THE ROLE OF FINANCIAL MANAGEMENT IN USING DATA MINING IN AGRICULTURAL COMPANIES

Tanja Arapović Ilić[1], Snežana Krstić[2], Miloš Dašić [3], Bojan Brajkovic [4], Radovan Damnjanović [5], Dragana Trnavac [6], Aleksandar Savić [7]

**Corresponding author E-mail: snezana.krstic1966@gmail.com*

ARTICLE INFO

Review Article

Received: *07 December 2022*

Accepted: *15 February 2023*

doi:10.59267/ekoPolj2301267A

UDC 005:336]:[004.8:631.151

Keywords:

agroeconomy, financial management and business, analytical methods, Data mining, group method of data acceptance - Group Method of Data Handling -GMDH

***JEL**: M41, M49*

ABSTRACT

Analytical methods are an indispensable method of auditing. Auditors typically use classical methods such as horizontal, vertical, regression analysis, such as the Z-score. Very few data mining methods are used at all, which are significantly more accurate in their results than the ones mentioned. The subject of this paper is the application of one of the most efficient methods of data so-called. Group Method of Data Handling –GMDH in agro entities.

1 Tanja Arapović Ilić, PhD, Ministry of Defence Republic of Serbia, Birčaninova 5, 11000 Belgrade, Serbia, +38163222158, E-mail: tanja.arapovic@mod.gov.rs, ORCID ID (https://orcid.org/0000-0002-4726-0112)

2 Snežana Krstić, Associate Professor, Military Academy, University of Defence, Veljka Lukića Kurjaka 33, 11000 Belgrade, Serbia, E-mail: snezana.krstic1966@gmail.com, ORCID ID(https://orcid.org/0000-0001-6666-6401)

3 Miloš Dašić, Academy of Vocational Studies South Serbia, Blace Business School Department, st. Kralja Peta I number 1, 18420 Blace, Serbia, Phone: +381 60 41 00 109, E-mail: dasicmil@yahoo.com, ORCID ID (https://orcid.org/0000-0002-6099-8381)

4 Bojan Brajkovic, The University of Business Studies Banja Luka, PhD student, 78000 Banja Luka,Bosnia and Herzegovina E-mail: brajkoboj@gmail.com, ORCID ID (https://orcid.org/0000-0002-8043-9237)

5 Radovan Damnjanović, Assistent Professor, Military Academy, University of Defence, Veljka Lukića Kurjaka 33, 11000 Belgrade, Serbia, e-mail: radovandam78@gmail.com, ORCID ID (https://orcid.org/0000-0002-4574-5546)

6 Dragana Trnavac, Associate Professor, "MB" University, Faculty of Business and Law, Teodora Drajzera 27, 11000 Belgrade, Serbia, E-mail: draganatrbnavac@yahoo.com, ORCID ID (https://orcid.org/0000-0003-1571-4500)

7 Aleksandar Savić, assistant, Military Technical Institute, Ratka Resanovića 2, 11000 Belgrade, Serbia, E-mail: aleksandar22071993@gmail.com, ORCID ID (https://orcid.org/0000-0002-9640-1583)

Introduction

Civilization is nowadays become dependent on large-scale systems of machines, environment and men. Everyday work, performed by auditors and financial experts, in addition to routine-repetitive work, requires knowledge that can be extracted from existing databases, not only data, but also knowledge represented in financial reports. Revision process nowadays, besides classical analytical methods (Z-score, regression analysis, horizontal analysis, vertical analysis includes modern methods of Data Mining (finding implicit - finding hidden knowledge) (Witten & Frank E. (2005)). Based on the cybernetic principle of self-organization, by learning the unknown relations between the outputs and inputs of a given system (in our case financial statements of agricultural entities) based on the evolutionary principle, which understands the initial very simple model of organization to optimally complex model (Savić & Obradović, 2020).

Various studies on agricultural system reveal that fulfilling agricultural production forecasts, particularly in large irrigation systems remains a difficult problem. Difficulties arise which seemingly cannot be overcome by conventional modeling techniques.

Accounting planning deals with the design, systematization, processing and presentation of data related to the future business of the agro company (Ilić & Tasić, 2021). Thus, this constitutive element of management accounting refers to the economic transactions of the agro company that are yet to take place. The end product of accounting planning is accounting estimates that include data on projected balances or changes in the future. Accounting analysis is the judgment and explanation of the state and success of agro business, determining deviations from it, the causes of these deviations and their consequences. The formation of proposals for improving business processes and the situation is also an integral part of accounting analysis (Mihajlović et al., 2020). The main subject of analysis is data and information provided by costing and analysis. The data provided by accounting planning and analysis are necessary inputs for conducting business control. Accountability accounting is a system that connects the plans and actions of each responsibility center in the company (Tekić et al., 2021). The responsibility center is a part of the company whose manager is in charge of a specific set of company activities.

Methodology

„There is no consensus in the literature and practice on the applicability and importance of traditional accounting planning and budgeting in terms of successful management of agro business activities" (Dukić-Mijatović et al., 2021). Conceptual definitions of these two ways of planning reflect differences in understanding the importance of one or another way of planning for the successful operation of the agro company, but also for its survival. The budget, viewed as a result of budgeting activities, is an instrument of effective short-term planning and control. At the same time, the purpose of budgeting, which consists in refining the strategic plan (Spathis et al., 2003), coordination, delegation of responsibilities and creating a basis for assessing performance indicates

the understanding of business budget as an instrument for achieving strategic goals (Vićentijević, 2021).

Putting the achievement of strategic goals in the forefront as budgeting as an instrument represents a broader view of the essence of budgeting. However, the broader understanding of budgeting highlights that planning is becoming the key to good management, that is, that without disciplined professional planning, the agro company goes into disrepair. Profit planning is a budgeting process (Ristić et al. 2021). This approach practically gives maximum importance to budgeting as a process to which the fate of the agro company is tied. Budgeting in a broader sense implies accounting planning and control, whereby budgeting is the responsibility of accounting responsibility (Avakumović et al., 2021). Control in this view implies a comparison with the planned values, which is the essence of control, but budgeting is still more than accounting planning and control, since the essence of budgeting is expressed by flexibility in relation to short-term changes. Machine learning is "a set of processes, which includes: collecting new declarative knowledge, developing and improving motor and cognitive skills through practice, structure of existing knowledge and discovering new facts and theories through observation and active experimentation" (Bogavac et al., 2021).

Learning can be viewed through two basic forms (Green & Choi, 1997):

- knowledge acquisition, which is the learning of new, symbolic information so that it can be effectively applied (so one learns theoretical knowledge, eg physics);
- training, which involves improving some acquired knowledge, mental or motor coordination, through practical repetition and correction of deviations from the desired behavior (so a person learns different skills, with the first phase of learning is the collection of knowledge);
- It is considered that human learning is a mixture of both forms, with mental activities emphasizing the first form, and motor activities to a greater extent the second form of learning.

Machine learning systems are most often divided according to the chosen learning strategy, the way of presenting knowledge and the area of application. Inductive (machine, automated), which is the subject of application in the analysis of financial statements, ie. their learning can be seen as a process in which the system improves its performance on a given task without additional programming, using two methods. We must mention that there are other methods of learning strategy, for example: learning by rote learning, learning by being told, learning by analogy. Learning by examples, which requires inductive reasoning. By analyzing and generalizing solved examples and counter examples of a class of phenomena (financial statements as a concept), we come to a rule, theory or description of the term, which explains all examples and no counter examples. Such learning methods can be further classified according to the choice of examples, source and way of using examples. The methods used by Data mining, which

are not the subject of our consideration, can be: production rules, decision lists and decision trees, which are examples of an understandable way of presenting empirical knowledge. Our focus is on applying a self-organized Data Mining model (SOHK).

Analytical model of self-organized discovery of hidden knowledge (SOHK)

Budgeting has a crucial role to play in agro business. Without a budget (plan), it is difficult to expect managers and their employees to achieve business growth and cost management goals. In developed market economies, financial managers spend 20% -30% of their time on budget-related jobs (Rakić et al., 2021).

In our conditions, the budgeting process is largely excluded from agro business practice. Which is definitely wrong. Agro companies should develop the skills of employees in this field because it will help them not only to discover development opportunities and control their business, and develop an adequate system of rewarding employees, but also to avoid unnecessary costs of paying external consultants when making business plans. (Vukša & Pantić, 2020).

Accountability accounting is a system that connects the plans and actions of each responsibility center in the agro company. The responsibility center is a part of the company whose manager is in charge of a specific set of company activities. Accounting for responsibilities (and thus business control) requires data on the amount and structure of costs of cost centers, ie. control units, as well as data on the amount and structure of the cost price of their services. This data is needed to determine differences in the amount and structure of costs between two or more successive periods, or differences between actual and planned costs, etc., as well as to take appropriate measures to eliminate the causes that led to increased costs. In some cases, measures should be taken against unjustified reduction of costs, if it harms the quality of the product (Mihajlović et al., 2018). For almost every expense, it can be determined where it was incurred. For the purposes of control, responsibility for the amount of incurred costs is essential. When determining the responsibility for the incurred costs and the amount of income, it should be borne in mind that the influence of the competent manager is partially limited, ie. the possibility of their control must be taken into account. Namely, the amount of costs and revenues is partly influenced by external factors, such as purchase or sale prices.

The SOHK model is a model that has significant advantages over classical neural networks and genetic algorithm, because it is based on the principle of evolutionary, mutational and selective approach to generate structure network systematically enabling automated structure synthesis and model validation until the optimal complex model is established (Michalski et al., 1983; Stice, 1991). The self-organization model performs data reduction, reprocessing and validation of model results that are corrected during the self-organization process, which is called self-organization data mining. The SODM model is presented with a group of methods, which is administratively called the Group Method of Data Handling - GMDH).

Group Method of Data Handling

GMDH was developed in the Institute of Cybernetics in Kyiv by Prof. A.G. Ivakhnenko in 1967, with improvements in 1970 and 1980. Group method of data handling represents a system of inductive algorithms for computer-based mathematical modeling in various multicycled systems such as: neural networks, noise immunity, clusterization, economic systems etc. (Ivakhnenko, 1995). Nowadays, there are a wide range of software that uses GMDG, such as: FAKE GAME Project, Gevom, GMDH Shell, KnowledgeMiner, PNN Discovery client, Sciengy RPF!, wGMDH, R Package, R Package for regression tasks, Python library of MIA algorithm.

"Self-organization modeling of inductive algorithms in the basis are using seven fundamental steps:

1. Data sample of *N* observations corresponding to the system under study is required; Split them into training set *A* and testing set *B (N = NA + NB)*.
2. Build up a "reference function" as a general relationship between dependent (output) and independent (input) variables.
3. Identify problem objectives like regularization or prediction. Choose the objective rule from the standard selection criteria list which is developed as "external complements."
4. Sort out various partial functions based on the "reference function."
5. Estimate the weights of all partial functions by a parameter estimation technique using the training data set A.
6. Compute quality measures of these functions according to the objective rule chosen using the testing data set B.
7. Choose the best measured function as an optimal model. If you are not satisfied, choose F number of partial functions which are better than all (this is called "freedom of choice") and do further analysis." (Madala, 1994)

The main advantages of GMDH methods are, in short, solving the problem of applying neural network learning techniques whose algorithms are slower and less efficient than highly optimized algorithms used in statistical software, as well as overcoming the problem of accumulation (Cherkassky & Mulier, 2007). The GMDH algorithm selects a model of optimal complexity by applying an inductive approach.

GMDH algorithm can be presented in Multi - layer artificial neural networks ("There are three main inductive learning networks: multilayer, combinatorial and harmonic. The network structures differ as per the interconnections among the units and their hierarchical levels.") where the structure consider the number of layers and neurons in each layer. Each sumulated unit k receives input variables $(x_i, x_j) \subset x$, $i \neq j$, and generates a function *f()* which is partial form of the reverence function.

$$f(x_i, x_j) = v_0^{(k)} + v_1^{(k)} x_i + v_2^{(k)} x_j + v_3^{(k)} x_i x_j + v_4^{(k)} x_i^2 + v_5^{(k)} x_j^2$$

Where $v^{(k)}$ are the connecting weights. If we denote *0* as the desired values and *y* as the estimated values of the outputs for the function being considered, the output errors would be given by

$$e_p = y_p - O_p; p \in N_A$$

The total squared error for that input vector is

$$E = \sum_{p \in N_A} e_p^2$$

This corresponds to the minimization of the average error *E* in estimating the weights $v^{(k)}$; this is the least squares technique. The weights are computed using a specific traninig sample N_A which is a part of the whole data points *N* specified for this purpose.

Each layer includes a group of units that are interconnected to the units in the next layer. The proces continues layer after layer. Each layer contrains a group of units that are interconnected to the units in the next layer. The weights at each unit are estimated by minimizing the error *E*. The measure of an abjective function is used as the threshold value to make the unit "on" or "off" in comparison with the testing data N_B which is another part of *N* and, at the same time, it is considered to obtain the optimum output response. On this level of modeling, a hardly avoidable error can be reduced to a minimum and finding the solution to optimize the problem, by using the assumed objective function to successive iterations.

"A GMDH network is made up of a number (m) of single neurons (the structure of a single neuron is shown in figure which process the input signal – vector x – and turn it into an output signal, y. The signal is processed when at least two input signals are stimuli, according to the following relation:" (Mrowczynska M., 2019)

$$y = f(x) = f(x_1, x_2, \dots, x_m)$$

where *f* is the transfer function.

"The transfer function must not be too complex, as this would extend the time required for training and would prevent an accurate assessment of the training error. Therefore, although the GMDH algorithm permits the application of various forms of the transfer function, the function is most considered a discrete form of the Volterra functional series which is also called Kolmogorov - Gabor polynomial, defined as: (Mrowczynska M., 2019)"

$$y = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j + \cdots$$

"where *i, j, a, 0* are polynomial parameters. Assuming that the polynomial has a degree of *n=2*, the transfer function is defined as: (Mrowczynska M., 2019)"

$$y = a_o + a_1 x_1 + a_2 x_2 + a_{11} x_1^2 + a_{22} x_2^2 + a_{12} x_1 x_2$$

This includes the application of external information that was not used to estimate the coefficients of the model.

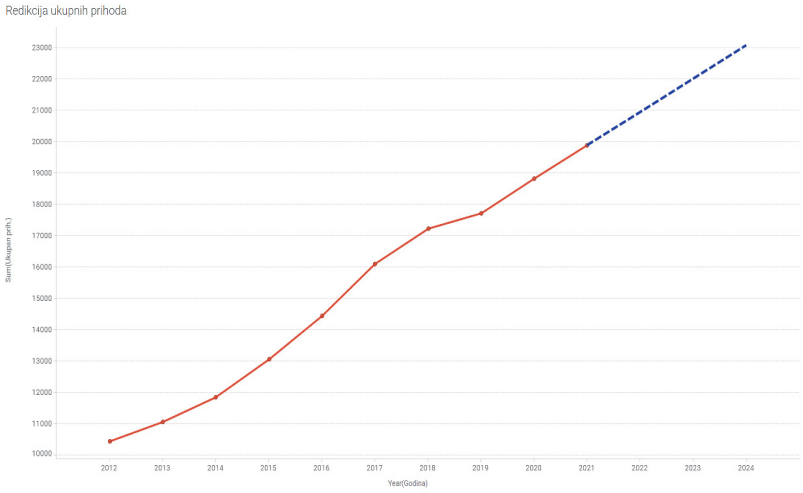
To illustrate the application of this method, we will take the following company data.

Table 1. Financial indicators

Year	Total fixed assets	Cap. investment	Current assets	Capital	Commission	Obligations	Total assets	Total revenue	Operating expenses	Net operating profit	Operating profit
2012	6000	68	2134	1424	2020	3239	8252	10437	11333	235	134
2013	6417	61	2292	2265	2241	2625	8815	11065	11990	313	207
2014	7705	64	2581	2276	2550	3614	10411	11845	12754	303	241
2015	8824	78	3331	3262	2700	4045	12281	13055	14093	314	260
2016	10831	78	3574	3144	2992	6237	14540	14447	15229	305	142
2017	11781	83	4025	2687	3433	7926	15954	16101	17110	11	-301
2018	12459	79	4270	2281	4193	8907	16870	17239	18135	67	-310
2019	12033	91	5183	2137	4429	9210	17410	17731	18597	325	-8
2020	12857	103	5105	3650	6371	7225	18138	18836	19702	596	399
2021	12455	116	5788	4943	6843	6618	18419	19900	20598	837	756

Source: Author's calculation

Figure 1. Reduction of total revenues



Source: Author's calculation

Table 2. Statistics

Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
Really	10437	11065	11845	13055	14447	16101	17239	17731	18836	19900			
Prediction	10069	11179	12289	13400	14569	15870	16925	17944	18928	19876	20789	21960	23392
Accumulation													
Final prediction	10069	11179	12289	13400	14569	15870	16925	17944	18928	19876	20789	21960	23392
Lower	9545	10655	11765	12876	14045	15346	16401	17420	18404	19352	20265	21437	22868
Higher	10593	11703	12813	13924	15093	16394	17449	18468	19452	20400	21313	22484	23916

Source: Author's calculation

The results indicate a high level of forecasting accuracy, which qualifies GBDH as an acceptable method of forecasting the movement of the balance of financial statements in every entity, includin agro company.

Conclusion

The changes caused by the information revolution, the development of technology pose new challenges to traditional accounting planning, and make the budgeting process more complex, bearing in mind that these are very significant changes that are happening in the agro business environment. At the same time, it should be borne in mind that the new business conditions are characteristic, among other things, of the increase in the mass of general costs that are not caused by increased production. An approach to budgeting that does not take this fact into account shows a weakness that can have negative effects on the company's operations and the achievement of strategic goals. In this sense, traditional accounting planning works well for activities that show a clear link between inputs and outputs, while otherwise, traditional budgeting serves only to approve a certain level of spending for each cost item.

The article is an attempt to show the success and applicability of Neural Networks, as well as the self-organized Data Mining System. For short-term forecasts of financial indicators. The great advantage of this method is in the fact that it supports nonlinear data forms, such as administrative and financial data. It is to be expected that Data Mining, especially the GBDH method, will become a standard technique for auditors and financial analysts. For directions of further research, the results of the research are expected to be improved by combining models with additional information of macroeconomic categories of the agro economy and building multi-input models for the extended database.

Conflict of interests

The authors declare no conflict of interest.

References

1. Alvin K. Mulashani, Chuanbo Shen, Baraka M. Nkurlu, Christopher N. Mkono, Martin Kawamala, Enhanced group method of data handling (GMDH) for permeability prediction based on the modified Levenberg Marquardt technique from well log data, Energy, Volume 239, Part A, 2022, https://doi.org/10.1016/j.energy.2021.121915
2. Amiri, M., Soleimani, S. ML-based group method of data handling: an improvement on the conventional GMDH. Complex Intell. Syst. 7, 2949–2960 (2021). https://doi.org/10.1007/s40747-021-00480-0
3. Avakumović, J., Obradović, L., & Božić, G. (2021). Management and organization of teamwork in the function of sustainable development. *Održivi razvoj*, *3*(2), 69-80. [*in Serbian*: Avakumović, J., Tešić, R., & Karić, D. (2021). Menadžment tranzicionim procesima u funkciji održivog razvoja].
4. Bogavac, M., Todorović, V., Karić, D., & Rosić, M. (2021). Proof in the accounting control procedure. *Oditor*, *7*(2), 25-35. https://doi.org/10.5937/Oditor2102025B
5. Cherkassky V. & Mulier F. M. (2007). *Learning from Data: Concepts, Theory, and Methods*, 2nd edition, John Wiley - IEEE Press
6. Dukić-Mijatović, M., Uzelac, O., & Stoiljković, A. (2021). Agricultural sustainability and social responsibility. *Ekonomika poljoprivrede*, *68*(4), 1109-1119.
7. Green, B. P., & Choi, J. H. (1997). Assessing the risk of management fraud through neural-network technology. *Auditing: A Journal of Practice and Theory*, 16(1), 8–28.
8. Ilić, B., & Tasić, S. (2021). Quantitative analysis of the role of production in value creation. *Održivi razvoj*, *3*(1), 17-33. [*in Serbian*: Ilić, B., & Tasić, S. (2021). Kvantitativna analiza uloge proizvodnje u stvaranju vrednosti].
9. Ivakhnenko, O.G.; Ivakhnenko, G.A. (1995). *"The Review of Problems Solvable by Algorithms of the Group Method of Data Handling (GMDH)"*. Pattern Recognition and Image Analysis. 5 (4): 527–535.
10. Ivakhnenko A.G., *"Polynomial Theory of Complex Systems,"* in IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-1, no. 4, pp. 364-378, Oct. 1971, https://doi:10.1109/TSMC.1971.4308320
11. Madala, H.R.; Ivakhnenko, O.G. (1994). *Inductive Learning Algorithms for Complex Systems Modeling.* Boca Raton: CRC Press. ISBN 978-0849344381
12. Michalski. R., Carbonell.J. &.Mitchell T. (Eds.) (1983). *Machine learning: An artifcial intelligence approach (Vol. I).* Morgan Kaufmann, San Francisco.

13. Mihajlović, M., Nikolić, S., & Tasić, S. (2020). Sustainability of the economic model of contemporary economy. *Održivi razvoj*, *2*(2), 7-13. [*in Serbian*: Mihajlović, M., Nikolić, S., & Tasić, S. (2020). Održivost ekonomskog modela savremene privrede].
14. Mihajlović, M., Stanojević, P. & Tešić, A. (2018). Knowledge management as a factor for increasing organization efficiency. *Akcionarstvo*, 24(1), 5-14 [*in Serbian*: Михајловић, М., Станојевић, П., & Тешић, А. (2018), Менаџмент знања као фактор повећања ефикасности организације].
15. Mrowczynska M., Sztubecki J., *Group Method of Data Handling as a Tool to Determine Vertical Displacements*, IOP Conf. Series: Materials Science and Engineering 471 (2019) 052077, IOP Publishing, https://doi:10.1088/1757-899X/471/5/052077
16. Parnicki, P. P., Živković-Petrović, D., & Tucaković, J. (2021). Financial malpractice as a destabilization factor of real financial reporting. *Anali Ekonomskog fakulteta u Subotici*, *57*(45), 149-160. https://doi.org/10.5937/AnEkSub2145149P
17. Persons, O. (1995). Using financial statement data to identify factors associated with fraudulent financial reporting. *Journal of Applied Business Research*, 11(3), 38–46.
18. Rakić, S., Adamović, V., & Špiler, M. (2021). Challenges and directions of developing business communication on the global market. *Oditor*, *7*(3), 67-103.
19. Ristić, K., Miljković, Lj., & Milunović, M. (2021). Investment in the banking sector as a basis for money laundering. *Akcionarstvo*, 27(1), 55-70
20. S.J. Farlow. *Self-Organizing Methods in Modelling: GMDH Type Algorithms.* New-York, Bazel: Marcel Decker Inc., 1984, 350 p.
21. Savić, B., & Obradović, N. (2020). Challenges in financial reporting about biological assets. *Oditor*, *6*(1), 7-21.
22. Spathis, C., Doumpos, M., & Zopounidis, C. (2003). Using client performance measures to identify pre-engagement factors associated with qualified audit reports in Greece. *The International Journal of Accounting*, 38(3), 267–284.
23. Stice, J. (1991). Using financial and market information to identify pre- engagement market factors associated with lawsuits against auditors. *The Accounting Review,* 66(3), 516–533.
24. Tekić, D., Mutavdžić, B., Milić, D., Novković, N., Zekić, V., & Novaković, T. (2021). Credit risk assessment of agricultural enterprises in the Republic of Serbia: Logistic regression vs discriminant analysis. *Ekonomika poljoprivrede*, *68*(4), 881-894.
25. Vićentijević, K. (2021). Challenges of accounting in the modern digital environment. *Revizor*, *24*(93), 27-36. https://doi.org/10.5937/rev2193027V
26. Vukša. S. & Pantić, N. (2020). Profit and profit rate. *Akcionarstvo*, 26(1), 7-25 [*in Serbian*: Вукша, С., & Пантић, Н. (2020), Профит и профитна стопа].
27. Witten I. H. & Frank E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*, Morgan Kaufmann Publishers.