UDC 004.623

ARTIFICIAL INTELLIGENCE IN THE DIAGNOSIS, PROGNOSIS AND TREATMENT OF DIABETIC NEOVASCULAR GLAUCOMA

Dr.Sci. V. Vychuzhanin ORCID: 0000-0002-6302-1832
Odessa Polytechnic National University, Ukraine
E-mail: vychuzhanin.o.v@op.edu.ua
Ph.D. O. Guzun ORCID: 0009-0003-6873-8503
1SI "The Filatov Institute of Eye Diseases and Tissue Therapy of the National Academy of Medical Sciences of Ukraine", Ukraine
E-mail: olga.v.guzun@gmail.com
Ph.D. M. Rudnichenko ORCID: 0000-0002-7343-8076
Odessa Polytechnic National University, Ukraine
E-mail: nickolay.rud@gmail.com
Ph.D. O. Zadorozhnyy ORCID: 0000-0003-0125-2456
1SI "The Filatov Institute of Eye Diseases and Tissue Therapy of the National Academy of Medical Sciences of Ukraine", Ukraine
E-mail: zadoroleg2@gmail.com
Ph.D. A.Korol ORCID: 0000-0003-0516-308X
1SI "The Filatov Institute of Eye Diseases and Tissue Therapy of the National Academy of Medical Sciences of Ukraine", Ukraine
E-mail: andrii.r.korol@gmail.com

Abstract. *This work is focused on key aspects of the diagnosis, prognosis and treatment of neovascular glaucoma of diabetic origin based on machine learning approaches and, in particular, various architectures artificial neural models. An analysis of the relevance, priority provisions and advantages of using machine learning methods is carried out, the existing approaches used in modern literature in the context of the topic under study are considered, the specifics of their integration into the process of diagnostic analysis of the feature space of an aggregated and labeled by the authors data set on patients with visual problems are described, in particular, those suffering from neovascular glaucoma of diabetic origin. A correlation analysis of input features was carried out, 3 different models of artificial neural networks were built, trained and tested, metrics for assessing the accuracy of their work were experimentally calculated and studied, and statistical indicators were determined, including errors and losses, characterizing their generalizing ability. Analysis of the results obtained from the studies made it possible to identify the prevailing input features and evaluate their impact on the target output variable and the overall significance in the feature space of the data set, as well as to establish the most suitable models for data analysis in terms of their accuracy and speed. The conducted research made it possible to establish the fact of a greater degree deep learning artificial neural networks models fully connected adaptability for the analyzed data set.*

Keywords. *Artificial intelligence, neovascular glaucoma, diagnosis eye treatment, neural networks, data analysis, data mining, machine learning*

1. Introduction

Currently, the volumes of heterogeneous information collected during the examination of patients in various fields of medicine are constantly increasing [1]. This is due to the. tendency to integrate various technical means and software systems for diagnosing human body organs diseases, especially vision organs, which make it possible to measure and process numerous quantitative and qualitative parameters characterizing the nature, stage and disease complexity. The development of artificial intelligence (AI) technology in medical decision making allows for the simultaneous analysis of an infinite variables number. AI technologies provide improving patient care in medical decision making and require physicians and computer scientists to have a successful mutual understanding of both technology and medical practice. Recently, thanks to AI methods, significant advances have been made in the diagnosis of ophthalmic pathologies.

With their help, data from visual fields, optic nerve structure (for example, obtained using optical coherence tomography (OCT), biomechanical ocular properties and their combination are successfully analyzed and classified to determine the severity of the disease, its progression and/or referral for specialized treatment. Diagnostic search algorithms for patients with glaucoma have undergone further changes, becoming more and more complex, and to improve existing methodologies it is necessary to use more reliable classification diseases signs.

A significant contribution in this direction is made by modern methods and models for assessing and predicting the effectiveness of patient treatment, actively developed and implemented by specialists in the field of prevention and eye pathologies diagnosis. It should be noted that in the process of inflammation assessing indicators and intraocular circulation in patients collected during statistical data examination, an important aspect of the procedure for their analysis is indicators diagnostic significance [2]. This is relevant both for identifying and formalizing key factors that have the greatest impact on the visual organs integral state, and for developing and implementing preventive measures and developing protocols ophthalmic diseases effective treatment.

Due to the lack of clearly identified patterns and correlations in the feature space of aggregated diagnostic data for the patients' vision organs as well as due to the high degree of uncertainty in mutual influence, carrying out the processes of assessing and analyzing data in a manual format is difficult and labor-intensive. Therefore, it is rational to use modern methods of data mining, in particular machine learning (ML) algorithms, in order to automate the process of searching for hidden patterns in data and assessing the level of feature space diagnostic significance, which is especially justified in the presence of large data volumes [3]. This allows us to provide solutions to various applied problems in the field of ophthalmology. In particular, it helps to increase the efficiency of the decision support process for assessing the patient's condition and the use of targeted treatments for diabetic neovascular glaucoma [4] through the use of ML models to automate data research and search for individual signs correlations with each other. There are significant results from the use of ML in the analysis of such pathological eye conditions as diabetic retinopathy, age-related macular degeneration, macular edema, glaucoma, cataracts, etc. In this case, color fundus photographs, images obtained during fluorescein angiography and autofluorescence, OCT scans, fields of view [5, 6, 7].

The defining characteristic of ML algorithms is the quality of their predictions, which improves with experience. The more data provided, the better the forecasting model. In this context, difficulties arise due to various factors, including [8]:

- insufficiently high generalization ability of classical ML models;
- absence of memory effect in models;
- insufficient adaptability of models to data, taking into account the specifics under consideration;
- the need to compile ML models ensembles to ensure a higher degree of accuracy of the forecast values they generate;
- time costs for preliminary data scaling and preprocessing.

In this regard, a perspective way is to use more advanced, modern and efficient deep ML models, which are artificial neural networks (ANN) of various architectures, topologies and hierarchically interconnected. An important aspect in the problem we are considering is the study of different ML models among themselves to experimentally identify an estimate of the most appropriate hyperparameters values, which affect the output metrics for assessing used models quality [9]. In this context, an approach based on deep learning and ANN makes it possible to more effectively solve classification and regression problems, taking into account the specifics of the application area under consideration, as well as to form various hybrid models that can have greater generalization ability and be more flexible in comparison with classical ML approaches, eliminate the imbalance of the feature space, perform the functions of data normalization and reduce the likelihood of overtraining models through the introduction of regularization methods [1].

An additional important aspect is the ability to minimize the risk of generating contradictory results from ANN models, for example, by integrating bagging and boosting techniques when creating model's ensembles. In this regard, ANN using is more appropriate within the framework of the topic under consideration.

2. Analysis existing researches

In the medical field, there is an active use of various approaches, most often based on statistical methods and ML models for the analysis and evaluation of experimentally obtained results of diagnosis and various diseases treatment, including vision organs.

Thus, the authors [10] adapt ML learning algorithms to solve the classification problem, achieving significant results of both high accuracy and completeness on the collected dataset of ophthalmological indicators.

However, the data analyzed by the authors is often not complete and contains signs of synthesized subsamples, which complicates the procedure for assessing metric indicators. In [11], the authors use fully connected ANN models, comparing their work with each other and existing methods for classifying ML data, which is appropriate to demonstrate the capabilities and performance of these models in the context of their deployment in real diagnostic software and hardware tools for data analysis and evaluation. The effectiveness of using ML was also established by the authors [12] in studies in the field of treating respiratory system evaluating methods when analyzing the significance of individual signs and diseases course intensity indicators. According to a number of authors [13, 14], the use of methods for preventive intelligence medical data analysis in the context of assessing statistical indicators and correlations between individual data sets features allows the formation of an effective consistent data basis, thus organizing the process of automatic dimensionality reduction. As follows from the works [15, 16] on the use of ML in practice, the most effective are algorithms united in committees, this allows for models balancing and error values averaging for various metrics.

Thus, analyzing the research results in the reviewed authors works [10-16], it should be noted that more often they consider the possibility of using statistical approaches, as well as ML and ANN models, primarily to solve classification and regression problems in a heterogeneous feature space. However, outside the scope of research, questions remain related to assessing the individual medical indicators diagnostic significance, identifying correlations between input signs and assessing their specific weight values in the context of considering ophthalmological disease problems in patients with diabetic neovascular glaucoma.

The relevance of such studies is undoubted, due to the increasing disability of patients with this pathology throughout the world. In this regard, this work goal is to create and study different ANN models with generalization abilities and architecture for assessing the significance of features based on the collected data set on diagnostic indicators of inflammation and intraocular circulation in patients with diabetic neovascular glaucoma.

3. Collected data set description

The created set of the most significant target input and output features with medical interpretation was selected as a data set for analysis in applying ML process.

A retrospective cohort study of 127 patients (127 eyes) with a painful form of neovascular glaucoma (NVG) of diabetic origin was conducted at the Institute of Eye Diseases and Tissue Therapy named after. V.P. Filatov NAMS of Ukraine. The study protocol complied with the principles of the Declaration of Helsinki and was approved by the local institute’s bioethical committee. Written informed consent was obtained from all study participants.

The average age of patients is 65.0 years (62-68). The patients had uncontrolled intraocular pressure - IOP $\geq$ 30 mm Hg, and also had ocular pain syndrome. The main criterion for assessing the success of treatment was control of IOP and achieving its reduction by 20% from the initial value after 12 months of observation and maintaining best-corrected visual acuity (BCVA) - “success of treatment".

The preoperative visit (V0) was carried out on the eve of transscleral (TSC) cyclophotocoagulation (CPC), which was performed according to standard techniques. The laser power ranged from 850 to 1500 mW (Me 1100 mW), exposure time was 1.5-2 seconds, and laser coagulates average number was 22.

In all patients, the need for repeat CFC TSC was determined at each postoperative visit. The criterion for repeated laser treatment was maintaining high IOP values. Preoperative laboratory parameters were determined: neutrophils (N), lymphocytes (L), platelets (Throm), monocytes (M), glycosylated hemoglobin (HbA1c, %), volumetric intraocular circulation (RQ, ‰) according to rheoophthalmography. The systemic immune inflammation index (SII = Throm×[N/L]) and systemic inflammation response index (SIRI = N×[M/L]) were calculated. The SIRI and SII scores were further divided into quartiles, and the RQ score into quintiles. The data has been imported, divided into input and output characteristics. The target variable is the indicator “success of treatment”, which is a binary attribute. A fragment of the generated dataset is shown in Fig. 1. The results of related correlation analysis of the most significant features are shown in Fig. 2.

Name	Type	Missing	Statistics		
complications after CPT	Integer	0	0	1	0.591
LYM (109/L)	Real	0	Min 1.500	Max 2.400	Average 1.857
SIRI	Real	0	Min 0.460	Max 1.150	Average 0.727
SIRI kvartils	Integer	0	Min 1	Max 4	Average 2.244
SII	Real	0	Min 338	Max 673.900	Average 474.002
SII kvartils	Integer	0	Min 1	Max 4	Average 2.394
RQ critical - 0	Integer	0	Min 0	Max 1	Average 0.551
RQ V0	Real	0	Min 1.400	Max 5.650	Average 3.168
Success VOT	Polynominal	0	Least false (44)	Most true (83)	Values true (83), false (44)

***Figure 1**. Fragment of the input data set for analysis*

There is some imbalance in the output class values, which is not critical for research, and therefore the use of synthetic balancing methods is not advisable due to introducing additional noise risk into the data. As part of the reconnaissance data analysis, a correlation analysis of the feature space was performed in order to identify obvious patterns between individual input variables.

Attribut...	sex	age	HbA1	laserc...	Caute...	Laser ...	V0 VOT	GZ V0	com...	LYM (...	SIRI	SIRI ...	SII	SII kv...	RQ cr...	RQ V0	Succes...
sex	1	-0.048	0.021	0.206	-0.056	0.069	0.102	0.131	-0.078	0.014	0.016	0.048	-0.024	-0.009	-0.097	0.148	-0.026
age	-0.048	1	-0.273	-0.241	-0.018	-0.021	-0.041	-0.295	0.103	-0.202	-0.156	-0.103	-0.074	-0.081	0.183	0.285	-0.115
HbA1	0.021	-0.273	1	-0.189	0.306	0.363	0.573	-0.219	0.355	-0.149	0.815	0.758	0.774	0.633	-0.620	-0.156	0.751
lasercou...	0.206	-0.241	-0.189	1	0.004	-0.191	-0.193	0.349	-0.163	0.133	-0.163	-0.203	-0.209	-0.263	0.062	-0.110	-0.230
Cauteriz...	-0.056	-0.018	0.306	0.004	1	-0.093	0.141	-0.153	0.193	-0.040	0.316	0.301	0.356	0.377	-0.203	-0.187	0.304
Laser 81...	0.069	-0.021	0.363	-0.191	-0.093	1	0.251	-0.066	0.209	-0.144	0.360	0.409	0.410	0.365	-0.213	0.046	0.332
V0 VOT	0.102	-0.041	0.573	-0.193	0.141	0.251	1	-0.494	0.504	-0.115	0.643	0.590	0.657	0.563	-0.430	-0.123	0.611
GZ V0	0.131	-0.295	-0.219	0.349	-0.153	-0.066	-0.494	1	-0.574	0.107	-0.385	-0.397	-0.379	-0.334	0.279	0.018	-0.344
complica...	-0.078	0.103	0.355	-0.163	0.193	0.209	0.504	-0.574	1	-0.148	0.536	0.465	0.510	0.408	-0.397	-0.086	0.472
LYM (10...	0.014	-0.202	-0.149	0.133	-0.040	-0.144	-0.115	0.107	-0.148	1	-0.259	-0.281	-0.265	-0.307	0.025	-0.034	-0.183
SIRI	0.016	-0.156	0.815	-0.163	0.316	0.360	0.643	-0.385	0.536	-0.259	1	0.954	0.917	0.813	-0.744	-0.138	0.874
SIRI kvar...	0.048	-0.103	0.758	-0.203	0.301	0.409	0.590	-0.397	0.465	-0.281	0.954	1	0.874	0.806	-0.683	-0.061	0.822
SII	-0.024	-0.074	0.774	-0.209	0.356	0.410	0.657	-0.379	0.510	-0.265	0.917	0.874	1	0.893	-0.682	-0.047	0.876
SII kvartils	-0.009	-0.081	0.633	-0.263	0.377	0.365	0.563	-0.334	0.408	-0.307	0.813	0.806	0.893	1	-0.533	-0.001	0.736
RQ critic...	-0.097	0.183	-0.620	0.062	-0.203	-0.213	-0.430	0.279	-0.397	0.025	-0.744	-0.683	-0.682	-0.533	1	0.287	-0.807
RQ V0	0.148	0.285	-0.156	-0.110	-0.187	0.046	-0.123	0.018	-0.086	-0.034	-0.138	-0.061	-0.047	-0.001	0.287	1	-0.175
Success...	-0.026	-0.115	0.751	-0.230	0.304	0.332	0.611	-0.344	0.472	-0.183	0.874	0.822	0.876	0.736	-0.807	-0.175	1

Figure 2. *Related correlation analysis of the most significant features*

How can we note the indicators SIRI, SIRI quartiles. SII, SII quartiles have a high correlation with each other and with another input feature HbA1, which can introduce additional noise into the operation of models, and therefore it is necessary to take into account the nature of the influence of these features on the output target variable.

The correlation analysis showed a high positive relationship between the success of treatment and the indicators HbA1c (r = 0.751), SII (r = 0.876), SIRI quartiles (r = 0.874), IOP V0 (r = 0.611) and a high negative relationship with the RQ quintiles indicator (r=-0.807).

4. Development of artificial neural network models

We construct a computational process for constructing and studying ANN models to assess the accuracy of their training and classification. To do this, we use the Rapidminer system and blocks for importing a data set from a *.csv file, setting a target variable, and a subsystem for dividing the sample into test and validation. In order to conduct a systemic study and determine the ANN model type influence nature used on the final classification accuracy, assessed by different metrics, it is necessary to create several ANN models with different architectures.

In this regard, we use 3 types of ANN models: single perceptron (SP), multilayer perceptron (MP) and deep multilayer neural network (DNN).

In RapidMiner, ANN modeling can be implemented using the Neural Net operator, which allows for the creation, training, and evaluation of neural networks. Here's an overview of how ANN modeling works in RapidMiner:

Key Concepts of ANN in RapidMiner:

1. Neurons and Layers. Developed neural network models are consist of interconnected neurons, arranged in layers: an input layer, one or more hidden layers, and an output layer. Neurons in one layer are connected to neurons in the next, with each connection having a weight that is learned during the training process.

2. Activation Functions. Each neuron in a hidden layer applies an activation function to the weighted sum of its inputs to introduce non-linearity, which helps the model capture complex relationships. We've used Sigmoid functions and ReLU.

The structure of the developed process for creating and modeling ANN operation is shown in the figure 3. All blocks are unitary commands aimed at processing data and performing processes for modeling the operation of ANN.

Training and test data are divided among themselves in the proportion of 75% to 25%, respectively. The structure of the Validation block of the model is shown at fig.4.

Also, we used backpropagation learning algorithm. We build ANN training process on the base of backpropagation algorithm, where the network adjusts the weights of connections based on the error between predicted and actual outputs.

This is done iteratively over the training data to minimize the error, usually measured using a loss function (e.g., mean squared error for regression). During training, data is fed into the network, and the model adjusts weights to minimize prediction error. This is repeated over several iterations or epochs.

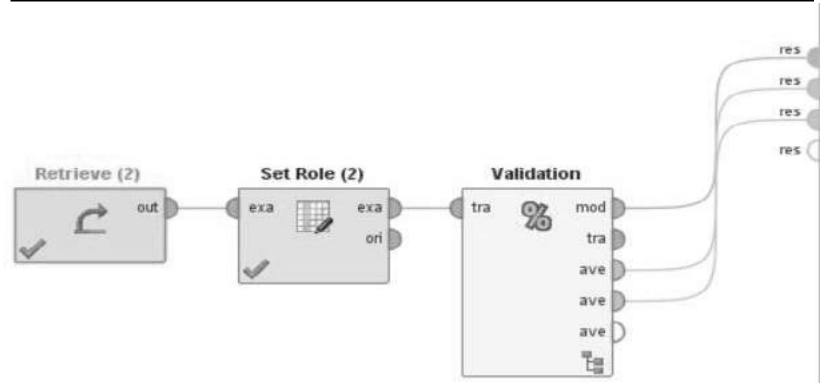

Figure 3. *The general structure of the developed process for creating and modeling the operation of an ANN*

Hyperparameters like learning rate, momentum, and number of hidden layers are critical in determining the success of the model.

During to our research in RapidMiner, model combination tools (also called ensemble techniques) are used to combine the predictions of multiple models to improve the overall performance.

These techniques aim to reduce errors, improve accuracy, and provide more robust predictions by leveraging the strengths of individual models.

We also provide a distinct module for confusion matrix calculation. It help us to provide different additional types of logical model's estimation, because of:

1. Correctly Identified Classes. True Positives (TP) t_p. These are instances where the ANN model correctly predicts the positive class. True Negatives (TN) t_n. These are instances where the model correctly predicts the negative class. TPs and TNs indicate the correct predictions made by the model, and their sum reflects the model's accuracy on correctly classified instances.

2. Misclassified Instances. False Positives (FP) f_p. These occur when the model incorrectly predicts a positive class when it should have predicted a negative one. In binary classification, this is known as a Type I error. False Negatives (FN) f_n. These occur when the model incorrectly predicts a negative class when the true label is positive. This is referred to as a Type II error. FPs and FNs indicate the errors made by the model. The distribution of these errors helps in identifying where the model is going wrong (e.g., if it's more prone to false alarms or missing positive cases).

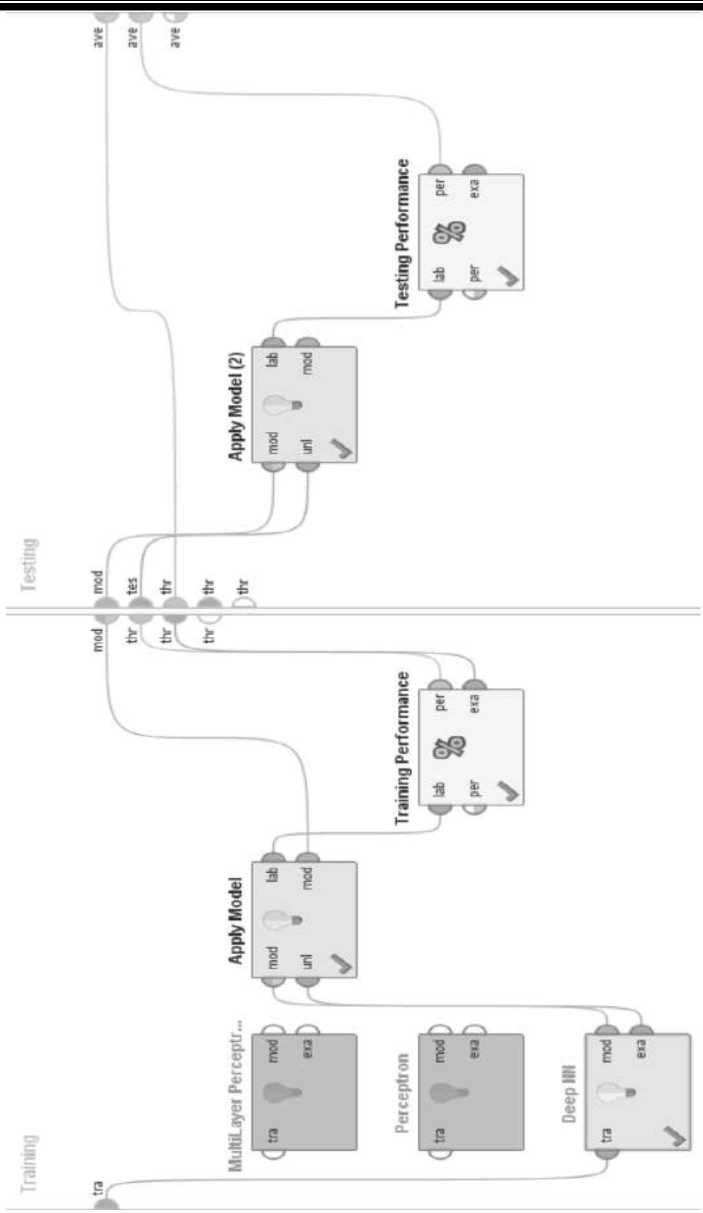

Figure 4. *The structure of the Validation block of the model*

3. Class Imbalance. Class Distribution. It is essential to consider the distribution of the classes when constructing the confusion matrix, especially in datasets where the classes are imbalanced (i.e., one class appears much more frequently than the other). In imbalanced datasets, accuracy alone can be misleading. For example, in a dataset with 90% negative and 10% positive examples, predicting every instance as negative would give a high accuracy but would fail on the positive class. The confusion matrix provides a more nuanced view of how the model is performing across different classes. As part of the process of studying the operation of the developed software, an assessment of the classification accuracy of the ANN models was carried out. ACCURACY was used as a numerical characteristic for assessing the operation of the software – a reliability metric that allows assessing the accuracy of classification, i.e. determining the proportion of correctly classified examples.

$$Accuracy = \frac{t_p + t_n}{t_p + f_p + f_n + t_n}$$

5. Experiments and results

10 computational experiments were carried out with different sets of parameters (hidden layers number, activation functions, learning rate coefficients, error rates). The average obtained metrics values for assessing model’s accuracy are shown in the table below. In particular, for a single perceptron, the value of the learning rate coefficient varied from 0.05 to 0.3. For a multilayer perceptron, 2 hidden layers of 5 and 3 neurons were created, respectively. For the deep ANN, 50 hidden layers were created. The results of obtained ANN model’s metrics evaluation to the first experiment are given in Table 1. The confusion matrices for each of the created neural network models for the training and test samples are shown in the following (Fig.5-7). Visualization of the results of comparing estimates of accuracy and error metrics for the created ANN models of the first experiment is shown in Fig.8.

As we can see (Fig. 8), the least accurate classification results on both the training (about 84%) and training set were shown by the SP model (about 87%), which is due to its simplified architecture, the absence of hidden layers and a number of hyperparameters. affecting model’s final accuracy.

accuracy: 86.90%

	true true	true false	class precision
pred. true	50	6	89.29%
pred. false	5	23	82.14%
class recall	90.91%	79.31%	

accuracy: 83.72%

	true true	true false	class precision
pred. true	22	1	95.65%
pred. false	6	14	70.00%
class recall	78.57%	93.33%	

Figure 5. *Confusion of training and test sample matrices for the Single Perceptron model*

Table 3
The results of obtained ANN model's metrics evaluation to the first experiment

	Metric values of ANN models					
	Single Perceptron (SP)		Multilayer Perceptron (MP)		Deep NN (DNN)	
Metric name	Training set	Test ıple	Training set	Test sample	Training set	Test sample
Accuracy	86.90%	83.72 %	100.00%	95.35%	100.00%	97.67%
Classification error	13.10%	16.28 %	0.00%	4.65%	0.00%	2.33%
Kappa	0.708	0.667	1.000	0.898	1.000	0.948
Weighted mean recall	85.11%	85.95 %	100.00%	94.88%	100.00%	96.67%
Weighted mean precision	85.71%	82.83 %	100.00%	94.88%	100.00%	98.28%
Spearman rho	0.708	0.667	1.000	0.898	1.000	0.949
Kendall tau	0.708	0.667	1.000	0.898	1.000	0.949
Absolute error	0.131 +/- 0.337	0.163 '- 0.369	0.004 +/- 0.013	0.038 +/- 0.171	0.029 +/- 0.070	0.036 +/- 0.150
Relative error	13.10% +/- 33.73%	16.28 % +/- 6.92%	0.37% +/- 1.33%	3.78% +/- 17.14%	2.85% +/- 6.99%	3.57% +/- 14.99%
Squared error	0.131 +/- 0.337	16.28 % +/- 6.92%	0.001 +/- 0.001	0.031 +/- 0.157	0.006 +/- 0.020	0.024 +/- 0.143
Corellation	0.708	0.667	1.000	0.898	1.000	0.949
Logistic loss	0.363	0.375	0.314	0.327	0,293	0,305

accuracy: 100.00%

	true true	true false	class precision
pred. true	55	0	100.00%
pred. false	0	29	100.00%
class recall	100.00%	100.00%	

accuracy: 95.35%

	true true	true false	class precision
pred. true	27	1	96.43%
pred. false	1	14	93.33%
class recall	96.43%	93.33%	

Figure 6. *Confusion of training and test sample matrices for the Multilayer Perceptron model*

accuracy: 100.00%

	true true	true false	class precision
pred. true	55	0	100.00%
pred. false	0	29	100.00%
class recall	100.00%	100.00%	

accuracy: 97.67%

	true true	true false	class precision
pred. true	28	1	96.55%
pred. false	0	14	100.00%
class recall	100.00%	93.33%	

Figure 7. *Confusion of training and test sample matrices for the Deep NN model*

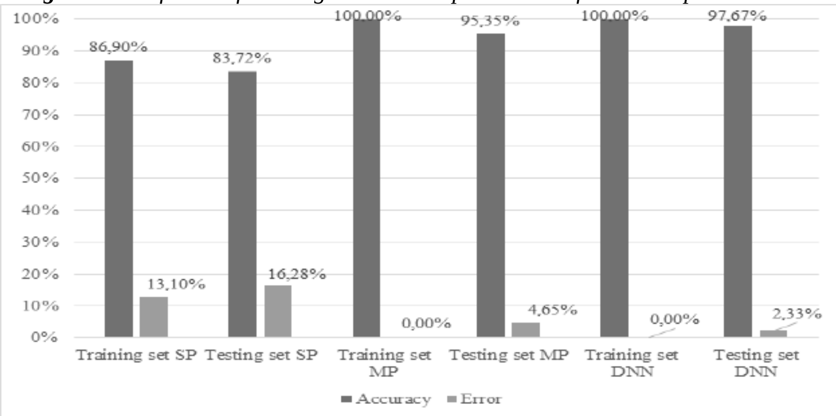

Figure 8. *Visualization of the results of comparing estimates of accuracy and error metrics for the created ANN models of the first experiment*

The MP and DNN models showed approximately the same results for metric evaluations on the training subset of data, however, the accuracy of DNN was 2.3% higher on the test subset, which may be due to a more complex model structure and a larger number of computational iterations.

An additional experiment was aimed at reducing the dimension of input features. A unique identifier, calculated values of SIRI quartiles, SII, HbA1c, RQ quintiles, IOP V0 were used as input features; the output feature is “success of treatment”. Results of calculating weighting coefficients for input parameters of the data set for the first set of experiments are shown in Fig.9. As we can see, the values are scaled to the range {0,1}, with occupies the smallest value – SII, and the largest – RQ critical -0.

attribute	weight	attribute	weight
id	0.626	complications after CPT	0.241
sex	0.430	LYM (109/L)	0.684
age	0.604	SIRI	0.019
HbA1	0.100	SIRI kvartils	0.030
lasercougul (1-yes)	0.662	SII	0
Cauterization	0.302	SII kvartils	0.043
Laser 810 power	0.280	RQ critical - 0	1
V0 VOT	0.140	RQ V0	0.533
GZ V0	0.801	Success VOT	0.065

Figure 9. *Results of calculating weighting coefficients for input parameters of the data set for the first set of experiments*

Results of calculating weighting coefficients for initial input parameters of the data set for the second set of experiments are shown in Fig.10.

As can be seen, the spread of values is quite significant, which is associated with stochastic processes in the formation of weighting coefficients of initial data in the process of initializing the ANN models within the framework of the experimental studies conducted when launching the Rapidminer working environment.

Attribute	Weight
ID	497.739
SIRI kvar...	-804.910
SII	80.686
RQ critic...	771.697
Success...	1482.795
VOT V5	-4124.037

Figure 10. *Results of calculating weighting coefficients for initial input parameters of the data set for the second set of experiments*

Research modelling was performed on the created three ANN models with the same parameters. The results of evaluating the metrics of the obtained NN models for the second experiment are given in Table 2.

Table 2
Results of evaluation of metrics of the obtained ANN models for the second experiment

Metric name	Metric values of ANN models Single Perceptron (SP) Training set	 Test ample	 Multilayer Perceptron (MP) Training set	 Test sample	 Deep NN (DNN) Training set	 Test sample
Accuracy	6.90%	3.81%	100.00%	93.02%	100.00%	90.70%
Classification error	2.56%	6.19%	0.00%	6.98%	0.00%	9.30%
Kappa	.355	.487	1.000	0.822	1.000	0.769
Weighted mean recall	2.31%	0.42%	100.00%	90.05%	100.00%	88.44%
Weighted mean precision	7.97%	4.89%	100.00%	92.33%	100.00%	88.44%
Spearman rho	.400	.550	1.000	0.824	1.000	0.769
Kendall tau	.400	.550	1.000	0.824	1.000	0.769
Absolute error	.326 +/- .469	.262 +/ .440	0.009 +/- 0.017	0.049 +/- 0.153	0.029 +/- 0.070	0.112 +/- 0.21
Relative error	2.56% +/- 6.86%	6.19% +/ 3.97%	0.94% +/- 1.65%	4.91% +/- 15.26%	2.85% +/- 6.99%	11.18% +/- 21.13%
Squared error	.326 +/- .469	.262 +/ .440	0.001 +/- 0.000	0.026 +/- 0.094	0.006 +/- 0.020	0.057 +/- 0.163
Corellation	.400	.550	1.000	0.824	1.000	0.769
Logistic loss	.437	.413	0.316	0.329	0,293	0.349

The nature of the obtained results differs somewhat from the previously calculated metric values for the better, however, due to the presence of noise effects, the values are not ideal. Adjacent correlation analysis of the most significant features of experiment 2 is shown in Fig.11.

As we can see, a number of features have a significant correlation with each other, which may be an additional reason for the presence of noise effects during training and testing of models. Visualization of the results of comparing estimates of accuracy and error metrics for the created ANN models for the second experiment is shown in Fig.12.

In the third experiment, gender was excluded from the original sample. This reduces noise effects because this feature does not carry a significant semantic load in the context of data analysis. Results of calculating weighting coefficients for initial input parameters of the data set for the third set of experiments are shown in Fig.13.

Attribut...	SIRI kva...	SII	RQ criti...	Succes...	VOT V5	Succes...
SIRI kvar...	1	0.874	-0.683	-0.822	0.423	-0.163
SII	0.874	1	-0.682	-0.876	0.415	-0.165
RQ critic...	-0.683	-0.682	1	0.807	-0.250	-0.006
Success...	-0.822	-0.876	0.807	1	-0.376	0.130
VOT V5	0.423	0.415	-0.250	-0.376	1	-0.781
Success...	-0.163	-0.165	-0.006	0.130	-0.781	1

Figure 11. *Adjacent correlation analysis of the most significant features of experiment 2*

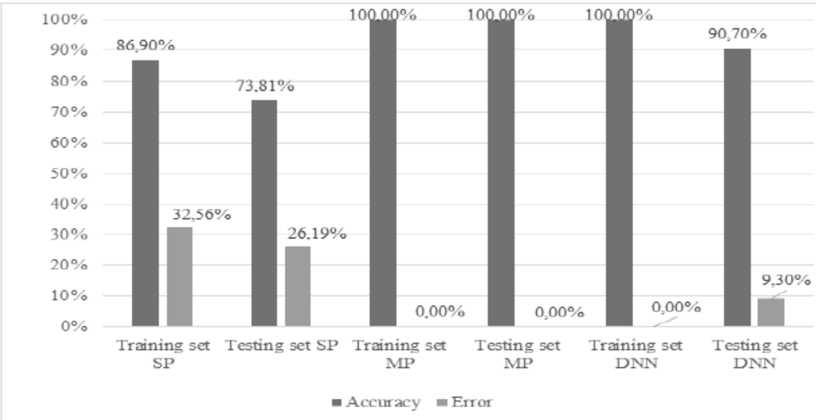

Figure 12. *Visualization of the results of comparing estimates of accuracy and error metrics for the created ANN models for the second experiment*

The results of evaluation of the metrics of the obtained NN models for the third experiment are given in Table 3.

Attribute	Weight	Attribute	Weight
id	3412.931	complica...	453.048
age	-9980.359	LYM (10...	-512.371
HbA1	-402.500	SIRI	243.611
lasercou...	-366.595	SIRI kvar...	967.704
Cauteriz...	-4811.274	SII	1441.607
Laser 81...	-73.848	SII kvartils	-245.604
V0 VOT	-3350.753	RQ critic...	-1097.827
GZ V0	-27.132	RQ V0	-1436.026

Figure 13. *Results of calculating weighting coefficients for initial input parameters of the data set for the third set of experiments*

The final values of the weighting coefficients of the most significant features determined by DNN are shown in Fig.14.

attribute	weight
ID	0
SIRI kvar...	0.481
SII	0.595
RQ critic...	0.592
Success...	1
VOT V5	0.643
Success...	0.806

Figure 14. *The final values of the weighting coefficients of the most significant features determined by DNN*

Visualization of the results of comparing estimates of accuracy and error metrics for the created ANN models for the third experiment is shown in Fig.15. As a result of ANN constructing significant input parameters were established that most influence the effectiveness of treatment: SIRI quartiles, SII, HbA1c and RQ quintiles. During inflammation, leukocytes accumulate in the lesion and the speed of local blood flow decreases. Impaired retinal microcirculation is observed in patients with type 2 diabetes mellitus (T2DM) without clinically significant diabetic retinopathy [17].

Table 3
Results of evaluation of metrics of the obtained ANN models for the third experiment

	Metric values of ANN models					
	Single Perceptron (SP)		Multilayer Perceptron (MP)		Single Perceptron (SP)	
Metric name	Training set	Metric name	Training set	Metric name	Training set	Metric name
Accuracy	89.29%	95.35%	100.00%	97.67%	98.81%	97.67%
Classification error	10.71%	4.65%	0.00%	2.33%	1.19%	2.33%
Kappa	0.769	0.898	1.000	0.948	0.974	0.948
Weighted mean recall	89.37%	94.88%	100.00%	96.67%	99.09%	96.67%
Weighted mean precision	87.74%	94.88%	100.00%	98.28%	98.33%	98.28%
Spearman rho	0.771	0.898	1.000	0.949	0.974	0.949
Kendall tau	0.771	0.898	1.000	0.949	0.974	0.949
Absolute error	0.107 +/- 0.309	0.047 +/- 0.211	0.005 +/- 0.016	0.024 +/- 0.151	0.034 +/- 0.101	0.037 +/- 0.147
Relative error	10.71% +/- 30.93%	4.65% +/- 21.06%	0.47% +/- 1.57%	2.39% +/- 15.06%	3.41% +/- 10.11%	3.70% +/- 14.69%
Squared error	0.107 +/- 0.309	0.047 +/- 0.211	0.000 +/- 0.001	0.023 +/- 0.151	0.011 +/- 0.058	0.023 +/- 0.140
Corellation	0.771	0.898	1.000	0.949	0.974	0.949
Logistic loss	0.354	0.331	0.315	0.322	0.324	0.326

Altered blood flow in patients with type 2 diabetes contributes to macrovascular (peripheral vascular disease and coronary artery disease) and microvascular (diabetic retinopathy and diabetic nephropathy) complications. Our results confirm that higher values of SIRI, SII, HbA1c and RQ indicate the severity of the disease and determine the need for diabetes mellitus stabilization [4], additional anti-inflammatory and antiischemic treatment in patients with painful NVG of diabetic origin to improve treatment prognosis.

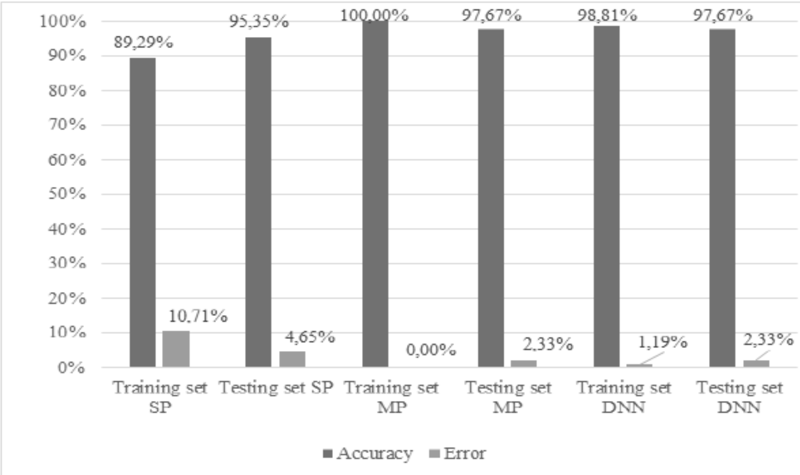

Figure 15. *Visualization of the results of comparing estimates of accuracy and error metrics for the created ANN models for the third experiment*

6. Conclusion

The conducted research made it possible to establish a higher adaptability of deep learning models, in particular deep fully connected ANN models, for the analyzed data set.

The SP model demonstrates unstable accuracy on the training and test samples in different experiments, which may be due to shortcomings in determining the weight values of features at different iterations, as well as limitations in generalization ability.

The MP model is more stable in all experiments, demonstrating high accuracy and completeness; its construction speed is 25% higher than the SP model. Presumably, the accuracy of this model can be increased through a more efficient organization model hyperparameter values automating selection process. The DNN model is the most accurate, but its training process is the most resource-intensive and takes almost 4 times longer than the SP model.

A promising direction for further research is the search for optimizing the ANN training process in order to minimize training and testing errors, as well as increase the generalization ability of models in general. It has been established that the most significant diagnostic features of the collected dataset in the context of the problems under consideration are the indicator of glycosylated hemoglobin HbA1c of the volumetric intraocular circulation (RQ), the systemic inflammatory response index (SIRI) and the systemic inflammation index (SII).

The data obtained allow us to conclude that the use of neural networks to predict the effectiveness of treatment is justified and timely. By taking into account and stabilizing blood sugar, as well as adjusting indicators of inflammation and microcirculation in patients with diabetic neovascular glaucoma, it is possible to ensure a timely significant reduction in intraocular pressure, maintain visual acuity, and therefore the quality of life of patients.

The data obtained allow us to conclude that the use of different ANN to predict the effectiveness of transcleral cyclophotocoagulation is justified. Timely prescribed treatment of identified changes in glycolyzed hemoglobin (HbA1c), volumetric intraocular circulation (RQ), systemic inflammatory response index (SIRI) and systemic inflammation index (SII) against the background of transcleral cyclophotocoagulation can provide a significant reduction in intraocular pressure of at least 20%, maintaining visual acuity and patients life quality with painful diabetic neovascular glaucoma.

7. References

[1] V. Vychuzhanin, N. Rudnichenko, Z. Sagova, M. Smieszek, V. V. Cherniavskyi, A. I. Golovan, M. V. Volodarets, Analysis and structuring diagnostic large volume data of technical condition of complex equipment in transport. IOP Conference Series: Materials Science and Engineering, Volume 776, 24th Slovak-Polish International Scientific Conference on Machine Modelling and Simulations - MMS 2019, 3-6 September 2019, Liptovský Ján, Slovakia, 2019 pp.1-11. DOI:10.1088/1757-899X/776/1/012049.

[2] N. Rudnichenko, V. Vychuzhanin, I. Petrov, D. Shibaev, Decision Support System for the Machine Learning Methods Selection in Big Data Mining, in: Proceedings of The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020), CEUR-WS, 2608, 2020, pp. 872-885.

[3] M. Alloghani, D. Al-Jumeily, J. Mustafina, A. Hussain, A.J. Aljaaf, A systematic review on supervised and unsupervised machine learning algorithms for data. Supervised and Unsupervised Learning for Data Science, Unsupervised and Semi-supervised Learning, Springer International Publishing, Cham (2014) 3-21. DOI: 0.1007/978-3-030-22475-2_1.

[4] O. Guzun, O. Zadorozhnyy, L. Velychko, O. Bogdanova, L. Dumbrăveanu, V. Cuşnir, A. Korol, The effect of the intercellular adhesion molecule-1 and glycated haemoglobin on the management of diabetic neovascular glaucoma, Romanian Journal of Ophthalmology, 68 2 (2024) 135-142. DOI:10.22336/rjo.2024.25.

[5] J.De. Fauw, J.R. Ledsam, B. Romera-Paredes, Clinically applicable deep learning for diagnosis and referral in retinal disease, Nature Medicine, 24 9 (2018) 1342–1350. DOI: 10.1038/s41591-018-0107-6.

[6] B. Al-Bander, W. Al-Nuaimy, B.M. Williams, Y. Zheng, Multiscale sequential convolutional neural networks for simultaneous detection of fovea and optic disc, Biomed Signal Process Control, 40 (2018) 91–101. DOI: 10.1016/j.bspc.2017.09.008.
[7] B.J. Antony, S. Maetschke, R. Garnavi, Automated summarisation of SDOCT volumes using deep learning: Transfer learning vs de novo trained networks, PLoS ONE, 14 5 (2019). DOI: 10.1371/journal. pone.0203726.
[8] J. Bertomeu, Machine learning improves accounting: discussion, implementation and research opportunities, Review of Accounting Studies, 25 3 (2020) 1135-1155. DOI:10.1007/s11142-020-09554-9
[9] M. Solanki, Ms. Sharma, A Review of Data Mining Techniques and Its Applications, International Journal of Innovative Research in Computer Science & Technology, 122 (2021) 100-104. DOI:10.55524/ijircst.2021.9.6.23.
[10] S. Sheshkal, M. Gundersen, M. Riegler, Ø. Utheim, K. Gundersen, H. Hammer, Classifying Dry Eye Disease Patients from Healthy Controls Using Machine Learning and Metabolomics Data (2024).
[11] A. Abdullah, A. Aldhahab, H. Abboodi, Review of Eye Diseases Detection and Classification Using Deep Learning Techniques, BIO Web of Conferences, 97 (2024). DOI:10.1051/bioconf/20249700012.
[12] A. Selim, M. Barz, O. Bhatti, H. Alam Tusfiqur, D. Sonntag, A review of machine learning in scanpath analysis for passive gaze-based interaction, Frontiers in Artificial Intelligence, 7 (2024). DOI:10.3389/frai.2024.1391745.
[13] J. Van Eijgen, V. Schuhmann, E.-L. Fingerroos, M. Renier, H. Burchert, J. Kröpfl, A. Craenenbroeck, V. Cornelissen, K. Gugleta, I. Stalmans, H. Hanssen, High-intensity interval training in patients with glaucoma (HIT-GLAUCOMA): protocol for a multicenter randomized controlled exercise trial, Frontiers in Physiology, 15 (2024). DOI:1349313. 10.3389/fphys.2024.1349313.
[14] C.-H. Moon, Ah-Jun Lee, H.-Y. Jeon, E.-B. Kim, K.-S. Ha, Therapeutic effect of ultra-long-lasting human C-peptide delivery against hyperglycemia-induced neovascularization in diabetic retinopathy, Theranostics, 13 (2023) 2424-2438. DOI:10.7150/thno.81714.
[15] E. Tokuç, M. Eksi, R. Kayar, S. Demir, R. Topaktaş, Y. Bastug, M. Akyuz, M. Ozturk, Inflammation indexes and machine-learning algorithm in predicting urethroplasty success, Investigative and Clinical Urology, 65 (2024) 240. DOI: 10.4111/icu.20230302.
[16] A. Abujaber, Y. Imam, I. Albalkhi, S. Yaseen, A. Nashwan, N. Akhtar, Utilizing machine learning to facilitate the early diagnosis of posterior circulation stroke, BMC Neurology, 24 (2024) 1-12. DOI:10.1186/s12883-024-03638-8.
[17] V. Cankurtaran, M. Inanc, K. Tekin, F. Turgut, Retinal Microcirculation in Predicting Diabetic Nephropathy in Type 2 Diabetic Patients without Retinopathy, Ophthalmologica, 243 4 (2020) 271-279. DOI: 10.1159/000504943.

ADVANCES
IN INFORMATION-CONTROL SYSTEMS AND TECHNOLOGIES

ШТУЧНИЙ ІНТЕЛЕКТ В ДІАГНОСТИЦІ, ПРОГНОЗІ ТА ЛІКУВАННІ ДІАБЕТИЧНОЇ НЕОВАСКУЛЯРНОЇ ГЛАУКОМИ

Dr.Sci. В. Вичужанін ORCID: 0000-0002-6302-1832
Національний університет «Одеська політехніка», Україна
E-mail: vychuzhanin.o.v@op.edu.ua
Ph.D. О. Гузун ORCID: 0009-0003-6873-8503
1ДУ «Інститут очних хвороб і тканинної терапії ім. Філатова НАМН України», Україна
E-mail: olga.v.guzun@gmail.com
Ph.D. М. Рудниченко ORCID: 0000-0002-7343-8076
Національний університет «Одеська політехніка», Україна
E-mail: nickolay.rud@gmail.com
Ph.D. О. Задорожний ORCID: 0000-0003-0125-2456
1ДУ «Інститут очних хвороб і тканинної терапії ім. Філатова НАМН України», Україна
E-mail: zadoroleg2@gmail.com
Ph.D. А.Король ORCID: 0000-0003-0516-308X
1ДУ «Інститут очних хвороб і тканинної терапії ім. Філатова НАМН України», Україна
E-mail: andrii.r.korol@gmail.com

Анотація. *Ця робота зосереджена на ключових аспектах діагностики, прогнозу та лікування неоваскулярної глаукоми діабетичного походження на основі підходів машинного навчання та, зокрема, різних архітектур штучних нейронних моделей. Проведено аналіз актуальності, пріоритетних положень і переваг використання методів машинного навчання, розглянуто існуючі підходи, які використовуються в сучасній літературі в контексті досліджуваної теми, особливості їх інтеграції в процес діагностичного аналізу Описано простір ознак агрегованого та позначеного авторами набору даних про пацієнтів із проблемами зору, зокрема тих, хто страждає на неоваскулярну глаукому діабетичного походження. Проведено кореляційний аналіз вхідних ознак, побудовано, навчено та протестовано 3 різні моделі штучних нейронних мереж, експериментально розраховано та досліджено метрики для оцінки точності їх роботи, визначено статистичні показники, у тому числі похибки та втрати, що характеризують їх узагальнюючу здатність. Аналіз результатів, отриманих у ході досліджень, дав змогу визначити переважаючі вхідні характеристики та оцінити їхній вплив на цільову вихідну змінну та загальну значимість у просторі ознак набору даних, а також встановити найбільш прийнятні моделі для даних аналізу з точки зору їх точності та швидкості. Проведене дослідження дозволило встановити факт повної адаптованості моделей глибокого навчання штучних нейронних мереж до аналізованого набору даних.*

Ключові слова. *Штучний інтелект, неоваскулярна глаукома, діагностика лікування очей, нейронні мережі, аналіз даних, аналіз даних, машинне навчання*