

ABSTRACT AND REFERENCES

MATHEMATICS AND CYBERNETICS – APPLIED ASPECTS

DOI: 10.15587/1729-4061.2023.280355
DEVELOPMENT OF A METHOD FOR THE SEARCH OF SOLUTIONS IN THE SPHERE OF NATIONAL SECURITY USING BIO-INSPIRED ALGORITHMS (p. 6–13)

Volodymyr Koval

General Staff of the Armed Forces of Ukraine, Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0002-6209-6779>

Andrii Shyshatskyi

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0001-6731-6390>

Ruslan Ranssevych

State Enterprise «Central Research Institute of Navigation and Management», Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0002-6076-5633>

Viktoriya Gura

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0002-4870-4037>

Oleksii Nalapko

Central Scientifically-Research Institute of Armaments and Military Equipment of the Armed Forces of Ukraine, Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0002-3515-2026>

Larysa Shypilova

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0002-5682-3570>

Nadiia Protas

Poltava State Agrarian University, Poltava, Ukraine
ORCID: <https://orcid.org/0000-0003-0943-0587>

Oleksandr Volkov

Scientific-Research Institute of Military Intelligence, Kyiv, Ukraine
ORCID: <https://orcid.org/0009-0003-3777-6195>

Oleksandr Stanovskyi

Odessa Polytechnic National University, Odessa, Ukraine
ORCID: <https://orcid.org/0000-0002-0360-1173>

Olena Chaikovska

The National Defence University of Ukraine named after Ivan Cherniakhovskiy, Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0002-0656-9105>

The object of research are decision support systems. The subject of research is the decision-making process in management problems using bio-inspired algorithms.

A method for the search of solutions in the field of national security using bio-inspired algorithms is proposed. The proposed method is based on a combination of an artificial bat algorithm and evolving artificial neural networks. The method has the following sequence of actions:

- input of initial data;
- processing of initial data taking into account the degree of uncertainty;

- numbering of bat agents (BA);
- placement of bat agents taking into account the degree of uncertainty about the state of the analysis object in the search space;
- setting the initial BA speed and the echolocation frequency of each BA;
- starting a local search;
- launching a global search;
- training knowledge bases of bat agents.

The originality of the proposed method consists in the arrangement of bat agents taking into account the uncertainty of initial data, improved global and local search procedures taking into account the noise degree of data about the state of the analysis object.

Another feature of the proposed method is the use of an improved procedure for training bat agents. The training procedure consists in learning the synaptic weights of an artificial neural network, the type and parameters of the membership function, the architecture of individual elements and the architecture of the artificial neural network as a whole. The method makes it possible to increase the efficiency of data processing at the level of 13–21 % due to the use of additional improved procedures. The proposed method should be used to solve the problems of evaluating complex and dynamic processes in the interests of solving national security problems.

Keywords: efficiency of decision-making, decision support systems, complex processes, national security.

References

1. Bashkyrov, O. M., Kostyna, O. M., Shyshatskyi, A. V. (2015). Rozvytok intehrovanykh system zviazku ta peredachi danykh dlia potreb Zbroinykh Syl. Ozbroiennia ta viyskova tekhnika, 1, 35–39. Available at: http://nbuv.gov.ua/UJRN/ovt_2015_1_7
2. Dudnyk, V., Sinenko, Y., Matsyk, M., Demchenko, Y., Zhyvotovskiy, R., Repilo, I. et al. (2020). Development of a method for training artificial neural networks for intelligent decision support systems. Eastern-European Journal of Enterprise Technologies, 3 (2 (105)), 37–47. doi: <https://doi.org/10.15587/1729-4061.2020.203301>
3. Sova, O., Shyshatskyi, A., Salnikova, O., Zhuk, O., Trotsko, O., Hrokholskyi, Y. (2021). Development of a method for assessment and forecasting of the radio electronic environment. EUREKA: Physics and Engineering, 4, 30–40. doi: <https://doi.org/10.21303/2461-4262.2021.001940>
4. Pievtsov, H., Turinskyi, O., Zhyvotovskiy, R., Sova, O., Zvieriev, O., Lanetskii, B., Shyshatskyi, A. (2020). Development of an advanced method of finding solutions for neuro-fuzzy expert systems of analysis of the radioelectronic situation. EUREKA: Physics and Engineering, 4, 78–89. doi: <https://doi.org/10.21303/2461-4262.2020.001353>
5. Zuiev, P., Zhyvotovskiy, R., Zvieriev, O., Hatsenko, S., Kuprii, V., Nakonechnyi, O. et al. (2020). Development of complex methodology of processing heterogeneous data in intelligent decision support systems. Eastern-European Journal of Enterprise Technologies, 4 (9 (106)), 14–23. doi: <https://doi.org/10.15587/1729-4061.2020.208554>
6. Shyshatskyi, A. (2020). Complex Methods of Processing Different Data in Intellectual Systems for Decision Sup-

- port System. *International Journal of Advanced Trends in Computer Science and Engineering*, 9 (4), 5583–5590. doi: <https://doi.org/10.30534/ijatcse/2020/206942020>
7. Yeromina, N., Kurban, V., Mykus, S., Peredrii, O., Voloshchenko, O., Kosenko, V. et al. (2021). The Creation of the Database for Mobile Robots Navigation under the Conditions of Flexible Change of Flight Assignment. *International Journal of Emerging Technology and Advanced Engineering*, 11 (5), 37–44. doi: https://doi.org/10.46338/ijetae0521_05
 8. Rotshteyn A. P. (1999). *Intellectual'nye tekhnologii identifikatsii: nechetkie mnozhestva, geneticheskie algoritmy, neyronnye seti*. Vinnitsa: «UNIVERSUM», 320.
 9. Alpeeva, E. A., Volkova, I. I. (2019). The use of fuzzy cognitive maps in the development of an experimental model of automation of production accounting of material flows. *Russian Journal of Industrial Economics*, 12 (1), 97–106. doi: <https://doi.org/10.17073/2072-1633-2019-1-97-106>
 10. Zagranovskaya, A. V., Eissner, Y. N. (2017). Simulation scenarios of the economic situation based on fuzzy cognitive maps. *Modern economics: problems and solutions*, 10, 33–47. doi: <https://doi.org/10.17308/meps.2017.10/1754>
 11. Simankov, V. S., Putyato, M. M. (2013). Issledovanie metodov kognitivnogo analiza. *Sistemnyy analiz, upravlenie i obrabotka informatsii*, 13, 31–35.
 12. Ko, Y.-C., Fujita, H. (2019). An evidential analytics for buried information in big data samples: Case study of semiconductor manufacturing. *Information Sciences*, 486, 190–203. doi: <https://doi.org/10.1016/j.ins.2019.01.079>
 13. Ramaji, I. J., Memari, A. M. (2018). Interpretation of structural analytical models from the coordination view in building information models. *Automation in Construction*, 90, 117–133. doi: <https://doi.org/10.1016/j.autcon.2018.02.025>
 14. Pérez-González, C. J., Colebrook, M., Roda-García, J. L., Rosa-Remedios, C. B. (2019). Developing a data analytics platform to support decision making in emergency and security management. *Expert Systems with Applications*, 120, 167–184. doi: <https://doi.org/10.1016/j.eswa.2018.11.023>
 15. Chen, H. (2018). Evaluation of Personalized Service Level for Library Information Management Based on Fuzzy Analytic Hierarchy Process. *Procedia Computer Science*, 131, 952–958. doi: <https://doi.org/10.1016/j.procs.2018.04.233>
 16. Chan, H. K., Sun, X., Chung, S.-H. (2019). When should fuzzy analytic hierarchy process be used instead of analytic hierarchy process? *Decision Support Systems*, 125, 113114. doi: <https://doi.org/10.1016/j.dss.2019.113114>
 17. Osman, A. M. S. (2019). A novel big data analytics framework for smart cities. *Future Generation Computer Systems*, 91, 620–633. doi: <https://doi.org/10.1016/j.future.2018.06.046>
 18. Gödri, I., Kardos, C., Pfeiffer, A., Váncza, J. (2019). Data analytics-based decision support workflow for high-mix low-volume production systems. *CIRP Annals*, 68 (1), 471–474. doi: <https://doi.org/10.1016/j.cirp.2019.04.001>
 19. Harding, J. L. (2013). Data quality in the integration and analysis of data from multiple sources: some research challenges. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-2/W1, 59–63. doi: <https://doi.org/10.5194/isprsarchives-xl-2-w1-59-2013>
 20. Kosko, B. (1986). Fuzzy cognitive maps. *International Journal of Man-Machine Studies*, 24 (1), 65–75. doi: [https://doi.org/10.1016/s0020-7373\(86\)80040-2](https://doi.org/10.1016/s0020-7373(86)80040-2)
 21. Gorelova, G. V. (2013). Cognitive approach to simulation of large systems. *Izvestiya YuFU. Tekhnicheskie nauki*, 3, 239–250.
 22. Koval, M., Sova, O., Shyshatskyi, A., Artabaiev, Y., Garashchuk, N., Yivzhenko, Y. et al. (2022). Improving the method for increasing the efficiency of decision-making based on bio-inspired algorithms. *Eastern-European Journal of Enterprise Technologies*, 6 (4 (120)), 6–13. doi: <https://doi.org/10.15587/1729-4061.2022.268621>
 23. Lagunova, A. D. (2019). Algoritm letuchikh myshey (BA) dlya zadachi global'noy bezuslovnoy optimizatsii. *Original Research Journal*, 9 (6), 101–116.
 24. Mahdi, Q. A., Shyshatskyi, A., Prokopenko, Y., Ivakhnenko, T., Kupriyenko, D., Golian, V. et al. (2021). Development of estimation and forecasting method in intelligent decision support systems. *Eastern-European Journal of Enterprise Technologies*, 3 (9 (111)), 51–62. doi: <https://doi.org/10.15587/1729-4061.2021.232718>
 25. Emel'yanov, V. V., Kureychik, V. V., Kureychik, V. M., Emel'yanov, V. V. (2003). *Teoriya i praktika evolyutsionnogo modelirovaniya*. Moscow: Fizmatlit, 432.
 26. Gorokhovatsky, V., Stiahlyk, N., Tsarevska, V. (2021). Combination method of accelerated metric data search in image classification problems. *Advanced Information Systems*, 5 (3), 5–12. doi: <https://doi.org/10.20998/2522-9052.2021.3.01>
 27. Levashenko, V., Liashenko, O., Kuchuk, H. (2020). Building Decision Support Systems based on Fuzzy Data. *Advanced Information Systems*, 4 (4), 48–56. doi: <https://doi.org/10.20998/2522-9052.2020.4.07>
 28. Meleshko, Y., Driev, O., Drieva, H. (2020). Method of identification bot profiles based on neural networks in recommendation systems. *Advanced Information Systems*, 4 (2), 24–28. doi: <https://doi.org/10.20998/2522-9052.2020.2.05>
 29. Kuchuk, N., Merlak, V., Skorodelov, V. (2020). A method of reducing access time to poorly structured data. *Advanced Information Systems*, 4 (1), 97–102. doi: <https://doi.org/10.20998/2522-9052.2020.1.14>
 30. Shyshatskyi, A., Tiurnikov, M., Suhak, S., Bondar, O., Melnyk, A., Bokhno, T., Lyashenko, A. (2020). Method of assessment of the efficiency of the communication of operational troop grouping system. *Advanced Information Systems*, 4 (1), 107–112. doi: <https://doi.org/10.20998/2522-9052.2020.1.16>
 31. Raskin, L., Sira, O. (2016). Method of solving fuzzy problems of mathematical programming. *Eastern-European Journal of Enterprise Technologies*, 5 (4 (83)), 23–28. doi: <https://doi.org/10.15587/1729-4061.2016.81292>
 32. Lytvyn, V., Vysotska, V., Pukach, P., Brodyak, O., Ugryn, D. (2017). Development of a method for determining the keywords in the slavic language texts based on the technology of web mining. *Eastern-European Journal of Enterprise Technologies*, 2 (2 (86)), 14–23. doi: <https://doi.org/10.15587/1729-4061.2017.98750>
 33. Stepanenko, A., Oliinyk, A., Deineha, L., Zaiko, T. (2018). Development of the method for decomposition of superpositions of unknown pulsed signals using the second-order adaptive spectral analysis. *Eastern-European Journal of Enterprise Technologies*, 2 (9 (92)), 48–54. doi: <https://doi.org/10.15587/1729-4061.2018.126578>
 34. Gorbenco, I., Ponomar, V. (2017). Examining a possibility to use and the benefits of post-quantum algorithms dependent on the conditions of their application. *Eastern-European Journal of Enterprise Technologies*, 2 (9 (86)), 21–32. doi: <https://doi.org/10.15587/1729-4061.2017.96321>
 35. Lovska, A. (2015). Peculiarities of computer modeling of strength of body bearing construction of gondola car during

transportation by ferry-bridge. *Metallurgical and Mining Industry*, 1, 49–54.

36. Lovska, A., Fomin, O. (2020). A new fastener to ensure the reliability of a passenger car body on a train ferry. *Acta Polytechnica*, 60 (6). doi: <https://doi.org/10.14311/ap.2020.60.0478>
37. Koval, M., Sova, O., Orlov, O., Shyshatskyi, A., Artabaiev, Y., Shknai, O. et al. (2022). Improvement of complex resource management of special-purpose communication systems. *Eastern-European Journal of Enterprise Technologies*, 5 (9 (119)), 34–44. doi: <https://doi.org/10.15587/1729-4061.2022.266009>
38. Koshlan, A., Salnikova, O., Chekhovska, M., Zhyvotovskiy, R., Prokopenko, Y., Hurskyi, T. et al. (2019). Development of an algorithm for complex processing of geospatial data in the special-purpose geoinformation system in conditions of diversity and uncertainty of data. *Eastern-European Journal of Enterprise Technologies*, 5 (9 (101)), 35–45. doi: <https://doi.org/10.15587/1729-4061.2019.180197>

DOI: 10.15587/1729-4061.2023.282130

STATISTICAL PROCESSING OF A SMALL SAMPLE OF RAW DATA USING ARTIFICIAL ORTHOGONALISATION TECHNOLOGY (p. 14–21)

Lev Raskin

National Technical University
«Kharkiv Polytechnic Institute», Kharkiv, Ukraine
ORCID: <https://orcid.org/0000-0002-9015-4016>

Larysa Sukhomlyn

Kremenchuk Mykhailo Ostrohradskiy
National University, Kremenchuk, Ukraine
ORCID: <https://orcid.org/0000-0001-9511-5932>

Viacheslav Karpenko

National Technical University
«Kharkiv Polytechnic Institute», Kharkiv, Ukraine
ORCID: <https://orcid.org/0000-0002-8378-129X>

Dmytro Sokolov

National Technical University
«Kharkiv Polytechnic Institute», Kharkiv, Ukraine
ORCID: <https://orcid.org/0000-0002-4558-9598>

This paper addresses the task to devise a statistical estimation procedure in an event where the volume of the array of initial data used in processing is insufficient to correctly determine the parameters of the response function. The object of research is the technology of statistical processing of a small sample of data. The subject of the study is the methods of statistical estimation under conditions of a small sample of initial data. The main direction is to devise a special procedure for statistical processing of a small sample of initial data, which provides a correct statistical estimation of the parameters of the response function. The method for solving the problem is the selection of the most representative orthogonal replica-like subplan from the plan of a complete factorial experiment obtained by artificially orthogonalizing the results of a passive experiment. The necessity and expediency of the proposed procedure is a consequence of the unpredictability and uneven distribution of points in the phase space of coordinates. The result of the implementation of the corresponding procedure is a truncated orthogonal plan of the full factorial experiment, which provides the

possibility of independent estimation of all coefficients of the regression polynomial describing the response function. Under conditions of a severe shortage of the number of measurements, the procedure makes it possible to isolate a representative orthogonal replica from the resulting plan of a complete factorial experiment. Using this subplan of the full factorial experiment plan makes it possible to evaluate all the coefficients of the regression polynomial that describes the desired response function. The corresponding computational procedure is based on solving the triaxial Boolean assignment problem.

Keywords: statistical data processing, small sample, artificial orthogonalization, triaxial assignment problem.

References

1. Shoba, K. (2019). Multiple diskriminant analysis. Arlington: Inst. for statistics Education, 112.
2. Aouati, M. (2017). Improvement of accuracy of parametric classification in the space of $N \times 2$ factors-attributes on the basis of preliminary obtained linear discriminant function. *EUREKA: Physics and Engineering*, 3, 55–68. doi: <https://doi.org/10.21303/2461-4262.2017.00362>
3. Luna-Romera, J. M., Martínez-Ballesteros, M., García-Gutiérrez, J., Riquelme, J. C. (2019). External clustering validity index based on chi-squared statistical test. *Information Sciences*, 487, 1–17. doi: <https://doi.org/10.1016/j.ins.2019.02.046>
4. Everitt, B. S., Landau, S., Leese, M., Stahl, D. (2011). *Cluster Analysis*. Wiley. doi: <https://doi.org/10.1002/9780470977811>
5. Aouati, M. (2018). Improving the accuracy of classifying rules for controlling the processes of deculfuration and dephosphorization of Fe-C melt. *Technology Audit and Production Reserves*, 2 (3 (46)), 10–18. doi: <https://doi.org/10.15587/2312-8372.2019.169696>
6. Mourad, A. (2017). Parametric identification in the problem of determining the quality of dusulfuration and dephosphoration processes of Fe-C alloy. *Technology Audit and Production Reserves*, 2 (1 (34)), 9–15. doi: <https://doi.org/10.15587/2312-8372.2017.99130>
7. Uotermen, D. (1989). *Rukovodstvo po ekspertnym sistemam*. Moscow: Mir, 388.
8. Dzhekson, P. (2014). *Vvedenie v ekspertnye sistemy*. Moscow: Vil'yams, 624.
9. Dzharratano, D., Rayli, G. (2016). *Ekspertnye sistemy*. Moscow: Vil'yams, 1152.
10. Gavrilova, T. A., Khoroshevskiy, V. F. (2011). *Bazy znaniy intellektual'nykh sistem*. Sankt-Peterburg: Piter, 384.
11. Oimoen, S. (2019). *Classical Designs: Full Factorial Designs*. STAT Center of Excellence, 26. Available at: https://www.ait.edu/stat/statcoe_files/Classical%20Designs-Full%20Factorial%20Designs_Final.pdf
12. Montgomery, D. (2013). *Design and analysis of experiments*. Wiley.
13. Burman, L. E., Reed, W. R., Alm, J. (2010). A Call for Replication Studies. *Public Finance Review*, 38 (6), 787–793. doi: <https://doi.org/10.1177/1091142110385210>
14. Hari, R. (2022). *Replication study*.
15. Narayan, C., Das, M. (2009). *Design and analysis of Experiments*. New of Experiments. Wiley, 53–76.
16. Kullback, S. (1959). *Information Theory and statistics*. Willey.
17. Seraya, O. V., Demin, D. A. (2012). Lineynyy regressionnyy analiz maloy vyborki nechetkikh iskhodnykh dannykh. *Problemy upravleniya i informatiki*, 4, 129–142.

18. Domin, D., Sira, O., Raskin, L. (2021). Artificial orthogonalization of a passive experiment for a small sample of fuzzy data for constructing regression equations. Available at: <https://ingraph.org/en/products/212>
19. Raskin, L. G. (1976). Analiz slozhnykh sistem i elementy teorii optimal'nogo upravleniya. Moscow: Sov. Radio, 344.
20. Domin, D. (2013). Artificial orthogonalization in searching of optimal control of technological processes under uncertainty conditions. *Eastern-European Journal of Enterprise Technologies*, 5 (9(65)), 45–53. doi: <https://doi.org/10.15587/1729-4061.2013.18452>
21. Adler, Yu. P., Markova, E. V., Granovskiy, Yu. V. (1971). Planirovanie eksperimenta pri poiske optimal'nykh usloviy. Moscow: Nauka, 282.
22. Domin, D., Sira, O., Raskin, L. (2021). Technology for constructing regression equations for a small sample of passive experiment data. Available at: <https://ingraph.org/en/products/211>
23. Raskin, L. G. (1982). Mnogoindeksnyye zadachi lineynogo programmirovaniya. Moscow: Radio i svyaz', 246.

DOI: 10.15587/1729-4061.2023.282645
DETECTION OF FIXED FOOTBALL MATCHES
BASED ON THE THEORY OF CONFORMAL
PREDICTORS USING THE MODIFIED STEPANETS
INDICATOR FUNCTION (p. 22–32)

Oleg Chertov

National Technical University of Ukraine
 «Igor Sikorsky Kyiv Polytechnic Institute», Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0003-0087-1028>

Ivan Zhuk

National Technical University of Ukraine
 «Igor Sikorsky Kyiv Polytechnic Institute», Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0001-8440-427X>

An urgent issue of modern football competitions is the detection of fixed matches. Known methods for predicting the outcome of a match by analyzing bets on a match or analyzing the actions of football players on the field use a large amount of data that is not always available. To overcome this obstacle, there can be applied a method for detecting suspicious fixed match results based on conformal predictors and power martingales, which uses publicly available public data. But in practice, this method does not always detect such matches with high precision. An improved method for determining a suspicious match is proposed, based on the theory of conformal predictors using a modified Stepanets indicator function, which is compared with a threshold. The modified Stepanets indicator function is applied to the power martingale and shows the relative change in the martingale value of the current match compared to the previous match. The threshold value was determined experimentally according to the criterion of the maximum of the $F1$ metric. Data from the 2013–2014 season of the French II League were used as a training sample, and data from the 2014–2015 season of Serie B in Italy were used as a test sample. Team clustering was performed on all samples. For each of the formed classes of matches on both samples, the measure of non-conformity, the degree of non-conformity, the power martingale, and the modified Stepanets indicator function were calculated. The resulting indicators of precision metrics and $F1$ are higher (average values of metrics $P=0.84$, $F1=0.87$) than the same indica-

tors of martingale and p -value rules (average values of metrics $P=0.75$, $F1=0.78$), applied to the same data. The proposed method reveals 4 out of 5 matches of the 2014–2015 Serie B season in Italy, which are considered fixed according to information from official Italian law enforcement sources.

Keywords: fixed result, power martingale, measure of non-conformity, offline algorithm, p -value, $F1$ metric.

References

1. Abarbanel, B., Johnson, M. R. (2018). Esports consumer perspectives on match-fixing: implications for gambling awareness and game integrity. *International Gambling Studies*, 19 (2), 296–311. doi: <https://doi.org/10.1080/14459795.2018.1558451>
2. Lilley, E. (2015). A Review of the recommendations of the 'Report of the Sports Betting Integrity Panel' in assessing the progress towards tackling Match-fixing in Sport. *Laws of the Game*, 1 (1).
3. Huggins, M. (2018). Match-Fixing: A Historical Perspective. *The International Journal of the History of Sport*, 35 (2-3), 123–140. doi: <https://doi.org/10.1080/09523367.2018.1476341>
4. Forrest, D., McHale, I. G. (2019). Using statistics to detect match fixing in sport. *IMA Journal of Management Mathematics*, 30 (4), 431–449. doi: <https://doi.org/10.1093/imaman/dpz008>
5. Razali, N., Mustapha, A., Yatim, F. A., Ab Aziz, R. (2017). Predicting Football Matches Results using Bayesian Networks for English Premier League (EPL). *IOP Conference Series: Materials Science and Engineering*, 226, 012099. doi: <https://doi.org/10.1088/1757-899x/226/1/012099>
6. Anfilets, S., Bezobrazov, S., Golovko, V., Sachenko, A., Komar, M., Dolny, R., Kasyanik, V. et al. (2020). Deep multilayer neural network for predicting the winner of football matches. *International Journal of Computing*, 19 (1), 70–77. doi: <https://doi.org/10.47839/ijc.19.1.1695>
7. Spapens, T., Olfers, M. (2015). Match-fixing: The Current Discussion in Europe and the Case of The Netherlands. *European Journal of Crime, Criminal Law and Criminal Justice*, 23 (4), 333–358. doi: <https://doi.org/10.1163/15718174-23032077>
8. Stepanets, A. I. (2005). *Methods of Approximation Theory*. Boston. doi: <https://doi.org/10.1515/9783110195286>
9. Chertov, O., Zhuk, I., Serdyuk, A. (2021). Search of the Deviation from the Natural Process Using Stepanets Approach for Classification of Functions. 2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). doi: <https://doi.org/10.1109/idaacs53288.2021.9660997>
10. Vovk, V. (2014). The Basic Conformal Prediction Framework. *Conformal Prediction for Reliable Machine Learning*, 3–19. doi: <https://doi.org/10.1016/b978-0-12-398537-8.00001-8>
11. Ho, S.-S., Wechsler, H. (2010). A Martingale Framework for Detecting Changes in Data Streams by Testing Exchangeability. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32 (12), 2113–2127. doi: <https://doi.org/10.1109/tpami.2010.48>
12. Zhuk, I., Chertov, O. (2023). Framework based on conformal predictors and power martingales for detection of fixed football matches. *Eastern-European Journal of Enterprise Technologies*, 2 (4 (122)), 6–15. doi: <https://doi.org/10.15587/1729-4061.2023.276977>

13. Historical analysis of closing odds. Available at: <https://github.com/Lisandro79/BeatTheBookie>
14. Catania's owner admits to match fixing in five Serie B games. The Guardian. Available at: <https://www.theguardian.com/football/2015/jun/30/catania-match-fixing-serie-b>

DOI: 10.15587/1729-4061.2023.279622

DEVELOPMENT OF A NEW MULTI-CRITERIA DECISION-MAKING METHOD (p. 33–38)

Tran Van Dua

Hanoi University of Industry, Hanoi, Vietnam

ORCID: <https://orcid.org/0000-0001-6435-720X>

Choosing the best option out of the many available options is always the goal to be achieved in all areas. However, the parameters (criteria) in each alternative are not the same, sometimes contradictory. In this situation, choosing the best option is an extremely difficult decision for the decision maker. Multi-criteria decision making (MCDM) is the ranking of alternatives based on the criteria of each alternative. More than one hundred multi-criteria decision-making methods have been proposed by the inventors. They are being used in many different fields. However, for decision makers, choosing an appropriate method to use in each specific case is a difficult task. CURLI (Collaborative Unbiased Rank List Integration) is a multi-criteria decision making method that distinguishes it from all others. That difference is reflected in the fact that when applying this method, the decision maker does not need to normalize the data nor determine the weights for the criteria. However, it will take a long time for decision makers to apply this method, especially when the number of options to rank is large. This study carried out the development of a new MCDM method based on the CURLI method. This new method is named CURLI-2. Many different examples are presented to evaluate the effectiveness of the proposed method. In each example, the result of ranking the alternatives using the CURLI-2 method has been compared with those using other different MCDM methods. The best alternative determined when using the CURLI-2 method always coincides with the use of existing MCDM methods. Using CURLI-2 method to rank alternatives will be much faster and simpler than using CURLI method. This is the advantage of CURLI-2 method compared with CURLI method.

Keywords: new multi criteria decision making method, CURLI-2 method, CURLI method.

References

1. Biswas, S., Pamucar, D. (2021). Combinative distance based assessment (CODAS) framework using logarithmic normalization for multi-criteria decision making. *Serbian Journal of Management*, 16 (2), 321–340. doi: <https://doi.org/10.5937/sjm16-27758>
2. Alrababah, S. A. A., Atyeh, A. J. (2019). Effect of Normalization Techniques in VIKOR Approach for Mining Product Aspects in Customer Reviews. *International Journal of Computer Science and Network Security*, 19 (12), 112–118. Available at: http://paper.ijcns.org/07_book/201912/20191216.pdf
3. Vafaei, N., Ribeiro, R. A., Camarinha-Matos, L. M. (2022). Assessing Normalization Techniques for Simple Additive Weighting Method. *Procedia Computer Science*, 199, 1229–1236. doi: <https://doi.org/10.1016/j.procs.2022.01.156>
4. Ersoy, N. (2021). Application of the PIV method in the presence of negative data: an empirical example from a real-world case. *Hitit Sosyal Bilimler Dergisi*, 14 (2), 318–337. doi: <https://doi.org/10.17218/hititsbd.974522>
5. Ersoy, N. (2021). Selecting the Best Normalization Technique for ROV Method: Towards a Real Life Application. *Gazi University Journal of Science*, 34 (2), 592–609. doi: <https://doi.org/10.35378/gujs.767525>
6. Stanujkić, D., Đorđević, B., Đorđević, M. (2013). Comparative analysis of some prominent MCDM methods: A case of ranking Serbian banks. *Serbian Journal of Management*, 8 (2), 213–241. doi: <https://doi.org/10.5937/sjm8-3774>
7. Vafaei, N., Ribeiro, R. A., Camarinha-Matos, L. M. (2016). Normalization Techniques for Multi-Criteria Decision Making: Analytical Hierarchy Process Case Study. *IFIP Advances in Information and Communication Technology*, 261–269. doi: https://doi.org/10.1007/978-3-319-31165-4_26
8. Mathew, M., Sahu, S., Upadhyay, A. K. (2017). Effect Of Normalization Techniques In Robot Selection Using Weighted Aggregated Sum Product Assessment. *International Journal of Innovative Research and Advanced Studies*, 4 (2), 59–63. Available at: http://www.ijiras.com/2017/Vol_4-Issue_2/paper_12.pdf
9. Duc Trung, D. (2022). Multi-criteria decision making of turning operation based on PEG, PSI and CURLI methods. *Manufacturing Review*, 9, 9. doi: <https://doi.org/10.1051/mfreview/2022007>
10. Trung, D. D. (2022). Development of data normalization methods for multi-criteria decision making: applying for MARCOS method. *Manufacturing Review*, 9, 22. doi: <https://doi.org/10.1051/mfreview/2022019>
11. Trung, D., Ba, N., Tien, D. (2022). Application of the Curli method for multi-critical decision of grinding process. *Journal of Applied Engineering Science*, 20 (3), 634–643. doi: <https://doi.org/10.5937/jaes0-35088>
12. Tran, D. V. (2022). Application of the Collaborative Unbiased Rank List Integration Method to Select the Materials. *Applied Engineering Letters: Journal of Engineering and Applied Sciences*, 7 (4), 133–142. doi: <https://doi.org/10.18485/aeletters.2022.7.4.1>
13. Ilangkumaran, M., Avenash, A., Balakrishnan, V., Kumar, S. B., Raja, M. B. (2013). Material selection using hybrid MCDM approach for automobile bumper. *International Journal of Industrial and Systems Engineering*, 14 (1), 20. doi: <https://doi.org/10.1504/ijise.2013.052919>
14. Prasenjit, C., Arnab, B., Supraksh, M., Soumava, B., Shankar, C. (2018). Development of a Hybrid Meta-Model for Material Selection Using Design of Experiments and EDAS Method. *Engineering Transactions*, 66 (2), 187–207. Available at: <https://et.ippt.gov.pl/index.php/et/article/view/812/>
15. Milani, A. S., Shanian, A., Madoliat, R., Nemes, J. A. (2004). The effect of normalization norms in multiple attribute decision making models: a case study in gear material selection. *Structural and Multidisciplinary Optimization*, 29 (4), 312–318. doi: <https://doi.org/10.1007/s00158-004-0473-1>
16. Chatterjee, P., Chakraborty, S. (2012). Material selection using preferential ranking methods. *Materials & Design*, 35, 384–393. doi: <https://doi.org/10.1016/j.matdes.2011.09.027>
17. Trung, D. D. (2022). Comparison r and curli methods for multi-criteria decision making. *Advanced Engineering Letters*, 1 (2), 46–56. doi: <https://doi.org/10.46793/adeletters.2022.1.2.3>
18. Ghorabae, M. K., Zavadskas, E. K., Turskis, Z., Antucheviciene, J. (2016). A new combinative distance-based assessment (CODAS) method for multi-criteria decision-making.

ing. Economic Computation and Economic Cybernetics Studies and Research, 50 (3), 25–44. Available at: https://www.researchgate.net/publication/308697546_A_new_combinative_distance-based_assessment_CODAS_method_for_multi-criteria_decision-making

DOI: 10.15587/1729-4061.2023.281007

DEVELOPMENT OF THE SOLUTION SEARCH METHOD USING THE POPULATION ALGORITHM OF GLOBAL SEARCH OPTIMIZATION (p. 39–46)

Stepan Yakymiak

The National Defence University of Ukraine named after Ivan Cherniakhovskiy, Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0002-1530-271X>

Yevhenii Vdovytskyi

The National Defence University of Ukraine named after Ivan Cherniakhovskiy, Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0003-0930-525X>

Yurii Artabaiev

Research Center for Trophy and Perspective Weapons and Military Equipment, Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0001-9446-3011>

Larisa Degtyareva

Poltava State Agrarian University, Poltava, Ukraine
ORCID: <https://orcid.org/0000-0001-5927-9550>

Yuliia Vakulenko

Poltava State Agrarian University, Poltava, Ukraine
ORCID: <https://orcid.org/0000-0002-6315-0116>

Serhii Nevhad

The National Defence University of Ukraine named after Ivan Cherniakhovskiy, Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0002-1557-9831>

Vitalii Andronov

Scientific-Research Institute of Military Intelligence, Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0002-1122-710X>

Roman Lazuta

Military Institute of Telecommunications and Information Technologies named after Heroes of Kruty, Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0003-3254-9690>

Petro Shapoval

Central Scientifically-Research Institute of Armaments and Military Equipment of the Armed Forces of Ukraine, Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0002-8892-0228>

Yevhen Artamonov

National Aviation University, Kyiv, Ukraine
ORCID: <https://orcid.org/0000-0002-9875-7372>

The research objects are decision making support systems. subject research is a decision making process in management tasks using bio-inspired algorithms. A method of finding solutions using the population algorithm of global search optimization is proposed. Joint use of invasive weed algorithm is proposed, genetic algorithm and evolving ar-

tificial neural networks are improved. The method has the following sequence of actions:

- an input of initial data;
- processing of initial data taking into account the degree of uncertainty;
- formation of the optimization vector;
- creation of descendant vectors;
- ordering of vectors in descending order;
- reducing the dimensionality of the feature space;
- teaching knowledge bases.

The peculiarity of the proposed method lies in the placement of agents-weeds, taking into account the uncertainty of the initial data, improved procedures for reducing the space of signs about the analysis object state.

Training of synaptic weights of an artificial neural network, type and parameters of the membership function and the architecture of individual elements, and the architecture of an artificial neural network as a whole is carried out. The proposed method was simulated in the MathCad 14 software environment. The task to be solved during the simulation was to determine the route of the ships in the operational zones of the Black and Azov seas in the conditions of hybrid actions of the enemy. The use of the method makes it possible to increase the efficiency of data processing at the level of 21–27 % due to the use of additional improved procedures. The proposed method should be used to solve the problems of evaluating complex and dynamic processes in the interests of solving national security problems.

Keywords: hybrid actions, artificial neural networks, bio-inspired algorithms, use of forces (troops) of the Naval Forces.

References

1. Bashkyrov, O. M., Kostyna, O. M., Shyshatskyi, A. V. (2015). Rozvytok integrovanykh system zviazku ta poredachi danykh dlia potreb Zbroinykh Syl. Ozbroyennia ta viyskova tekhnika, 1, 35–39. Available at: http://nbuv.gov.ua/UJRN/ovt_2015_1_7
2. Dudnyk, V., Sinenko, Y., Matsyk, M., Demchenko, Y., Zhyvotovskiy, R., Repilo, I. et al. (2020). Development of a method for training artificial neural networks for intelligent decision support systems. Eastern-European Journal of Enterprise Technologies, 3 (2 (105)), 37–47. doi: <https://doi.org/10.15587/1729-4061.2020.203301>
3. Sova, O., Shyshatskyi, A., Salnikova, O., Zhuk, O., Trotsko, O., Hrokholskyi, Y. (2021). Development of a method for assessment and forecasting of the radio electronic environment. EUREKA: Physics and Engineering, 4, 30–40. doi: <https://doi.org/10.21303/2461-4262.2021.001940>
4. Pievtsov, H., Turinskyi, O., Zhyvotovskiy, R., Sova, O., Zvieriev, O., Lanetskii, B., Shyshatskyi, A. (2020). Development of an advanced method of finding solutions for neuro-fuzzy expert systems of analysis of the radioelectronic situation. EUREKA: Physics and Engineering, 4, 78–89. doi: <https://doi.org/10.21303/2461-4262.2020.001353>
5. Zuiev, P., Zhyvotovskiy, R., Zvieriev, O., Hatsenko, S., Kuprii, V., Nakonechnyi, O. et al. (2020). Development of complex methodology of processing heterogeneous data in intelligent decision support systems. Eastern-European Journal of Enterprise Technologies, 4 (9 (106)), 14–23. doi: <https://doi.org/10.15587/1729-4061.2020.208554>
6. Shyshatskyi, A., Zvieriev, O., Salnikova, O., Demchenko, Ye., Trotsko, O., Neroznak, Ye. (2020). Complex Methods of Processing Different Data in Intellectual Systems for Decision

- Support System. *International Journal of Advanced Trends in Computer Science and Engineering*, 9 (4), 5583–5590. doi: <https://doi.org/10.30534/ijatcse/2020/206942020>
7. Yeromina, N., Kurban, V., Mykus, S., Peredrii, O., Voloshchenko, O., Kosenko, V. et al. (2021). The Creation of the Database for Mobile Robots Navigation under the Conditions of Flexible Change of Flight Assignment. *International Journal of Emerging Technology and Advanced Engineering*, 11 (05), 37–44. doi: https://doi.org/10.46338/ijetae0521_05
 8. Rotshteyn, A. P. (1999). *Intellectual'nye tekhnologii identifikatsii: nechetkie mnozhestva, geneticheskie algoritmy, neyronnye seti*. Vinnitsa: «UNIVERSUM», 320.
 9. Alpeeva, E. A., Volkova, I. I. (2019). The use of fuzzy cognitive maps in the development of an experimental model of automation of production accounting of material flows. *Russian Journal of Industrial Economics*, 12 (1), 97–106. doi: <https://doi.org/10.17073/2072-1633-2019-1-97-106>
 10. Zagranovskaya, A. V., Eyssner, Yu. N. (2017). Modelirovanie stsensariiev razvitiya ekonomicheskoy situatsii na osnove nechetkikh kognitivnykh kart. *Sovremennaya Ekonomika: Problemy i Resheniya*, 10, 33–47. doi: <https://doi.org/10.17308/meps.2017.10/1754>
 11. Simankov, V. S., Putyato, M. M. (2013). Issledovanie metodov kognitivnogo analiza. *Sistemnyy analiz, upravlenie i obrabotka informatsii*, 13, 31–35.
 12. Ko, Y.-C., Fujita, H. (2019). An evidential analytics for buried information in big data samples: Case study of semiconductor manufacturing. *Information Sciences*, 486, 190–203. doi: <https://doi.org/10.1016/j.ins.2019.01.079>
 13. Ramaji, I. J., Memari, A. M. (2018). Interpretation of structural analytical models from the coordination view in building information models. *Automation in Construction*, 90, 117–133. doi: <https://doi.org/10.1016/j.autcon.2018.02.025>
 14. Pérez-González, C. J., Colebrook, M., Roda-García, J. L., Rosa-Remedios, C. B. (2019). Developing a data analytics platform to support decision making in emergency and security management. *Expert Systems with Applications*, 120, 167–184. doi: <https://doi.org/10.1016/j.eswa.2018.11.023>
 15. Chen, H. (2018). Evaluation of Personalized Service Level for Library Information Management Based on Fuzzy Analytic Hierarchy Process. *Procedia Computer Science*, 131, 952–958. doi: <https://doi.org/10.1016/j.procs.2018.04.233>
 16. Chan, H. K., Sun, X., Chung, S.-H. (2019). When should fuzzy analytic hierarchy process be used instead of analytic hierarchy process? *Decision Support Systems*, 125, 113114. doi: <https://doi.org/10.1016/j.dss.2019.113114>
 17. A Osman, A. M. S. (2019). A novel big data analytics framework for smart cities. *Future Generation Computer Systems*, 91, 620–633. doi: <https://doi.org/10.1016/j.future.2018.06.046>
 18. Gödri, I., Kardos, C., Pfeiffer, A., Vánca, J. (2019). Data analytics-based decision support workflow for high-mix low-volume production systems. *CIRP Annals*, 68 (1), 471–474. doi: <https://doi.org/10.1016/j.cirp.2019.04.001>
 19. Harding, J. L. (2013). Data quality in the integration and analysis of data from multiple sources: some research challenges. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-2/W1, 59–63. doi: <https://doi.org/10.5194/isprsarchives-XL-2-W1-59-2013>
 20. Kosko, B. (1986). Fuzzy cognitive maps. *International Journal of Man-Machine Studies*, 24 (1), 65–75. doi: [https://doi.org/10.1016/S0020-7373\(86\)80040-2](https://doi.org/10.1016/S0020-7373(86)80040-2)
 21. Gorelova, G. V. (2013). Kognitivnyy podkhod k imitatsionnomu modelirovaniyu slozhnykh sistem. *Izvestiya YuFU. Tekhnicheskie nauki*, 3, 239–250.
 22. Koval, M., Sova, O., Shyshatskyi, A., Artabaiev, Y., Garashchuk, N., Yivzhenko, Y. et al. (2022). Improving the method for increasing the efficiency of decision-making based on bio-inspired algorithms. *Eastern-European Journal of Enterprise Technologies*, 6 (4 (120)), 6–13. doi: <https://doi.org/10.15587/1729-4061.2022.268621>
 23. Rad, H. S., Lucas, C. (2007). A recommender system based on invasive weed optimization algorithm. *2007 IEEE Congress on Evolutionary Computation*. doi: <https://doi.org/10.1109/cec.2007.4425032>
 24. Koshlan, A., Salnikova, O., Chekhovska, M., Zhyvotovskiy, R., Prokopenko, Y., Hurskyi, T. et al. (2019). Development of an algorithm for complex processing of geospatial data in the special-purpose geoinformation system in conditions of diversity and uncertainty of data. *Eastern-European Journal of Enterprise Technologies*, 5 (9 (101)), 35–45. doi: <https://doi.org/10.15587/1729-4061.2019.180197>
 25. Mahdi, Q. A., Shyshatskyi, A., Prokopenko, Y., Ivakhnenko, T., Kupriyenko, D., Golian, V. et al. (2021). Development of estimation and forecasting method in intelligent decision support systems. *Eastern-European Journal of Enterprise Technologies*, 3 (9 (111)), 51–62. doi: <https://doi.org/10.15587/1729-4061.2021.232718>
 26. Emel'yanov, V. V., Kureychik, V. V., Kureychik, V. M., Emel'yanov, V. V. (2003). *Teoriya i praktika evolyutsionnogo modelirovaniya*. Moscow: Fizmatlit, 432.
 27. Gorokhovatsky, V., Stiahlyk, N., Tsarevska, V. (2021). Combination method of accelerated metric data search in image classification problems. *Advanced Information Systems*, 5 (3), 5–12. doi: <https://doi.org/10.20998/2522-9052.2021.3.01>
 28. Levashenko, V., Liashenko, O., Kuchuk, H. (2020). Building Decision Support Systems based on Fuzzy Data. *Advanced Information Systems*, 4 (4), 48–56. doi: <https://doi.org/10.20998/2522-9052.2020.4.07>
 29. Meleshko, Y., Drieiev, O., Drieieva, H. (2020). Method of identification bot profiles based on neural networks in recommendation systems. *Advanced Information Systems*, 4 (2), 24–28. doi: <https://doi.org/10.20998/2522-9052.2020.2.05>
 30. Kuchuk, N., Merlak, V., Skorodelov, V. (2020). A method of reducing access time to poorly structured data. *Advanced Information Systems*, 4 (1), 97–102. doi: <https://doi.org/10.20998/2522-9052.2020.1.14>
 31. Shyshatskyi, A., Tiurnikov, M., Suhak, S., Bondar, O., Melnyk, A., Bokhno, T., Lyashenko, A. (2020). Method of assessment of the efficiency of the communication of operational troop grouping system. *Advanced Information Systems*, 4 (1), 107–112. doi: <https://doi.org/10.20998/2522-9052.2020.1.16>
 32. Raskin, L., Sira, O. (2016). Method of solving fuzzy problems of mathematical programming. *Eastern-European Journal of Enterprise Technologies*, 5 (4 (83)), 23–28. doi: <https://doi.org/10.15587/1729-4061.2016.81292>
 33. Lytyvn, V., Vysotska, V., Pukach, P., Brodyak, O., Ugryn, D. (2017). Development of a method for determining the keywords in the slavic language texts based on the technology of web mining. *Eastern-European Journal of Enterprise Technologies*, 2 (2 (86)), 14–23. doi: <https://doi.org/10.15587/1729-4061.2017.98750>
 34. Stepanenko, A., Oliynyk, A., Deineha, L., Zaiko, T. (2018). Development of the method for decomposition of superpositions of unknown pulsed signals using the secondorder

- adaptive spectral analysis. *Eastern-European Journal of Enterprise Technologies*, 2 (9 (92)), 48–54. doi: <https://doi.org/10.15587/1729-4061.2018.126578>
35. Gorbenko, I., Ponomar, V. (2017). Examining a possibility to use and the benefits of post-quantum algorithms dependent on the conditions of their application. *Eastern-European Journal of Enterprise Technologies*, 2 (9 (86)), 21–32. doi: <https://doi.org/10.15587/1729-4061.2017.96321>
 36. Lovska, A. A. (2015). Peculiarities of computer modeling of strength of body bearing construction of gondola car during transportation by ferry-bridge. *Metallurgical and Mining Industry*, 1, 49–54. Available at: https://www.metaljournal.com.ua/assets/Journal/english-edition/MMI_2015_1/10%20Lovska.pdf
 37. Lovska, A., Fomin, O. (2020). A new fastener to ensure the reliability of a passenger car body on a train ferry. *Acta Polytechnica*, 60 (6). doi: <https://doi.org/10.14311/ap.2020.60.0478>
 38. Koval, M., Sova, O., Orlov, O., Shyshatskyi, A., Artabaiev, Y., Shknai, O. et al. (2022). Improvement of complex resource management of special-purpose communication systems. *Eastern-European Journal of Enterprise Technologies*, 5 (9 (119)), 34–44. doi: <https://doi.org/10.15587/1729-4061.2022.266009>
 39. Syrotenko, A. M. (Red.) (2020). *Voienni aspekty protydyi «hibrydniy ahresiyi»: dosvid Ukrainy*. Kyiv: NUOU imeni Ivana Cherniakhovskoho, 176. Available at: https://nuou.org.ua/assets/monography/mono_gibr_viin.pdf
 40. Yakymiak, S. (2021). Hybrid warfare in the Black Sea: lessons learned and training improvement. *Current issues of military specialists training in the security and defence sector under conditions of hybrid threats*. Warszawa: Wydawnictwo Instytutu Bezpieczeństwa I Rozwoju Międzynarodowego, 396–405. Available at: https://www.academia.edu/47706282/CURRENT_ISSUES_OF_MILITARY_SPECIALISTS_TRAINING_IN_THE_SECURITY_AND_DEFENCE_SECTOR_UNDER_CONDITIONS_OF_HYBRID_THREATS
 41. Yakymyak, S., Vdovitsky, E. (2021). Analysis of factors that may affect on the effectiveness of use the Navy duties during the protection of the economic activities of the state at sea in the context of hybrid enemy actions. *Збірник Наукових Prats Tsentru Voinno-Stratehichnykh Doslidzhen NUOU Imeni Ivana Cherniakhovskoho*, 1 (71), 75–81. doi: <https://doi.org/10.33099/2304-2745/2021-1-71/75-81>
 42. Vdovytskyi, Ye. A. (2022). Analiz isnuichoho naukovometodychnoho aparatu otsiniuvannia efektyvnosti zastovuvannia syl (viysk) VMS pid chas vykonannia zavdan zakhystu ekonomichnoi diyalnosti derzhavy na mori v umovakh hibrydnykh diy protyvnyka. *Trudy universytetu*, 3 (172), 202–208.

DOI: 10.15587/1729-4061.2023.279936
UNBALANCED CREDIT FRAUD MODELING
BASED ON BAGGING AND BAYESIAN
OPTIMIZATION (p. 47–53)

Mohammed A. Kashmoola

University of Al-Hamdaniya, Bartela, Hamdaniya, Iraq
ORCID: <https://orcid.org/0000-0003-3487-5908>

Samah Fakhri Aziz

University of Al-Hamdaniya, Bartela, Hamdaniya, Iraq
ORCID: <https://orcid.org/0000-0003-0754-5727>

Hasan Mudhafar Qays

University of Al-Hamdaniya, Bartela, Hamdaniya, Iraq
ORCID: <https://orcid.org/0000-0002-2040-207X>

Naors Y. Anad Alsaleem

University of Al-Hamdaniya, Bartela, Hamdaniya, Iraq
ORCID: <https://orcid.org/0000-0002-0785-2674>

Credit fraud modeling is a crucial area of research that is highly relevant to the credit loan industry. Effective risk management is a key factor in providing quality credit services and directly impacts the profitability and bad debt ratio of leading organizations in this sector. However, when the distribution of credit fraud data is highly unbalanced, it can lead to noise errors caused by information distortion, periodic statistical errors, and model biases during training. This can cause unfair results for the minority class (target class) and increase the risk of overfitting. While traditional data balancing methods can reduce bias in models towards the majority class in relatively unbalanced data, they may not be effective in highly unbalanced scenarios. To address this challenge, this paper proposes using Bagging algorithms such as Random Forest and Bagging to model highly unbalanced credit fraud data. Bayesian optimization is utilized to find hyperparameters and determine the accuracy of the minority class as an optimization function for the model, which is tested with real European credit card fraud data. The results of the proposed packing algorithms are compared with traditional data balancing methods such as Balanced Bagging and Balanced Random Forest. The study found that traditional data balancing methods may not be compatible with excessively unbalanced data, whereas Bagging algorithms show promise as a solution for modeling such data. The proposed method for finding hyperparameters effectively deals with highly unbalanced data. It achieved precision, recall, and F1-score for the minority category of 0.94, 0.81, and 0.87, respectively. The study emphasizes the importance of addressing the challenges associated with unbalanced credit fraud data to improve the accuracy and fairness of credit fraud models.

Keywords: unbalanced data, Bayesian optimization, random forest, majority and minority class.

References

1. Patil, Dr. A. (2018). Digital Transformation in Financial Services and Challenges and Opportunities. *International Journal of Trend in Scientific Research and Development*, Special Issue (Special Issue-ICDEBI2018), 10–12. doi: <https://doi.org/10.31142/ijtsrd18661>
2. McNulty, D., Milne, A. (2021). Bigger Fish to Fry: FinTech and the Digital Transformation of Financial Services. *Disruptive Technology in Banking and Finance*, 263–281. doi: https://doi.org/10.1007/978-3-030-81835-7_10
3. Lokanan, M. E., Sharma, K. (2022). Fraud prediction using machine learning: The case of investment advisors in Canada. *Machine Learning with Applications*, 8, 100269. doi: <https://doi.org/10.1016/j.mlwa.2022.100269>
4. Roberts, P. (2022). Risk Management Systems. *Simplifying Risk Management*, 43–66. doi: <https://doi.org/10.4324/9781003225157-3>
5. Meng, C., Zhou, L., Liu, B. (2020). A Case Study in Credit Fraud Detection With SMOTE and XGBoost. *Journal of Physics: Conference Series*, 1601 (5), 052016. doi: <https://doi.org/10.1088/1742-6596/1601/5/052016>

6. Lin, T.-H., Jiang, J.-R. (2021). Credit Card Fraud Detection with Autoencoder and Probabilistic Random Forest. *Mathematics*, 9 (21), 2683. doi: <https://doi.org/10.3390/math9212683>
7. Shamsudin, H., Yusof, U. K., Jayalakshmi, A., Akmal Khalid, M. N. (2020). Combining oversampling and under-sampling techniques for imbalanced classification: A comparative study using credit card fraudulent transaction dataset. 2020 IEEE 16th International Conference on Control & Automation (ICCA). doi: <https://doi.org/10.1109/icca51439.2020.9264517>
8. Ahammad, J., Hossain, N., Alam, M. S. (2020). Credit Card Fraud Detection using Data Pre-processing on Imbalanced Data – both Oversampling and Undersampling. *Proceedings of the International Conference on Computing Advancements*. doi: <https://doi.org/10.1145/3377049.3377113>
9. Praveen Mahesh, K., Ashar Afrouz, S., Shaju Areeckal, A. (2022). Detection of fraudulent credit card transactions: A comparative analysis of data sampling and classification techniques. *Journal of Physics: Conference Series*, 2161 (1), 012072. doi: <https://doi.org/10.1088/1742-6596/2161/1/012072>
10. Wang, S., Dai, Y., Shen, J., Xuan, J. (2021). Research on expansion and classification of imbalanced data based on SMOTE algorithm. *Scientific Reports*, 11 (1). doi: <https://doi.org/10.1038/s41598-021-03430-5>
11. Klikowski, J. Woźniak, M. (2022). Deterministic Sampling Classifier with weighted Bagging for drifted imbalanced data stream classification. *Applied Soft Computing*, 122, 108855. doi: <https://doi.org/10.1016/j.asoc.2022.108855>
12. Kim, M., Hwang, K.-B. (2022). An empirical evaluation of sampling methods for the classification of imbalanced data. *PLOS ONE*, 17 (7), e0271260. doi: <https://doi.org/10.1371/journal.pone.0271260>
13. Liu, Z., Cao, W., Gao, Z., Bian, J., Chen, H., Chang, Y., Liu, T.-Y. (2020). Self-paced Ensemble for Highly Imbalanced Massive Data Classification. 2020 IEEE 36th International Conference on Data Engineering (ICDE). doi: <https://doi.org/10.1109/icde48307.2020.00078>
14. Park, S., Lee, H.-W., Im, J. (2022). Resampling-Based Relabeling & Raking Algorithm to One-Class Classification. doi: <https://doi.org/10.36227/techrxiv.17712122.v2>
15. Juez-Gil, M., Arnaiz-González, Á., Rodríguez, J. J., García-Osorio, C. (2021). Experimental evaluation of ensemble classifiers for imbalance in Big Data. *Applied Soft Computing*, 108, 107447. doi: <https://doi.org/10.1016/j.asoc.2021.107447>
16. Wang, Q., Luo, Z., Huang, J., Feng, Y., Liu, Z. (2017). A Novel Ensemble Method for Imbalanced Data Learning: Bagging of Extrapolation-SMOTE SVM. *Computational Intelligence and Neuroscience*, 2017, 1–11. doi: <https://doi.org/10.1155/2017/1827016>
17. Singh, G., Chu, L., Wang, L., Pei, J., Tian, Q., Zhang, Y. (2022). Mining Minority-Class Examples with Uncertainty Estimates. *Lecture Notes in Computer Science*, 258–271. doi: https://doi.org/10.1007/978-3-030-98358-1_21
18. Rayhan, F., Ahmed, S., Mahbub, A., Jani, Md. R., Shatabda, S., Farid, D. Md., Rahman, C. M. (2017). MEBoost: Mixing estimators with boosting for imbalanced data classification. 2017 11th International Conference on Software, Knowledge, Information Management and Applications (SKIMA). doi: <https://doi.org/10.1109/skima.2017.8294128>
19. Meng, Q. (2022). Credit Card Fraud Detection Using Feature Fusion-based Machine Learning Model. *Highlights in Science, Engineering and Technology*, 23, 111–116. doi: <https://doi.org/10.54097/hset.v23i.3208>
20. Zhou, Z.-H. (2020). *Ensemble Learning: Foundations and Algorithms*. Beijing: Electronic Industry Press.
21. Chen, Z., Duan, J., Kang, L., Qiu, G. (2021). A hybrid data-level ensemble to enable learning from highly imbalanced dataset. *Information Sciences*, 554, 157–176. doi: <https://doi.org/10.1016/j.ins.2020.12.023>
22. Bader-El-Den, M., Teitei, E., Perry, T. (2019). Biased Random Forest For Dealing With the Class Imbalance Problem. *IEEE Transactions on Neural Networks and Learning Systems*, 30 (7), 2163–2172. doi: <https://doi.org/10.1109/tnnls.2018.2878400>
23. Borah, P., Gupta, D. (2021). Robust twin bounded support vector machines for outliers and imbalanced data. *Applied Intelligence*, 51 (8), 5314–5343. doi: <https://doi.org/10.1007/s10489-020-01847-5>
24. Gupta, D., Richhariya, B. (2018). Entropy based fuzzy least squares twin support vector machine for class imbalance learning. *Applied Intelligence*, 48 (11), 4212–4231. doi: <https://doi.org/10.1007/s10489-018-1204-4>
25. Husejinovic, A. (2020). Credit card fraud detection using naive Bayesian and C4.5 decision tree classifiers. *Periodicals of Engineering and Natural Sciences*, 8 (1), 1–5.
26. Kalid, S. N., Ng, K.-H., Tong, G.-K., Khor, K.-C. (2020). A Multiple Classifiers System for Anomaly Detection in Credit Card Data With Unbalanced and Overlapped Classes. *IEEE Access*, 8, 28210–28221. doi: <https://doi.org/10.1109/access.2020.2972009>
27. Lee, T.-H., Ullah, A., Wang, R. (2019). Bootstrap Aggregating and Random Forest. *Advanced Studies in Theoretical and Applied Econometrics*, 389–429. doi: https://doi.org/10.1007/978-3-030-31150-6_13
28. Denuit, M., Hainaut, D., Trufin, J. (2020). Bagging Trees and Random Forests. *Effective Statistical Learning Methods for Actuaries II*, 107–130. doi: https://doi.org/10.1007/978-3-030-57556-4_4
29. Genuer, R., Poggi, J.-M. (2020). Random Forests. *Random Forests with R*, 33–55. doi: https://doi.org/10.1007/978-3-030-56485-8_3
30. Syam, N., Kaul, R. (2021). Random Forest, Bagging, and Boosting of Decision Trees. *Machine Learning and Artificial Intelligence in Marketing and Sales*, 139–182. doi: <https://doi.org/10.1108/978-1-80043-880-420211006>
31. Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5–32. doi: <https://doi.org/10.1023/a:1010933404324>
32. Hatwell, J., Gaber, M. M., Azad, R. M. A. (2020). CHIRPS: Explaining random forest classification. *Artificial Intelligence Review*, 53 (8), 5747–5788. doi: <https://doi.org/10.1007/s10462-020-09833-6>
33. Lin, W., Gao, J., Wang, B., Hong, Q. (2021). An Improved Random Forest Classifier for Imbalanced Learning. 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA). doi: <https://doi.org/10.1109/icaica52286.2021.9497933>
34. Sudha, C., Akila, D. (2021). Credit Card Fraud Detection System based on Operational & Transaction features using SVM and Random Forest Classifiers. 2021 2nd International Conference on Computation, Automation and Knowledge Management (ICCAKM). doi: <https://doi.org/10.1109/iccakm50778.2021.9357709>
35. Aung, M. H., Seluka, P. T., Fuata, J. T. R., Tikoisuva, M. J., Cabealawa, M. S., Nand, R. (2020). Random Forest Classifier for Detecting Credit Card Fraud based on Performance

- Metrics. 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE). doi: <https://doi.org/10.1109/csde50874.2020.9411563>
36. Credit Card Fraud Detection. Available at: <https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud>
37. Carrington, A. M., Manuel, D. G., Fieguth, P. W., Ramsay, T., Osmani, V., Wernly, B. et al. (2023). Deep ROC Analysis and AUC as Balanced Average Accuracy, for Improved Classifier Selection, Audit and Explanation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45 (1), 329–341. doi: <https://doi.org/10.1109/tpami.2022.3145392>
38. Kamble, S., Desai, A., Vartak, P. (2014). Evaluation and Performance Analysis of Machine Learning Algorithms. *International Journal of Engineering Sciences & Research Technology*, 3 (5), 789–794.
39. Glisic, S. G., Lorenzo, B. (2022). Machine Learning Algorithms. *Artificial Intelligence and Quantum Computing for Advanced Wireless Networks*, 17–54. doi: <https://doi.org/10.1002/9781119790327.ch2>
40. Piegorsch, W. W. (2020). Confusion Matrix. *Wiley StatsRef: Statistics Reference Online*, 1–4. doi: <https://doi.org/10.1002/9781118445112.stat08244>

АНОТАЦІЇ

MATHEMATICS AND CYBERNETICS – APPLIED ASPECTS

DOI: 10.15587/1729-4061.2023.280355

РОЗРОБКА МЕТОДИКИ ПОШУКУ РІШЕНЬ В СФЕРІ НАЦІОНАЛЬНОЇ БЕЗПЕКИ З ВИКОРИСТАННЯМ БІОІНСПІРОВАНИХ АЛГОРИТМІВ (с. 6–13)**В. В. Коваль, А. В. Шишацький, Р. І. Рансевич, В. Л. Гура, О. Л. Налапко, Л. М. Шипілова, Н. М. Протас, О. Л. Становський, О. В. Волков, О. Є. Чайковська**

Об'єктом дослідження є системи підтримки прийняття рішень. Предметом дослідження є процес прийняття рішення в задачах управління за допомогою біоінспірованих алгоритмів.

Запропоновано методику пошуку рішень в сфері національної безпеки з використанням біоінспірованих алгоритмів. Запропонована методика заснована на поєднанні штучного алгоритму кажанів та штучних нейронних мереж, що еволюціонують. Методика має наступну послідовність дій:

- введення вихідних даних;
- оброблення вихідних даних з урахуванням ступеню невизначеності;
- нумерація агентів-кажанів (АК);
- розставлення агентів-кажанів з урахуванням ступеню невизначеності про стан об'єкту аналізу у просторі пошуку;
- задання початкової швидкості АК та частоти ехолокації кожного АК;
- запуск локального пошуку;
- запуск глобального пошуку;
- навчання баз знань агентів-кажанів.

Оригінальність запропонованої методики полягає у розставленні агентів-кажанів з урахуванням невизначеності вихідних даних, удосконаленими процедурами глобального та локального пошуку з урахуванням ступеню зашумленості даних про стан об'єкту аналізу.

Також особливістю запропонованої методики полягає в використанні удосконаленої процедури навчання агентів-кажанів. Процедура навчання полягає в тому, що відбувається навчання синаптичних ваг штучної нейронної мережі, типу та параметрів функції належності, а також архітектури окремих елементів і архітектури штучної нейронної мережі в цілому. Використання методики дозволяє досягти підвищення оперативності обробки даних на рівні 13–21 % за рахунок використання додаткових удосконалених процедур. Запропоновану методику доцільно використовувати для вирішення задач оцінки складних та динамічних процесів в інтересах вирішення завдань національної безпеки.

Ключові слова: оперативність прийняття рішень, системи підтримки прийняття рішень, складні процеси, національна безпека.

DOI: 10.15587/1729-4061.2023.282130

СТАТИСТИЧНЕ ОПРАЦЮВАННЯ МАЛОЇ ВИБІРКИ ВИХІДНИХ ДАНИХ ІЗ ВИКОРИСТАННЯМ ТЕХНОЛОГІЇ ШТУЧНОЇ ОРТОГОНАЛІЗАЦІЇ (с. 14–21)**Л. Г. Раскін, Л. В. Сухомлин, В. В. Карпенко, Д. Д. Соколов**

Розглянуто завдання розробки методики статистичного оцінювання ситуації, коли обсяг масиву вихідних даних, що використовуються при обробці, недостатній для коректного визначення параметрів функції відгуку. Об'єкт дослідження – технології статистичної обробки малої вибірки даних. Предмет дослідження – методи статистичного оцінювання за умов малої вибірки вихідних даних. Основний напрямок – розробка спеціальної методики статистичної обробки малої вибірки вихідних даних, що забезпечує коректне статистичне оцінювання параметрів функції відгуку. Метод розв'язання задачі – виділення максимально представницького ортогонального реплікоподібного підплану із плану повного факторного експерименту, отриманого шляхом штучної ортогоналізації результатів пасивного експерименту. Необхідність та доцільність запропонованої процедури є наслідком непередбачуваності та нерівномірності розподілу точок у фазовому просторі координат. Результатом реалізації відповідної методики є зрізаний ортогональний план повного факторного експерименту, що забезпечує можливість незалежного оцінювання всіх коефіцієнтів регресійного полінома, що описує функцію відгуку. У разі жорсткого дефіциту кількості вимірів методика дозволяє виділити з отриманого плану повного факторного експерименту представницьку ортогональну репліку. Використання цього плану повного факторного експерименту дозволяє здійснити оцінку всіх коефіцієнтів регресійного полінома, що описує потрібну функцію відгуку. Відповідна обчислювальна процедура заснована на розв'язанні триаксальної булевої задачі призначення.

Ключові слова: статистична обробка даних, мала вибірка, штучна ортогоналізація, триаксальне завдання призначення.

DOI: 10.15587/1729-4061.2023.282645

ВИЯВЛЕННЯ ДОГОВІРНИХ МАТЧІВ НА ОСНОВІ ТЕОРІЇ КОНФОРМНИХ ПРЕДИКТОРІВ З ВИКОРИСТАННЯМ МОДИФІКОВАНОЇ ФУНКЦІЇ-ІНДИКАТОРУ СТЕПАНЦЯ (с. 22–32)**О. Р. Чертов, І. С. Жук**

Актуальною проблемою сучасних футбольних змагань є виявлення договірних матчів. Відомі методи прогнозування результату матчу за рахунок аналізу ставок на матч чи аналізу дій футболістів на полі використовують велику кількість даних, які не завжди є доступними. Для подолання цієї перешкоди може бути застосовано метод виявлення підозрілих на фіксований результат матчів на основі конформних предикторів і степеневих мартингалів, який використовує загальнодоступні публічні дані. Але на практиці цей метод не завжди виявляє такі матчі з високою точністю. Запропоновано вдосконалений метод визначення підозрілого матчу на основі теорії конформних предикторів з використанням модифікованої функції-індикатора Степанця, яка порівнюється з порогом. Модифікована функція-індикатор Степанця застосовується до степеневого мартингалу і показує відносну зміну значення мартингалу

поточного матчу у порівнянні з попереднім матчем. Значення порогу визначено експериментально за критерієм максимуму метрики $F1$. Як навчальна вибірка бралися дані сезону 2013–2014 років Ліги П Франції, а як тестова вибірка – дані сезону 2014–2015 років Серії В Італії. На всіх вибірках було проведено кластеризацію команд. Для кожного з утворених класів матчів на обох вибірках було обчислено міру неконформності, ступінь неконформності, степеневий мартингал і модифіковану функцію-індикатор Степанця. Отримані показники метрик точності і $F1$ є вищими (середні значення метрик $P=0,84$, $F1=0,87$), ніж ці ж показники у правил мартингала і p -value (середні значення метрик $P=0,75$, $F1=0,78$), застосованих до цих же даних. Запропонованим методом виявляються 4 з 5 матчів сезону 2014–2015 Серії В Італії, які вважаються договірними за інформацією з офіційних правоохоронних джерел Італії.

Ключові слова: фіксований результат, степеневий мартингал, міра неконформності, офлайн алгоритм, p -value, метрика $F1$.

DOI: 10.15587/1729-4061.2023.279622

РОЗРОБКА НОВОГО БАГАТОКРИТЕРІАЛЬНОГО МЕТОДУ ПРИЙНЯТТЯ РІШЕНЬ (с. 33–38)

Tran Van Dua

Вибір найкращого варіанту з багатьох доступних варіантів завжди є метою, якої потрібно досягти в усіх сферах. Однак параметри (критерії) в кожній альтернативі неоднакові, іноді суперечливі. У цій ситуації вибір найкращого варіанту є надзвичайно важким рішенням для особи, яка приймає рішення. Багатокритеріальне прийняття рішень (БПР) – це ранжування альтернатив на основі критеріїв кожної альтернативи. Винахідниками запропоновано понад сто багатокритеріальних методів прийняття рішень. Вони використовуються в багатьох різних сферах. Однак для тих, хто приймає рішення, вибір відповідного методу для кожного конкретного випадку є складним завданням. CURLI (Collaborative Unbiased Rank List Integration) – багатокритеріальний метод прийняття рішень, який відрізняє його від усіх інших. Ця різниця відображається в тому факті, що при застосуванні цього методу особі, яка приймає рішення, не потрібно ні нормалізувати дані, ні визначати ваги для критеріїв. Однак особам, які приймають рішення, знадобиться багато часу, щоб застосувати цей метод, особливо коли кількість варіантів для ранжирування велика. Це дослідження здійснило розробку нового методу MCDM на основі методу CURLI. Цей новий метод отримав назву CURLI-2. Наведено багато різних прикладів для оцінки ефективності запропонованого методу. У кожному прикладі результат ранжування альтернатив за допомогою методу CURLI-2 порівнювався з тими, які використовували інші різні методи БПР. Найкраща альтернатива, визначена при використанні методу CURLI-2, завжди збігається з використанням існуючих методів БПР. Використання методу CURLI-2 для ранжирування альтернатив буде набагато швидшим і простішим, ніж використання методу CURLI. Це перевага методу CURLI-2 порівняно з методом CURLI.

Ключові слова: новий багатокритеріальний метод прийняття рішень, метод CURLI-2, метод CURLI.

DOI: 10.15587/1729-4061.2023.281007

РОЗРОБКА МЕТОДИКИ ПОШУКУ РІШЕНЬ З ВИКОРИСТАННЯМ ПОПУЛЯЦІЙНОГО АЛГОРИТМУ ГЛОБАЛЬНОЇ ПОШУКОВОЇ ОПТИМІЗАЦІЇ (с. 39–46)

С. В. Яким'як, Є. А. Вдовицький, Ю. З. Артабаєв, Л. М. Дегтярєва, Ю. В. Вакуленко, С. С. Невгад, В. В. Андронов, Р. Р. Лазута, П. І. Шаповал, Артамонов Є. Б.

Об'єктом дослідження є системи підтримки прийняття рішень. Предметом дослідження є процес прийняття рішення в задачах управління за допомогою біоінспірованих алгоритмів. Запропоновано методику пошуку рішень з використанням популяційного алгоритму глобальної пошукової оптимізації. Запропоноване спільне використання алгоритму інвазивних бур'янів, удосконаленого генетичного алгоритму та штучних нейронних мереж, що еволюціонують. Методика має наступну послідовність дій:

- введення вихідних даних;
- оброблення вихідних даних з урахуванням ступеню невизначеності;
- формування вектору оптимізації;
- створення векторів-нащадків;
- впорядкування векторів по убуванню;
- скорочення розмірності простору ознак;
- навчання баз знань.

Особливості запропонованої методики полягає у розставленні агентів- бур'янів з урахуванням невизначеності вихідних даних, удосконаленими процедурами зменшення простору ознак про стан об'єкту аналізу.

Здійснюється навчання синаптичних ваг штучної нейронної мережі, типу та параметрів функції належності, а також архітектури окремих елементів і архітектури штучної нейронної мережі в цілому. Проведено моделювання роботи запропонованої методики в програмному середовищі MathCad 14. В якості задачі, що вирішувалася при проведенні моделювання був визначення маршруту руху суден в операційних зонах Чорного та Азовського морів в умовах гібридних дій противника. Використання методики дозволяє досягти підвищення оперативності обробки даних на рівні 21–27 % за рахунок використання додаткових удосконалених процедур. Запропоновану методику доцільно використовувати для вирішення задач оцінки складних та динамічних процесів в інтересах вирішення завдань національної безпеки.

Ключові слова: гібридні дії, штучні нейронні мережі, біоінспіровані алгоритми, застосування сил (військ) Військово-Морських Сил.

DOI: 10.15587/1729-4061.2023.279936

МОДЕЛЮВАННЯ КРЕДИТНОГО ШАХРАЙСТВА ПРИ НЕЗБАЛАНСОВАНИХ ДАНИХ НА ОСНОВІ БЕГІНГУ ТА БАЙЄСІВСЬКОЇ ОПТИМІЗАЦІЇ (с. 47–53)

Mohammed A. Kashmoola, Samah Fakhri Aziz, Hasan Mudhafar Qays, Naors Y. Anad Alsaleem

Моделювання кредитного шахрайства є важливою областю досліджень, що має велике значення для сфери кредитування. Ефективне управління ризиками є ключовим фактором у наданні якісних кредитних послуг і безпосередньо впливає на рентабельність та

коефіцієнт безнадійної заборгованості провідних організацій у даному секторі. Однак при сильно незбалансованому розподілі даних про кредитне шахрайство можуть виникати шумові похибки, викликані спотворенням інформації, періодичними статистичними помилками та зміщеннями моделі під час навчання. Це може призвести до неправильних результатів для класу меншості (цільового класу) та збільшити ризик перенавчання. У той час як традиційні методи балансування даних дозволяють зменшити зміщення моделей у бік класу більшості при відносно незбалансованих даних, вони можуть виявитися неефективними у випадках сильно незбалансованих. Для вирішення цієї проблеми в роботі пропонується використовувати алгоритми бегінгу, такі як випадковий ліс та бегінг, для моделювання сильно незбалансованих даних про кредитне шахрайство. Байєсівська оптимізація використовується для пошуку гіперпараметрів та визначення точності класу меншості як функції оптимізації для моделі, що тестується на реальних даних про шахрайство з кредитними картками в Європі. Результати запропонованих алгоритмів пакетування порівнюються з традиційними методами балансування даних, такими як збалансований бегінг та збалансований випадковий ліс. Дослідження показало, що традиційні методи балансування даних можуть бути несумісними з надмірно незбалансованими даними, тоді як алгоритми бегінгу є перспективним рішенням для моделювання таких даних. Запропонований метод знаходження гіперпараметрів ефективний при сильно незбалансованих даних. Він дозволив досягти точності, повноти та F1-міри для категорії меншості 0,94, 0,81 і 0,87 відповідно. Дослідження підкреслює важливість вирішення задач, пов'язаних з незбалансованими даними про кредитне шахрайство, для підвищення точності та об'єктивності моделей кредитного шахрайства.

Ключові слова: незбалансовані дані, байєсівська оптимізація, випадковий ліс, класи більшості та меншості.