INFUS
International Conference on Intelligent and Fuzzy Systems

INFUS CANAKKALE

An intuitionistic fuzzy approach for blood data analysis

Evdokia Sotirova[1,2], Veselina Bureva[1], Sotir Sotirov[1,3], Hristo Bozov[4,5], Simeon Ribagin[1,6]

[1] Laboratory of Intelligent Systems, Prof. Asen Zlatarov University, Burgas, Bulgaria,
[2] University hospital for active treatment, Burgas, Bulgaria.
[3] Southeast Digital Innovation Hub, 2 Serdika str. TRIA City Center, floor 2
[4] Medical Faculty, Prof. Asen Zlatarov University, Burgas-8010, Bulgaria
[5] Oncology Complex Center - Burgas, 86 Demokratsiya Blvd, Burgas 8000 Bulgaria
[6] Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences

Email address: esotorova@btu.bg, vbureva@btu.bg, ssotirov@btu.bg, hr_bozov@yahoo.com, simribagin@gmail.com

1. Introduction

The transfusion of quality blood and blood components is of utmost importance. On the one hand, the high quality of these products is essential for the correct treatment of patients, and on the other hand, ensuring high quality is a complex task.

Actions are regularly organized to encourage and popularize regular, anonymous and voluntary blood donation. It is of particular importance to improve the quality of the produced blood components. In this sense, it would be useful to be able to identify people who are potentially suitable as blood donors. In this way, it can be determined how many of the people in a given region are potential blood donors. One of the diseases in which there are contraindications for donating blood is anemia.

1. Introduction

According to World Health Organization data, anemia affects half a billion women between 15 and 49 years of age and 269 million children between 6 and 59 months of age worldwide. Anemia is a condition in which the levels of red blood cells or hemoglobin are below normal. These cells are responsible for transporting oxygen to all tissues in the body. One of the main functions of donating blood is to provide the necessary elements to maintain the health of the blood system. This determines the impossibility of these people to be blood donors.

1. Introduction

In the present investigation, blood test information was analyzed to determine whether a person is suitable as a blood donor by examining the relationships between blood parameters to detect the presence of anemia.

The studied data are taken from a public domain with a free license. The datasets contain measurements (parameters) from complete blood count test performed by Hematology analyzer: hemoglobin, red blood cell count, packed cell volume test, mean corpuscular volume, mean corpuscular hemoglobin, mean corpuscular hemoglobin concentration, red cell distribution width, total leucocyte count, platelet/mm^3 (PLT/mm^3), etc.

1. Introduction

The input data set is analyzed by two intelligent tools - an InterCriteria Analysis (ICA) method and Deep learning neural network.

The ICA method is based on two mathematical tools - indexed matrices and intuitionistic fuzzy sets. ICA is an intelligent method by which the correlations between all the criteria by which objects are evaluated can be presented simultaneously. In some situations, some of these criteria are more expensive and more difficult to assess or measure. If they are highly dependent on some of the other measured (evaluated) criteria, it is possible to initially remove the first ones.

1. Introduction

Free licensed ICA software at http://intercriteria.net/software was used for programmatic testing of the data. After applying the ICA method, dependencies between each pair of the measured (evaluated) parameters of the blood donors are obtained. The resulting correlations are in the form of intuitionistic fuzzy pairs (IFPs). They have values between 0 and 1. An ICA approach was used to detect dependencies between parameters related to blood analysis and blood donation.

Using these measurements, a neural network was trained to diagnose anemia in potential blood donors. Applying DNNs to analyze blood test data can lead to new info into the relationships between various biomarkers and health outcomes.

1. Introduction

By identifying patterns and associations that may not be apparent to human observers, DNNs can contribute to the discovery of novel biomarkers, risk factors, and disease mechanisms, advancing our understanding of human physiology and pathology. The use of DNNs for anemia diagnosis can improve diagnostic accuracy compared to traditional methods. By leveraging large datasets and complex modeling techniques, DNNs can identify subtle patterns and features in blood test data that may be indicative of anemia, leading to earlier and more accurate diagnoses. This improved accuracy can reduce misdiagnosis rates and improve patient outcomes.

1. Introduction

DNNs can enhance existing predictive models for anemia and other hematological conditions by incorporating a broader range of input features and learning more complex relationships between variables.

DNNs enable personalized predictions and recommendations based on individual patient data, allowing healthcare providers to tailor interventions and treatments to each patient's unique characteristics and risk factors.

By analyzing large-scale datasets with DNNs, researchers can generate data-driven insights that inform future studies and contribute to the development of more effective diagnostic and therapeutic strategies for anemia and related conditions.

1. Introduction

The purpose of the investigation is to reduce the input data about blood donors at the inputs of a DNN. The reduction of the DNN inputs is done on the basis of the results obtained from the ICA. When two parameters are highly correlated in between, one can be reduced without significant loss of accuracy.

The organization of the article is as follows: Section 1 is introductory. Section 2 presents the application of ICA approach. Section 3 presents the neural network training with reduced input data. In Section 4 are the conclusions of the conducted research. Section 5 is acknowledgments. Section 6 is references.

2 Application of the ICA method

The dataset was organized in index matrix with 11 rows (for the mean square error of the parameters for evaluation the breast cancer) and 362 columns (for each blood donor).

The evaluating parameters (criteria) are as follows:

- C1: age,
- C2: gender,
- C3: hemoglobin (HGB),
- C4: red blood cell count (RBC),
- C5: packed cell volume test (PCV),
- C6: mean corpuscular volume (MCV),
- C7: mean corpuscular hemoglobin (MCH),
- C8: mean corpuscular hemoglobin concentration (MCHC),
- C9: red cell distribution width (RDW),
- C10: total leucocyte count,
- C11: platelet/mm^3 (PLT/mm^3)

2 Application of the ICA method

After applying the ICA method, 55 intuitionistic fuzzy pairs IFPs were obtained, reflecting the degree of correlation between the analyzed parameters. Indexed matrices with membership part and non-membership part of the IFPs are given in Table 1 and Table 2

Table 1. Result matrix with membership parts of the IFPs of the correlations between parameters for evaluation the blood donor.

μ	Age	Sex	RBC	PCV	MCV	MCH	MCHC	RDW	TLC	PLT/ mm^3	HGB
Age	1,00	0,22	0,45	0,47	0,52	0,51	0,49	0,48	0,58	0,48	0,47
Sex	0,22	1,00	0,17	0,16	0,21	0,18	0,19	0,31	0,25	0,28	0,13
RBC	0,45	0,17	1,00	0,85	0,40	0,36	0,44	0,44	0,53	0,51	0,76
PCV	0,47	0,16	0,79	1,00	0,56	0,48	0,44	0,39	0,54	0,50	0,85
MCV	0,52	0,21	0,40	0,56	1,00	0,75	0,51	0,46	0,48	0,45	0,55
MCH	0,51	0,18	0,36	0,48	0,75	1,00	0,73	0,39	0,43	0,43	0,57
MCHC	0,49	0,19	0,44	0,44	0,51	0,73	1,00	0,37	0,42	0,42	0,58
RDW	0,48	0,31	0,44	0,39	0,46	0,39	0,37	1,00	0,50	0,49	0,34
TLC	0,58	0,25	0,53	0,54	0,48	0,43	0,42	0,50	1,00	0,61	0,50
PLT/ mm^3	0,48	0,28	0,51	0,50	0,45	0,43	0,42	0,49	0,61	1,00	0,46
HGB	0,47	0,13	0,76	0,82	0,55	0,57	0,58	0,34	0,50	0,46	1,00

2 Application of the ICA method

Table 2. Result matrix with non-membership parts of the IFPs of the correlations between parameters for evaluation the blood donor.

ν	Age	Sex	RBC	PCV	MCV	MCH	MCHC	RDW	TLC	PLT/ mm^3	HGB
Age	0,00	0,27	0,53	0,51	0,46	0,47	0,48	0,48	0,40	0,50	0,50
Sex	0,27	0,00	0,32	0,33	0,29	0,31	0,30	0,18	0,25	0,21	0,36
RBC	0,53	0,32	0,00	0,15	0,59	0,62	0,54	0,54	0,46	0,48	0,22
PCV	0,51	0,33	0,20	0,00	0,43	0,51	0,54	0,59	0,46	0,49	0,14
MCV	0,46	0,29	0,59	0,43	0,00	0,24	0,47	0,52	0,51	0,54	0,43
MCH	0,47	0,31	0,62	0,51	0,24	0,00	0,25	0,59	0,55	0,56	0,40
MCHC	0,48	0,30	0,54	0,54	0,47	0,25	0,00	0,61	0,56	0,56	0,39
RDW	0,48	0,18	0,54	0,59	0,52	0,59	0,61	0,00	0,49	0,49	0,63
TLC	0,40	0,25	0,46	0,46	0,51	0,55	0,56	0,49	0,00	0,38	0,49
PLT/ mm^3	0,50	0,21	0,48	0,49	0,54	0,56	0,56	0,49	0,38	0,00	0,52
HGB	0,50	0,36	0,22	0,16	0,43	0,40	0,39	0,63	0,49	0,52	0,00

2 pairs of criteria are in positive consonance: PCV-HGB : ⟨0.853; 0.139⟩; RBC - PCV: ⟨0.851; 0.148⟩. 1 pair is in RBC - HGB : ⟨0.763; 0.219⟩. This means that these criteria are in very similar behavior.

5 pairs are in weak dissonance, 13 pairs are criteria in dissonance, 27 pairs are in strong dissonance and 7 pairs are in weak negative consonance. That means, that these pairs are independent from each other.

3 Application of the feed-forward neural network

We use a feed-forward neural network to detect anemia. The structure we use to detect anemia is 8 inputs, 24 neurons in the first neural layer, 12 neurons in the second neural layer, 8 neurons in the third neural layer, 5 neurons in the fourth neural layer, 3 neurons in the fifth neural layer, and 1 neuron in the last sixth neural layer. The transfer functions of all layers are logsig, and the last neural layer has a linear function.

The following parameters are given to the input of the neural network: age, gender, hemoglobin (HGB), mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), red cell distribution width (RDW), total leucocyte count and platelet/mm3 (PLT/mm3), and Presence of anemia (yes/no) are given to the output.

3 Application of the feed-forward neural network

Levenberg-Marquardt back-propagation algorithm is used for training the neural network, which is one of the fastest for training the neural network. Data from 1421 measurements are used to train the neural network, which are divided into three parts. The first part is for neural network training - 75%, the second part is for testing -13% and for verification 12%.

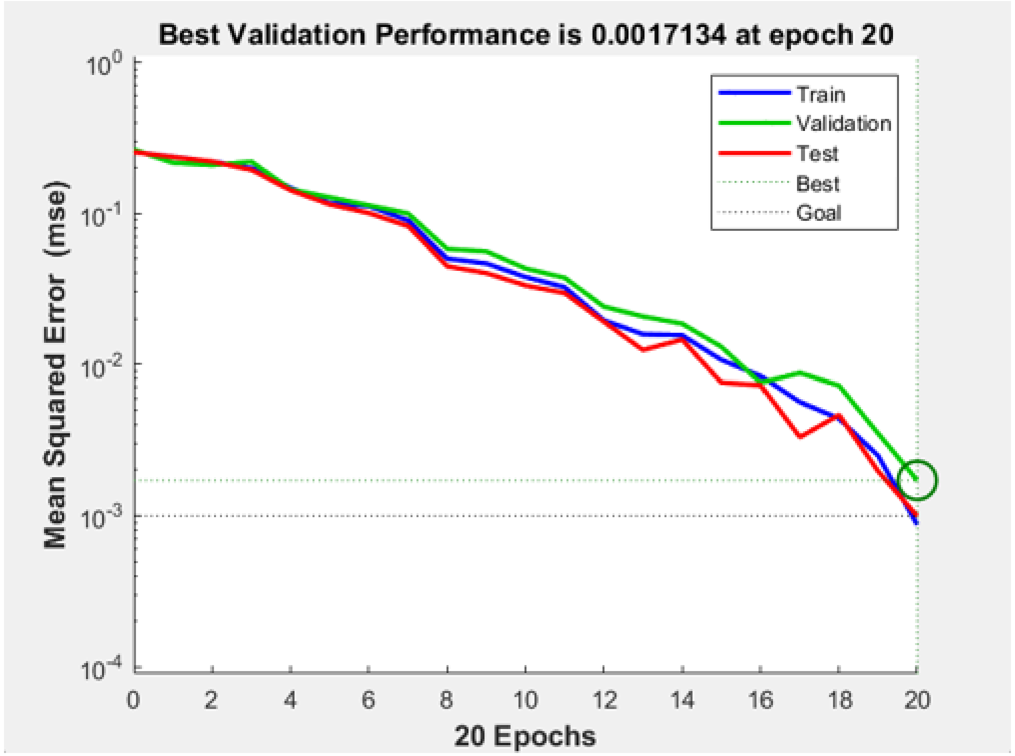

Fig. 1. Learning process of the neural network.

3 Application of the feed-forward neural network

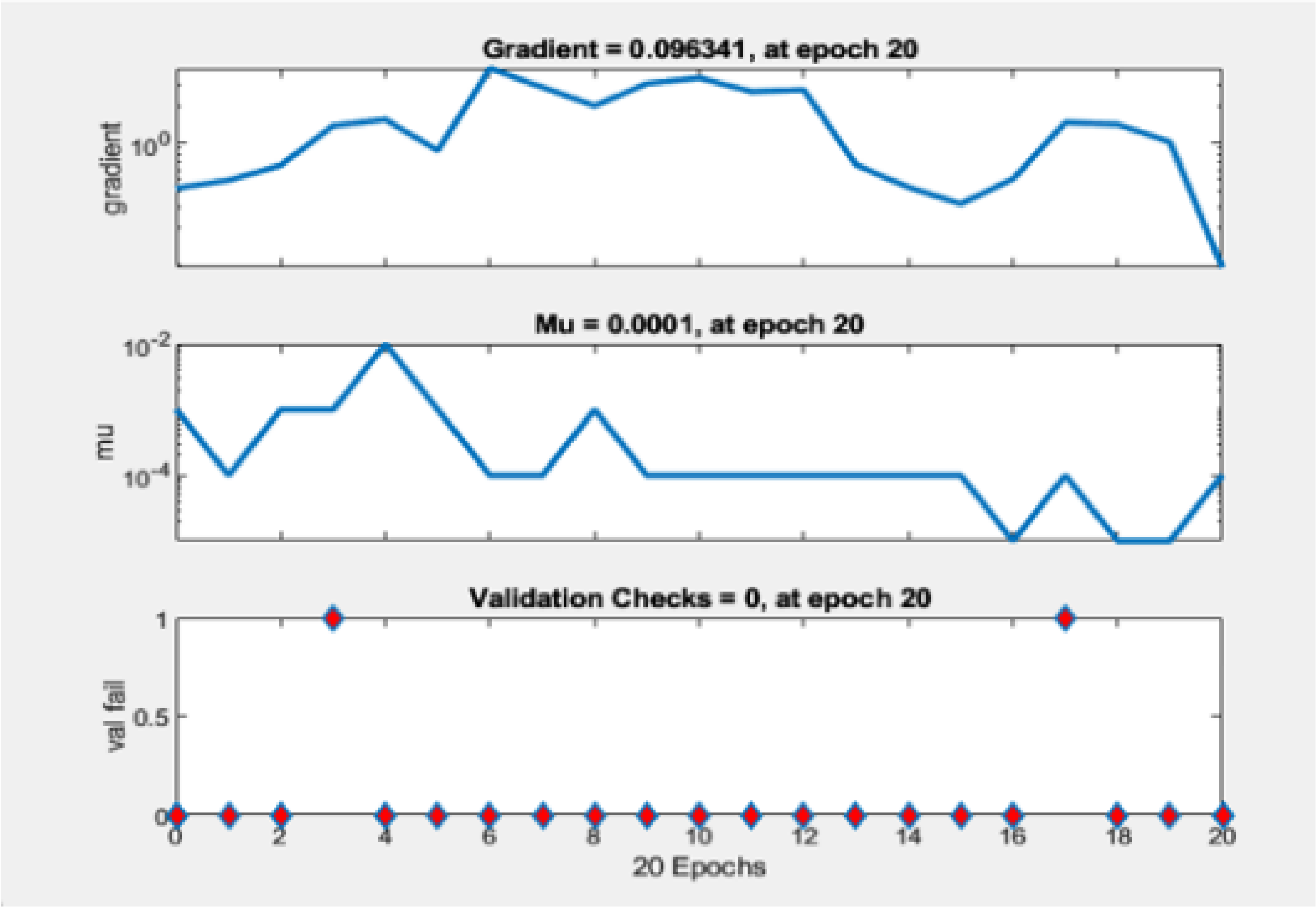

Fig. 2. Training parameters of the neural network.

3 Application of the feed-forward neural network

Most important in training and predicting data from Datasets is the regression coefficient: for the test data, the training data, the verification data, and all the training data.

The training of the deep neural network is for 20 epochs, having a root mean square error of 0.0017134, and the regression coefficients are shown at Fig. 3.

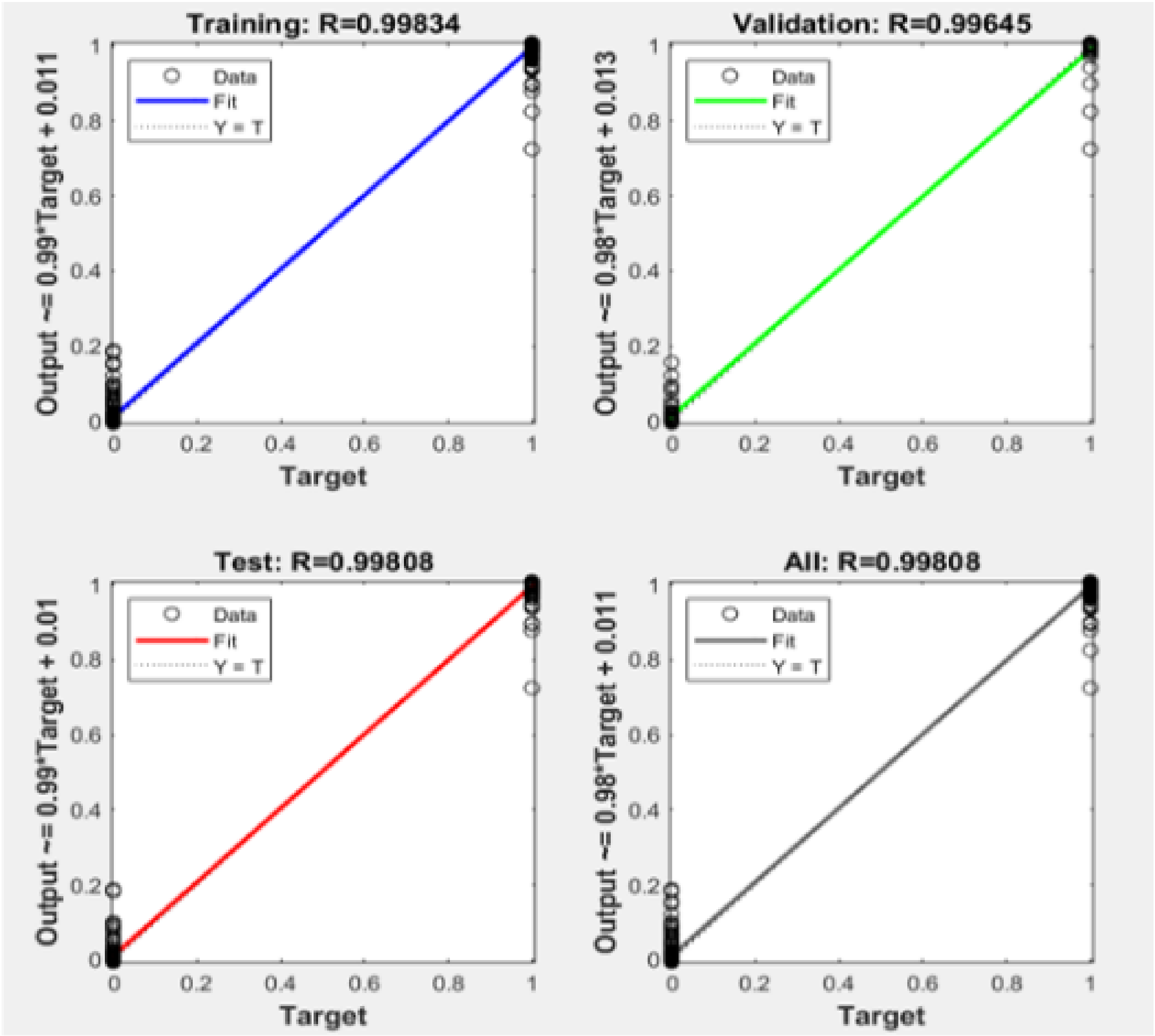

Fig. 3. The regression coefficients of the neural network - for the test data, the training data, the verification data, and all the training data.

4 Conclusion

In conclusion, the transfusion of high-quality blood and blood components is indispensable for effective patient treatment, emphasizing the importance of encouraging regular, anonymous, and voluntary blood donations. To enhance the quality of produced blood components, it is crucial to identify individuals who are potentially suitable as blood donors. This study investigates the relationships between various blood parameters to determine donor suitability, employing the (ICA) approach, which integrates indexed matrices and intuitionistic fuzzy logic.

4 Conclusion

Through the analysis of data from 362 patients, including age, gender, and various hematological parameters, the ICA method yielded intuitionistic fuzzy pairs, indicating the degree of correlation between these parameters. These findings enable the assessment of donor suitability and the identification of individuals at risk of developing conditions such as anemia. The application of the ICA method provides valua-ble insights into the complex interrelationships among blood parameters, aiding in the identification of potential blood donors and facilitating the production of high-quality blood components.

4 Conclusion

By using a data-driven deep neural network, healthcare professionals can make informed decisions about blood donor eligibility, ultimately ensuring the delivery of safe and effective blood transfusion therapies to patients in need. The collective's future research is aimed at proposing new methods and architectures to improve informed decisions by physicians.

Acknowledgment

This research was funded by the Bulgarian National Science Fund under the Grant KP-06-N72/8 from 14.12.2023 "Intuitionistic fuzzy methods for data analysis with an emphasis on the blood donation system in Bulgaria".

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

1. Acker, J. P., Marks, D. C., Sheffield, W. P. Quality assessment of established and emerging blood components for transfusion. Journal of blood transfusion, Vol. 2016, Article ID 4860284, 28 pages, http://dx.doi.org/10.1155/2016/4860284 (2016).
2. Andreev, N., Atanassov, K., Bureva, V.: InterCriteria analysis on data for blood collec-tion. Annual of the Informatics Section, Union of Scientists in Bulgaria, 10, pp. 30–53 (2019–2020). (in Bulgarian)
3. Andreev, N., Atanassova, V.: InterCriteria analysis of the blood group distribution of patients of Saint Anna hospital in 2015–2019. In: Atanassov, K.T., et al. (eds.) Ad-vances and New Developments in Fuzzy Logic and Technology (IWIFSGN 2019). AISC, vol. 1308, pp. 158–165. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-77716-6_14
4. Andreev, N., Sotirova, E., Ribagin, S. Intercriteria analysis of data from the centers for transfusion haematology in Bulgaria. Comptes rendus de l'Acade'mie bulgare des Sci, 72(7), 982-990 (2019).
5. Andreev, N., Vassilev, P., Ribagin, S., Sotirov, S.: InterCriteria analysis of data for blood collection in the transfusion hematology department, university hospital "St. Anna", Sofia. Notes Intuitionistic Fuzzy Sets 25(2), 88–95 (2019)
6. Anemia Diagnosis Dataset, https://www.kaggle.com/datasets/saurabhshahane/anemia-diagnosis-dataset?resource=download
7. Atanassov, K. Index Matrices Towards an Augmented Matrix Calculus. Studies in Computational Intelligence Series, Vol. 573, Springer, Cham (2014).
8. Atanassov, K. On Intuitionistic Fuzzy Sets Theory, Springer, Berlin (2012).
9. Atanassov, K., Mavrov, D., Atanassova, V. InterCriteria Decision Making: A New Ap-proach for Multicriteria Decision Making, Based on Index Matrices and Intuitionistic Fuzzy Sets. Issues in Intuitionistic Fuzzy Sets and Generalized Nets, 11, p. 1–8 (2014).
10. Atanassov, K., Szmidt, E., Kacprzyk, J. On intuitionistic fuzzy pairs, Notes on Intui-tionistic Fuzzy Sets, 19(3), p. 1–13 (2013).

References

11. Atanassova, V., Andreev, N., Dimitriev, A. InterCriteria Analysis of the Geographic Distribution of the ABO System Blood Groups in the Patients of the University Hospi-tal “Saint Anna”, Sofia, Bulgaria, from 2015 to 2021. BioInfoMed 2022. Lecture Notes in Networks and Systems, vol 658., pp. 84–97. Springer, Cham. (2023).
12. Atanassova, V., Andreev, N., Dimitriev, A. ABO system blood groups distribution in Bulgaria, based on a dataset of the patients of the university hospital “Saint Anna”, So-fia, Bulgaria, from 2015 to 2021. In The International Symposium on Bioinformatics and Biomedicine (pp. 72-83). Cham: Springer Nature Switzerland. (2022).
13. Cappellini, M. D., Motta, I. (Anemia in clinical practice—definition and classification: does hemoglobin change with aging?. In Seminars in hematology, Vol. 52, No. 4, pp. 261-269. WB Saunders (2015).
14. Exploring the Anemia dataset, https://www.kaggle.com/code/richel145/exploring-the-anemia-dataset
15. Hess, J. R., Sparrow, R. L., Van Der Meer, P. F., Acker, J. P., Cardigan, R. A., & Devine, D. V. (2009). Blood components: red blood cell hemolysis during blood bank storage: using national quality management data to answer basic scientific questions. Transfusion, 49(12), 2599-2603.
16. InterCriteria Analysis Portal, Software, https://intercriteria.net/software/
17. Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J., Müller, K. R. Explaining deep neural networks and beyond: A review of methods and applications. Proceedings of the IEEE, 109(3), pp. 247-278 (2021).
18. Sotirov, S., Atanassova, V., Sotirova, E., Bureva, V., Mavrov, D. Application of the in-tuitionistic fuzzy InterCriteria analysis method to a neural network preprocessing pro-cedure. In: 9th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT), 30 June–03 July 2015, Gijon, Spain, pp. 1559–1564 (2015).
19. Sotirov, S., Atanassova, V., Sotirova, E., Doukovska, L., Bureva, V., Mavrov, D., & Tomov, J. (2017). Application of the intuitionistic fuzzy InterCriteria analysis method with triples to a neural network preprocessing procedure. Computational Intelligence and Neuroscience, Vol. 2017. Article ID 2157852, 9 pages (2017).
20. World Health Organization, Anaemia, https://www.who.int/news-room/fact-sheets/detail/anaemia, 1 May 2023

Thank you for your attention!