Assessing the Impact of Expert Labelling of Training Data on the Quality of Automatic Classification of Lithological Groups Using Artificial Neural Networks

Yan Kuchin[1*], Ravil Mukhamediev[2], Kirill Yakunin[3], Janis Grundspenkis[4], Adilkhan Symagulov[5]
[1, 4, 5]*Riga Technical University, Riga, Latvia*
[1–3, 5]*Satbayev University, Almaty, Kazakhstan*

Abstract – **Machine learning (ML) methods are nowadays widely used to automate geophysical study. Some of ML algorithms are used to solve lithological classification problems during uranium mining process. One of the key aspects of using classical ML methods is causing data features and estimating their influence on the classification. This paper presents a quantitative assessment of the impact of expert opinions on the classification process. In other words, we have prepared the data, identified the experts and performed a series of experiments with and without taking into account the fact that the expert identifier is supplied to the input of the automatic classifier during training and testing. Feedforward artificial neural network (ANN) has been used as a classifier. The results of the experiments show that the "knowledge" of the ANN of which expert interpreted the data improves the quality of the automatic classification in terms of accuracy (by 5 %) and recall (by 20 %). However, due to the fact that the input parameters of the model may depend on each other, the SHapley Additive exPlanations (SHAP) method has been used to further assess the impact of expert identifier. SHAP has allowed assessing the degree of parameter influence. It has revealed that the expert ID is at least two times more influential than any of the other input parameters of the neural network. This circumstance imposes significant restrictions on the application of ANNs to solve the task of lithological classification at the uranium deposits.**

Keywords – **Assessment of expert influence, lithology, machine learning, SHAP, uranium mining.**

I. Introduction

Kazakhstan provides about 39 % of the world uranium production [1]. Production is carried out by in-situ leaching. In this method, uranium is extracted through a network of pumping-in and pumping-out wells, along which the leaching solution circulates. The purpose of the Geophysical Data Interpretation for Boreholes (GDIB) in this case is not only to determine the position of the ore body, but also to identify the type and parameters of the enclosing rocks, their filtration properties. Since the extraction process is carried out by spreading the leaching solution, the isolation of impermeable

and permeable rocks, the determination of their filtration properties are critically important.

At the same time, accurate lithological interpretation can be carried out only taking into account all available data for the given well, on neighbouring wells, and also taking into account the information obtained at the exploration stage [2]. Obviously, an accurate interpretation requires experience and time, while the extraction technology requires prompt decision-making.

Erroneous and inaccurate results from geophysical data analysis may lead to serious financial loses on different levels: from an overall decrease in the quantity of active boreholes to unjustified labour cost and low volume production.

The most complete information on the host rocks and their filtration properties can be obtained by a joint analysis of geophysical logging data based on various physical principles. There is no exhaustive theory for automatic lithological interpretation, which takes into account all types of logging, and it is unlikely that such a theory can be developed. Interpretation of logging data is a complex and poorly formalised task, especially if one tries to take into account all the parameters and their mutual influence, geological cross-sections and data obtained during the exploration phase. Statistical methods and machine learning are successfully used to solve such problems. From the point of view of machine learning, the majority of geophysical problems, in particular the problem of lithological interpretation, when it is necessary to determine the type of rock based on the recorded physical parameters, is a classification problem.

Different classifiers were successfully tested for GDIB interpretation automation on uranium deposits of Kazakhstan [3].

However, the success of data interpretation processes heavily depends on properly prepared input data. The accuracy of automated log data classification to a great extent depends on expert's manual assessment because it is used as input for training machine learning (ML) automatic classifiers and also

---

* Corresponding author's e-mail: ykuchin@mail.ru

as an etalon for comparison. However, cross-comparison of assessments for core sampling provided by different experts shows significant discrepancies (Table I). Specifically, the difference between assessments is higher for permeable rocks (e.g., sands have more deviations in manual classification than clays). In this paper, we aim to capture the nature of this phenomenon and to propose some measures to take it into account while applying artificial neural network (ANN) that is trained on input data assessed by experts.

The paper is organised as follows. Section 1 provides a literature review, with delineation of known issues for ANN classification of log data from uranium mines. In Section 2, the problem of inconsistency of expert's assessments is described. Section 3 provides the proposed method of dealing with inconsistency of expert's assessments. Section 4 demonstrates results of application of interpreter's ID and SHAP. In Section 5, conclusions, with summary of results and recommendations for the next set of analytical experiments design are given.

II. RELATED WORK

ANN is a part of large research area named machine learning [4]. ML has been successfully used to solve problems of cluster analysis and classification using a wide range of algorithms, including number of ANN types [5].

Since 1970, the artificial neural network modelling has been widely used in petrography to analyse lithologic log data, evaluate geo-mineral resources, make deep seismic sounding, and perform many other geophysical analyses [6]–[16]. Dr Kostikov thoroughly describes methods for geophysical data interpretation of boreholes based on transformation of log diagrams using a multi-layer neural network [17]. Several publications demonstrate the application of feedforward neural network for interpretation of geophysical data from uranium deposits [18]–[20].

There are a number of publications focused on tasks and issues related to automatic interpretation of log data from uranium deposits. For example, results of analytical testing with ANN as an approach for log data classification can be found in publications [21]–[23], while several ML methods and their comparative results are described in publications [24], [25]. There, it has been shown that a feedforward neural network demonstrates a much better classification quality when compared to *k*-nearest neighbour (*k*-NN) or support vector machine (SVM) algorithms. Furthermore, results from a combination of ML algorithms applied to a similar underlaying task are reviewed in publications [26], [27]. The above-mentioned challenges 2 and 4 have been addressed during our analytical experiments.

Our earlier experiments demonstrated that, on average, the automatic classification with a feedforward neural network performs at ~55 % accuracy [28]. It was also shown that ambiguity of expert assessment had a negative effect on auto-classification accuracy. Therefore, this particular study is focused on challenge #1: the evaluation of the impact the expert assessment of electric logging has on auto-classifiers, and more specifically on neural networks – during ML training.

To validate and confirm our hypothesis about a negative impact of expert opinion on interpretation accuracy, we tackled it from several angles. Firstly, we defined the range of values for auto-classifiers by using a synthetic dataset [28]. Then, we evaluated inconsistency of expert opinions in geophysical data assessment using one specific example. Thirdly, we calculated ANN qualitative indicators for two large datasets with subsequent assessment by various experts.

III. PROBLEM

Inconsistency of expert assessment introduces additional difficulties into the process of ML system training. Despite the fact that expert assessments are used for training, each expert does classification differently. Three boreholes with identification numbers 2100, 2104, and 4939 were used for expert assessment quality comparison. Log-data for these boreholes were assigned to three independent experts, named here D, L, and T. Moreover, boreholes 2100 and 4939 have actual core sampling data. The four main indicators of quality (accuracy, recall, precision, and Kappa) were calculated based on the assessment data for lithologic differentiation and core sampling received from experts D, L, and T. It was done as pair-wise comparison, when one of the experts was considered a standard and the other one – variable (Table I). The average value for experts' accuracy is 0.67 with scatter score F1 = 0.6. When similar comparison is done between core sampling and expert assessment, the accuracy is 0.5 and F1 score is 0.27. At the same time, taking into consideration expert biases, assessment of boreholes 2100 and 2104 done by experts D and L has the highest agreement (accuracy > 0.8). Expert D has the highest accuracy when compared with an average core value for borehole 2100 (accuracy 0.69), and experts D and T are in good agreement for borehole 4939 assessment (accuracy 0.77). For borehole 4939, the average core value is well aligned with the assessment by expert L (accuracy = 0.60). When comparison is done between expert assessment and core sampling, the quality values are significantly lower than comparison is done between experts. Our analysis showed that for some critical lithotypes (e.g., claystone) expert assessments aligned by 70–95 %, yet they much less agreed with core sampling data. For claystone extraction the core and expert assessment correlated by 30–50 %. This is largely due to critical requirement for in-situ leaching processing to define impenetrable lithotypes because uranium fields in Kazakhstan are comprised primarily by sands, claystone, siltstones, and sandstones.

It is known that clay exhibits minimal apparent resistivity (AR) and maximal spontaneous polarization (SP) values, which simplifies its identification. On the other hand, definition of various sub-layers in penetrable lithologic types (e.g., sub-layers of mixed sand at the medium sand layer) is not critical from a technological point of view and can easily be substituted by the biased opinion of an individual expert.

Comparison of expert assessments suggests several options for observed inconsistency minimisation. One of them could be grouping borehole data assessed by the one and the same expert only. Another option is to use one of the experts as an identification (ID) parameter for ML training.

TABLE I

PAIR-WISE COMPARISON

Experts	Accuracy	Kappa	Recall	Precision	T1 Score
2100					
D vs L	0.81	0.70	0.63	0.61	0.6180
D vs T	0.71	0.54	0.42	0.49	0.4553
L vs T	0.80	0.67	0.46	0.52	0.4930
4939					
D vs L	0.3317	0.16	0.5979	0.4975	0.5430
D vs T	0.7706	0.66	0.7941	0.6879	0.7371
L vs T	0.3762	0.19	0.6086	0.5495	0.5775
2104					
D vs L	0.8409	0.76	0.8445	0.8205	0.8323
D vs T	0.6551	0.49	0.5493	0.5424	0.5458
L vs T	0.7213	0.57	0.5845	0.6011	0.5926
Experts average	**0.67**	**0.53**	**0.61**	**0.59**	**0.60**
2100					
Core vs D	0.693	0.39	0.3624	0.3487	0.3554
Core vs L	0.6444	0.23	0.2775	0.2423	0.2587
Core vs T	0.6505	0.21	0.274	0.2441	0.2581
4939					
Core vs D	0.1749	0.04	0.2011	0.2295	0.2143
Core vs L	0.6089	0.37	0.3066	0.3986	0.3465
Core vs T	0.2096	0.04	0.2004	0.2231	0.2111
Core average	**0.50**	**0.21**	**0.27**	**0.28**	**0.27**

To perform automatic classification, it is necessary to select the type of classifier and its parameters. The results of work of several types of classifiers as applied to the logging data are given in Appendix (https://drive.google.com/open?id=1al9QqJ GydtAwNzapWribV3rBPzkSPdvC) [28]. The evaluation of the classifier parameters allows us to conclude that ANN and XGBoost demonstrate the best classification results. The same three boreholes were interpreted with a feedforward ANN with 26 input neurons (depth, coordinates, floating windows for AR and SP curves) and two hidden layers built using Keras library [29].

Log curves, after preliminary normalization, were presented as floating windows because for the accurate lithotype definition not only the value in a given point was important, but the curve shape itself (extremes, curvatures) was a key criterion. In addition, it was decided to add geographic coordinates of boreholes as training parameters, because lithologic structures of neighbouring boreholes were well correlated (this fact was also used by experts in defining lithotypes). This approach allowed us to achieve accuracy of identification comparable to the expert assessment of core sampling (Table II). However, for impenetrable layers (claystone, siltstone) in core boreholes the accuracy of ANN algorithm is lower than expert assessment: precision = 0.30, and recall = 0.22. (Table III), i.e., the majority

of impenetrable layers, critical from a technological point of view, are still poorly identified.

TABLE II

EXPERT ASSESSMENT OF CORE SAMPLING

Outputs from ANN Algorithm for Core Borehole number	Accuracy	Precision	Recall	T1 Score
2100	0.4273	0.4623	0.2701	0.3409
2104	0.7053	0.7961	0.6053	0.6877
4939	0.3277	0.2573	0.1181	0.1619
Average	**0.5092**	**0.5069**	**0.3409**	**0.4076**

Overall, low accuracy of assessment for these boreholes by both experts and ANN model can be explained by the fact that data interpretation was performed not based on log curve shapes as per regular procedure, but based on core sampling (description, laboratory samples), i.e., – the required input information was available neither to experts nor for ANN model.

The outcome of this experiment led us to an assumption that quality metrics for automatic classification would largely depend on the data analysis method used by an expert.

Borehole Number	Expert D		Expert L		Expert T		ANN	
	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
2100	0.3182	0.7000	0.2727	0.5455	0.2727	0.5455	0.4516	0.4375
2104	Comparison impossible as data belong to different horizons							
4939	0.4146	0.3366	0.4146	0.3301	0.3049	0.2907	0.1613	0.0060
Average	**0.3664**	**0.5183**	**0.3437**	**0.4378**	**0.2888**	**0.4181**	**0.3064**	**0.2275**

IV. ACCOUNTING METHOD OF EXPERT ASSESSMENTS INCONSISTENCY

During the analysis of well interpretation performed by three different experts, the following hypothesis was formulated: Expert estimates often do not agree with each other, and this is most relevant to the interpretation of all types of rocks. When only permeable rocks are singled out, the experts agree with each other much better. Inconsistency of expert estimations has a strong negative effect on the quality of classifiers' work and it should be considered.

To test this hypothesis, we performed experiments with data from "Budennovskoe" minefield: 57 boreholes were interpreted/assessed by expert C, and 40 boreholes – by expert E.

Data processing includes preprocessing, training and assessment of the quality indicators of the classifier. Feedforward artificial neural networks were used as a classifier.

After linear normalization, logging data were fed to the input of ANN in the form of a "floating window" (Fig. 1).

The use of "floating data windows" is a common method of analysing data sequences, for example, time series or, as in our case, the dependence of recorded physical parameters on the depth. Since the expert takes into account the form of the logging curve when evaluating the data, it is reasonable to submit to the input data in a form of floating window with the size of $n + 1 + n$ points, i.e., n measurements above, the current value and n measurements below the current value are considered. The next window is formed in a similar way, shifting one point below.

Since the logging probe length is 1 meter, which corresponds to 10 depth measurements, it is natural to set n = 5. Presenting data in the form of a floating window allows, to some extent, taking into account the form of the curve, and not just the value of recorded parameter at a specific depth.

The boreholes were divided into training and test sets in the ratio of 80/20 (%) (77/20). It should be noted that for the correct division into training and test sets in this case it was impossible to use sklearn.model_selection.train_test_split or similar functions due to the presence of a floating window.

The network topology was largely determined by the dataset. Since the dataset included data from two logs (AR and SP), the network had 23 inputs (due to the floating window, each type of logging was fed to 11 inputs, also one input was for depth).

The number of outlets was determined by the number of types of rocks allocated – 8. The network had two hidden layers.
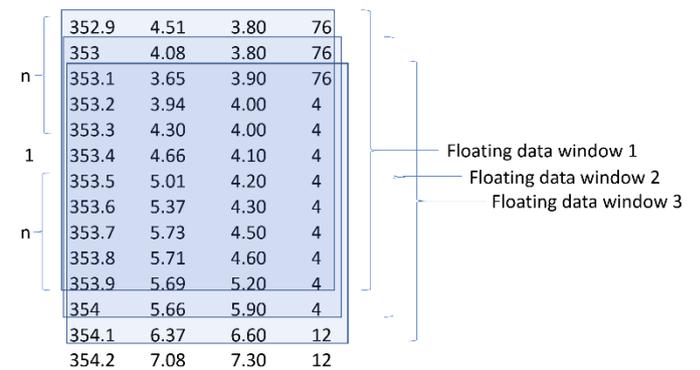

Fig. 1. Floating data window.

To assess the impact of an expert identifier, it is desirable to interpret the constructed machine learning model in order to answer the question "Why did one or another result of classification or regression occur?".

In particular, there are examples of models, when some features of the dataset led to the correct classification results for a specific case, while the model itself was in general completely wrong and the results were based on illogical features from the dataset [30]. That work describes the task of text classification between atheistic and religious content, when it turned out that the model correctly distinguished the former from the latter, but not based on content of the texts, but based on the presence of the word "posting", which occurred only two times in texts of religious content, although it was very common (21 %) in the training set.

In other words, an algorithm is needed that will help answer the question "Which variables and within what limits affect the prediction?" This allows us to make sure that the model does not overfit and that it does not generate a result in a random way.

The ideal interpretation may be, for example, when the general response of the model is the sum of the values of the parameters (X set) multiplied by the model coefficients (Θ set):

$$h_\theta = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + ... + \theta_n x_n, \qquad (1)$$

where $\theta_i \in \Theta$, $x_i \in X$, h_θ – a hypothesis function of a linear regression model.

In this case, it is easy to understand how individual features influence the result. In terms of explaining the result, linear regression models are easily interpretable.

However, this approach is not always possible. In the case of complex non-linear dependencies and significant correlation of properties, the model is a so-called "black box", i.e., a kind of closed system that receives data as input, "secretly" processes them and outputs the result. What is being done inside the "black box", how the model has linked the input data with the result remains incomprehensible.

To turn such a "black box" into a "white" one, i.e., to understand how individual parameters influence model prediction in regression and classification problems, the Treeinterpreter for decision trees, DeepLIFT, similar to Layer-Wise Relevance Propagation and the most recent method LIME (Local Interpretable Model-agnostic) and SHAP (SHapley Additive exPlanations) can be applied.

LIME is described in [31]–[33]. Application of the method is considered in [34], [35]. LIME assumes linear dependences of the classifier's output in a local area.

In turn, SHAP [36], [37] is qualitatively different from the simple search for correlations, in view of the fact that it uses the model for gaining knowledge about nonlinear and non-monotonic interdependencies of parameters that influence the final result. SHAP is widely used to explain the result of machine learning [38]–[42].

SHAP in the form of the shap python library was used by us to assess the influence of the features of the logging data classifier model.

V. RESULTS OF DATASET LABELLING

To assess the impact of expert inconsistency, experiments on training and classification of test data were carried out using an expert identifier (ID) and without using an ID.

At first, the classifier was applied to datasets with wells interpreted by only one expert (C or E). At the same time, on the first dataset the classification accuracy was quite high (accuracy 0.59, F1 0.55). The dataset was then merged into a single one that did not know which well had been interpreted by which expert. This resulted in a significant drop in classification quality (accuracy 0.44, F1 0.20). The use of the interpreter ID as a training parameter resulted in an improved quality of interpretation on this mixed data set (accuracy 0.50, F1 0.42) (Table IV).

Based on our experimental results, we make the following suggestions:

- the accuracy of lithotype identification (accuracy, precision and recall) significantly varies from expert to expert, and this fact should be taken seriously;
- the input data from each individual expert provide a much better ANN output (accuracy of lithotype identification), when compared with a mixed dataset of boreholes (manual classification, multiple experts);

- introduction of parameter 'expert ID' for ML training increases overall quality of classification, especially for recall values.

To test our hypothesis that the experts agree well on the allocation of impermeable rocks, and that each interpreter introduces a large share of subjectivity in the separation of permeable rocks, we analysed the classifier's work using SHAP in two cases: identifying all classes of rocks and identifying only impermeable rocks.

The results are shown in Figs. 2 and 3. They demonstrate the importance of features in descending order.

AR [x] and SP [x] here denote the values of recorded physical parameters of AR and SP within the boundaries of the floating window [−5; 5]. It can be seen that when identifying all classes of rocks, the most significant parameter is the interpreter ID, which reflects his subjective "style" and this must be taken into account when training the network. At the same time, when identifying only impermeable rocks, this parameter is not essential.

TABLE IV
RESULTS OF ANN APPLICATION TO ASSESSMENTS BY TWO EXPERTS

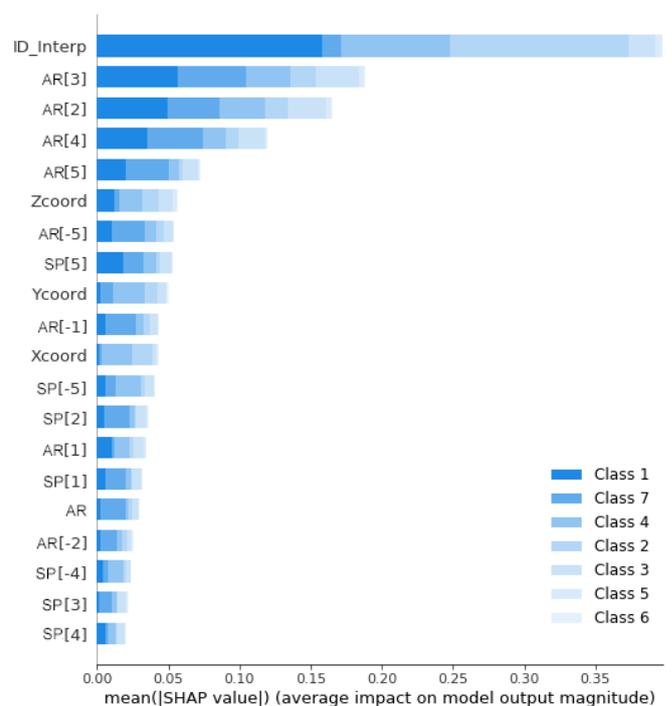| Methods	Train	Test	Accuracy	Precision	Recall	F1
Expert C	45	12	0.5869	0.6491	0.4751	0.5486
Expert E	35	5	0.4722	0.5409	0.1963	0.2880
Experts C&E w/out ID	77	20	0.4401	0.4078	0.1365	0.2045
Experts C&E with ID	77	20	0.4996	0.5157	0.3529	0.4190

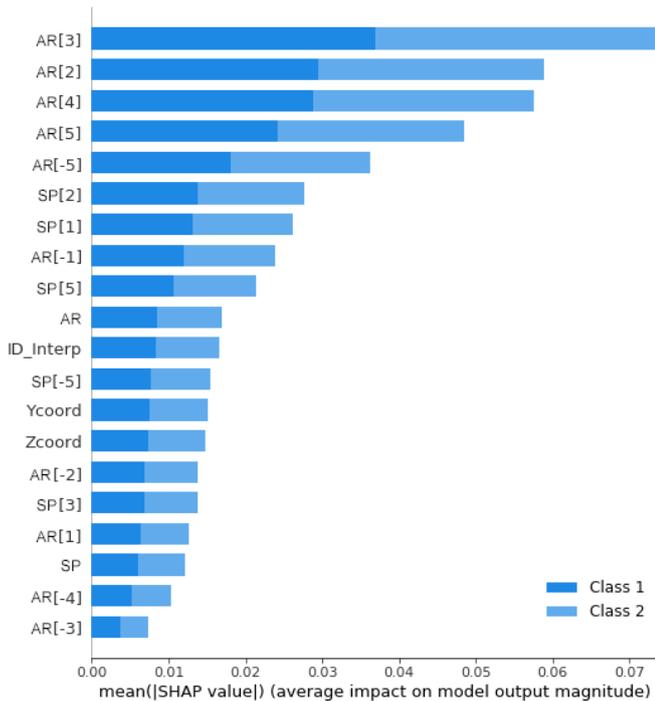Fig. 2. Comparative importance of parameters when identifying all classes of rocks.

Fig. 3. Comparative importance of parameters when identifying only impermeable rocks.

VI. CONCLUSION

Accurate interpretation of electric log data is vital for uranium production needs, specifically for selecting filter installation location when using the method of uranium extraction via sub-surface in-situ leaching of boreholes. Until now, such interpretation has been performed mostly manually.

The use of automatic classifiers in this case promises great advantages. However, the preparation of the initial data and especially the training data set is not trivial for this task. The reason is that the machine learning algorithm can be trained almost exclusively on data interpreted by experts. In turn, expert classification can hardly be verified. It is intuitively clear that in this case the quality of automatic classification depends on how much the experts agree with each other.

During the analysis of well interpretation carried out by three different experts, it was concluded that their estimates often did not agree with each other, and this was most relevant for the interpretation of all types of rocks. When only impermeable rocks are identified, the experts are much better aligned with each other. Labelling of data (introduction of the interpreter ID as a learning parameter) allowed improving the quality of classification (by 5 %) and recall (by 20 %).

The use of SHAP on a mixed data set confirmed that when selecting eight rock classes, the most important parameter was the interpreter ID, while when selecting only two rock classes (permeable/impermeable) its role was insignificant.

ACKNOWLEDGMENT

REFERENCES

[1] NAC Kazatomprom JSC, "Kazatomprom," 2020. [Online]. Available: https://www.kazatomprom.kz/en [Accessed: Mar. 02, 2020].

[2] GRK LLP, "Technical instruction for geophysical survey in wells at reservoir infiltration deposits of uranium," unpublished.

[3] E. N. Amirgaliev, S. Kh. Iskakov, Ya. I. Kuchin, and R. I. Muhamediev, "Machine learning methods for rock recognition problems in uranium deposits," in *Proc. of the National Academy of Sciences of Kazakhstan* 3, 2013, pp. 82–88.

[4] N. Giang, *et al.* "Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey," *Artificial Intelligence Review* 52.1, pp. 77–124, 2019. https://doi.org/10.1007/s10462-018-09679-z

[5] R. Muhamedyev, "Machine learning methods: An overview," *CMNT* 19, no. 6, pp. 14–29, 2015.

[6] M. van der Baan, *et al.* "Neural networks in geophysical applications," *Geophysics* 65(4), pp. 1032–1047, 2000. https://doi.org/10.1190/1.1444797

[7] J. L. Baldwin, R. M. Bateman, and C. L. Wheatley, "Application of a neural network to the problem of mineral identification from well logs," *The Log Analyst,* vol. 3, pp. 279–293, 1990.

[8] B. Benaouda, G. Wadge, R. B. Whitmarsh, R. G. Rothwell, and C. MacLeod, "Inferring the lithology of borehole rocks by applying neural network classifiers to downhole logs - an example from the Ocean Drilling Program." *Geophysical Journal International,* vol. 136, no. 2, pp. 477–491, 1999. https://doi.org/10.1046/j.1365-246X.1999.00746.x

[9] M. M. Saggaf, and Ed. L. Nebrija, "Estimation of missing logs by regularized neural networks." *AAPG Bulletin,* vol. 87, no. 8, pp. 1377–1389, 2003. https://doi.org/10.1306/03110301030

[10] V. A. Tenenev, B. A. Yakimovich, M. A. Senilov, and N. B. Paklin, "Intellectual systems for interpretation of well logging," *Shtnyi intelekt* vol. 3, p. 338, 2002.

[11] Y. Klaus, and T. Sven, "Computational Neural Networks for Geophysical Data Processing," Elsevier Science, 2001.

[12] M. Borsaru, B. Zhou, T. Aizawa, H. Karashima, and T. Hashimoto, "Automated lithology prediction from PGNAA and other geophysical logs," *Applied Radiation and Isotopes,* vol. 64, no. 2, pp. 272–282, 2006. https://doi.org/10.1016/j.apradiso.2005.07.012

[13] S. J. Rogers, H. C. Chen, D. C. Kopaska-Merkel, and J. H. Fang, "Predicting permeability from porosity using artificial neural networks," *AAPG Bulletin*, vol. 12, no. 12, pp. 1786–1797, 1995. https://doi.org/10.1306/7834DEFE-1721-11D7-8645000102C1865D

[14] L. Kapur, L. Lake, K. Sepehrnoori, D. Herrick, and C. Kalkomey, "Facies prediction from core and log data using artificial neural network technology," in *39th Society of Professional Well Log Analysts Annual Logging Symposium*, 1998.

[15] S. P. Aleshin, A. L. Lyakhov, "Neural network assessment of the mineral resource base of a region according to geophysical monitoring data," *New technologies,* vol. 1, no. 31, pp. 39–43, 2001.

[16] S. J. Rogers, J. H. Fang, C. L. Karr, D. A. Stanley, "Determination of lithology from well logs using a neural network," *AAPG Bulletin,* vol. 76, no. 5, pp. 731–739, 1992. https://doi.org/10.1306/BDFF88BC-1718-11D7-8645000102C1865D

[17] D. V. Kostikov, "Instrumental tools for interpretation of well logging based on converted logging data using a multilayer neural network," Ph.D. dissertation, p. 189, 2007.

[18] R. Muhamediyev, E. Amirgaliev, S. Iskakov, Y. Kuchin, and E. Muhamedyeva, "Integration of Results of Recognition Algorithms at the Uranium Deposits," *Journal of ACIII,* vol. 18, no. 3, pp. 347–352, 2014.

[19] E. N. Amirgaliev, S. Kh. Iskakov, Ya. I. Kuchin, and R. I. Muhamediev, "Integration of recognition algorithms of lithological types," *Informatics problems. Siberian Branch of the Russian Academy of Sciences* vol. 4, no. 21, pp. 11–20, 2013.

[20] E. N. Amirgaliev, S. Kh. Iskakov, Ya. I. Kuchin, R. I. Muhamediev, "Machine learning methods for rock recognition problems in uranium deposits," in *Proc. of the National Academy of Sciences of Kazakhstan 3*, 2013, pp. 82–88.

[21] "Development of methods of data boreholes interpretation by using artificial neural network (On request of Geotehnoserviss ltd)," unpublished.

[22] Y. I. Kuchin, R. I. Muhamedyev, E. L. Muhamedyeva, P. Gricenko, Zh. Nurushev, and K. Yakunin, "The analysis of the data of geophysical research of boreholes by means of artificial neural networks," *Computer Modelling and New Technologies,* vol. 15, no. 4, pp. 35–40, 2011.

[23] R. I. Muhamedyev, Y. I. Kuchin, and E. L. Muhamedyeva, "Geophysical research of boreholes: Artificial neural networks data analysis," in *IEEE 6th International Conference on Soft Computing and Intelligent Systems, and the 13th International Symposium on Advanced Intelligence Systems*, 2012, pp. 825–829. https://doi.org/10.1109/SCIS-ISIS.2012.6505183

[24] E. Amirgaliev, Z. Isabaev, S. Iskakov, Y. Kuchin, R. Muhamediyev, E. Muhamedyeva, K. Yakunin, "Recognition of rocks at uranium deposits by using a few methods of machine learning," *Soft Computing in Machine Learning*, pp. 33–40, 2014. https://doi.org/10.1007/978-3-319-05533-6_4

[25] R. I. Muhamediyev, *et al.*, "Comparative analysis of classification algorithms," in *IEEE 9th International Conference on Application of Information and Communication Technologies (AICT)*, 2015, pp. 96–101. https://doi.org/10.1109/ICAICT.2015.7338525

[26] R. Muhamediyev, E. Amirgaliev, S. Iskakov, Y. Kuchin, and E. Muhamedyeva, "Integration of Results of Recognition Algorithms at the Uranium Deposits," *JACIII,* vol. 8, no. 3, pp. 347–352, 2014.

[27] R. Muhamediyev, S. Iskakov, P. Gricenko, K. Yakunin, and Y. Kuchin, "Integration of results from Recognition Algorithms and its realization at the uranium production process," in *8th IEEE International Conference AICT*, 2014, pp. 188–191.

[28] Y. Kuchin, R. Mukhamediev, and K. Yakunin, "One method of generating synthetic data to assess the upper limit of machine learning algorithms performance," *Cogent Engineering*, p. 1718821, 2020. https://doi.org/10.1080/23311916.2020.1718821

[29] Keras Team, "Keras: Deep Learning for humans," [Online]. Available: https://github.com/keras-team/keras [Accessed: Mar. 02, 2020].

[30] M. Ribeiro, S. Singh, and C. Guestrin, "Local Interpretable Model-Agnostic Explanations (LIME): An Introduction A technique to explain the predictions of any machine learning classifier," 2019.

[31] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should i trust you? Explaining the predictions of any classifier," in *Proc. of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135–1144.

[32] L. Hulstaert, "Understanding Model Predictions with LIME. Learn about Lime and how it works along with the potential pitfalls that come with using it," [Online]. Available: https://www.datacamp.com/community/tutorials/understanding-model-predictions-lime [Accessed: Mar. 02, 2020].

[33] M. Ribeiro, "Lime," [Online]. Available: https://github.com/marcotcr/lime [Accessed: Mar. 02, 2020].

[34] M. Ribeiro, S. Singh, and C. Guestrin, "Local Interpretable Model-Agnostic Explanations (LIME): An Introduction," [Online]. Available: https://www.oreilly.com/learning/introduction-to-local-interpretable-model-agnostic-explanations-lime [Accessed: Mar. 02, 2020].

[35] W. Koehrsen, "A Complete Machine Learning Walk-Through in Python: Part Three Interpreting a machine learning model and presenting results," [Online]. Available: https://towardsdatascience.com/a-complete-machine-learning-walk-through-in-python-part-three-388834e8804b [Accessed: Mar. 02, 2020].

[36] S. M. Lundberg, S. I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, pp. 4765–4774, 2017.

[37] S. Lundberg, "SHAP (SHapley Additive exPlanations)," [Online]. Available: https://github.com/slundberg/shap [Accessed: Mar. 02, 2020].

[38] Mangalathu, Sujith, Seong-Hoon Hwang, and Jong-Su Jeon. "Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach," *Engineering Structures* 219, pp. 110927, 2020. https://doi.org/10.1016/j.engstruct.2020.110927

[39] A. B. Parsa, *et al.* "Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis," *Accident Analysis & Prevention,* vol. 136, pp. 105405, 2020. https://doi.org/10.1016/j.aap.2019.105405

[40] R. Muhamedyev, *et al.* "The use of machine learning "black boxes" explanation systems to improve the quality of school education." *Cogent Engineering,* vol. 7.1, pp. 1769349, 2020. https://doi.org/10.1080/23311916.2020.1769349

[41] García, María Vega, and José L. Aznarte. "Shapley additive explanations for NO2 forecasting," *Ecological Informatics* vol. 56, pp. 101039, 2020. https://doi.org/10.1016/j.ecoinf.2019.101039

[42] Rodríguez-Pérez, Raquel, and Jürgen Bajorath. "Interpretation of compound activity predictions from complex machine learning models using local approximations and shapley values," *Journal of Medicinal Chemistry* vol. 63, no. 16, pp. 8761−8777, 2020. https://doi.org/10.1021/acs.jmedchem.9b01101

Yan Kuchin was born in 1980 in Almaty, Kazakhstan. He received a bachelor's degree in physics in 2002 from the Kazakh National University, Almaty, Kazakhstan, a bachelor's degree in computer science in 2011 from IITU, Almaty, Kazakhstan and a master's degree in computer science in 2016. Currently he is a PhD student in Riga Technical University. From 2004 to 2018, he worked at Kazatomprom. From 2018, he is a programmer-engineer at the Institute of Information and Computing Technologies and researcher at Satbayev University. He is the author of more than 30 articles. His research interests include geophysics, machine learning, data processing, natural language processing, etc.
E-mail: ykuchin@mail.ru
ORCID iD: https://orcid.org/0000-0002-5271-9071

Ravil Mukhamediev was born in 1959 in Irkutsk, Russia. He received an engineering degree in radio electronics from the Riga Institute of Civil Aviation Engineers, Riga, Latvia, in 1983, and a PhD Degree in Information Systems from RCAEI in 1987. From 1987 to 1995 he worked at RCAII as an assistant, lecturer, assistant professor. From 2004 to 2010, he was Associate Professor and Head of the Department of Information Systems Management, Riga, Latvia. Since 2010, he has been a professor at various high schools in Almaty, Kazakhstan. Currently holds the position of professor of the Kazakh National Research Technical University. He is the author of three books, more than 200 articles and a leading researcher of four research projects. His research interests include machine learning, data processing, decision support systems, etc. He is a member of the editorial board of CMNT and CoS journals.
E-mail: ravil.muhamedyev@gmail.com
ORCID iD: https://orcid.org/0000-0002-3727-043X

Kirill Yakunin was born in 1992 in Almaty, Kazakhstan. He received a bachelor's degree in computer science in 2014 from the International University of Information Technology, Almaty, Kazakhstan, a master's degree in computer science and software development in 2016 from the International University of Information Technology, Almaty, Kazakhstan. Since 2015, she has been working at Epigraph OJSC as a lead software engineer. From 2018, he is a programmer-engineer at the Institute of Information and Computing Technologies. He is the author of two books, more than 20 articles. His research interests include optimization, machine learning, data processing, natural language processing, etc.
E-mail: yakunin.k@mail.ru
ORCID iD: https://orcid.org/0000-0002-7378-9212

Janis Grundspenkis received the qualification of Electrical Engineer in automation and telemechanics in 1965 and the degree of Candidate of Technical Sciences in 1972 (recognised in 1992 as the degree of Doctor of Engineering Sciences) from Riga Polytechnical Institute, Latvia. In 1993, he received the degree of Habilitated Doctor of Engineering Sciences from Riga Technical University, Latvia.
Since 1994, he has been Professor of Systems Theory at Riga Technical University. He also holds the positions of the Dean of the Faculty of Computer Science and Information Technology and the Head of the Department of Artificial Intelligence and Systems Engineering of Riga Technical University, both since 1994. His research interests include agent based and multiagent intelligent systems, knowledge acquisition and representation, causal domain models for complex cyber-physical systems, and structural modelling.
He is a full member of the Latvian Academy of Sciences, senior member of IEEE, and member of ACM.
E-mail: janis.grundspenkis@rtu.lv
ORCID iD: https://orcid.org/0000-0003-2526-4662

Adilkhan Symagulov was born in 1994 in Almaty, Kazakhstan. He received the bachelor degree in natural science in 2018 from Kazakh National Research University, Almaty, Kazakhstan. From 2018, he is a Programmer-Engineer in the Institute of Information and Computational Technologies. He is the author of 5 articles. His research interests include machine learning, expert systems, optimization, and distributed systems. His research interests include machine learning and natural language processing.

E-mail: asmogulove00@gmail.com

ORCID iD: https://orcid.org/0000-0001-9974-3215