Sectors: Organizational and Government Application. Information Science Reference, Hershey, New York, 2010.

[3] A.V. Senthil Kumar, Knowledge Discovery Practices and Emerging Application of Data Mining: Trends and New Domains. Information Science Reference, Hershey, New York, 2011.

[4] Peter Jackson. Introduction to Expert Systems. Williams, Moscow, 2001.

[5] A. Barsegyan, M. Kupriyanov, V. Stepanenko, and I. Kholod, Technologies of Data Analysis: Data Mining, Visual Mining, Text Mining, OLAP. Sec. ed. BHV-Petersburg, Saint-Petersburg, 2007.

[6] V. Duk and A. Samoilenko, Data Mining. Peter, Saint-Petersburg, 2001.

[7] N. Yusupova and G. Shakhmametova, Intelligent Information Technologies in the Decision Support System for Enterprises Bankruptcy Monitoring. UNC RAN, Ufa, 2010.

[8] G. Shakhmametova, D. Amineva, and V. Dolzhenko, “Expert System for Decision Support in Anti-Crisis Monitoring,” Proc. of the 13-th International Workshop on Computer Science and Information Technologies, Germany, Garmisch-Partenkirchen, 2011. Vol. 1. pp. 151-155.

[9] http://www.basegroup.ru 20.06.2012.

[10] http://EconPapers.repec.org/RePEc:ovi:oviste:v:xi:y:2011:i:9:p:233-236, 20.08.2012.

[11] V. Afanasyev, The Time Series Analysis and Forecasting, Finance and Statistics, Moscow, 2001.