ISSN 2518-170X (Online)
ISSN 2224-5278 (Print)

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN, SERIES OF GEOLOGY AND TECHNICAL SCIENCES

№5
2025

ISSN 2518-170X (Online)
ISSN 2224-5278 (Print)

CENTRAL ASIAN ACADEMIC
RESEARCH CENTER

N E W S
OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN, SERIES OF GEOLOGY AND TECHNICAL SCIENCES

5 (473)

SEPTEMBER – OCTOBER 2025

THE JOURNAL WAS FOUNDED IN 1940

PUBLISHED 6 TIMES A YEAR

ALMATY, 2025

«Central Asian Academic Research Center» LLP is pleased to announce that "News of NAS RK. Series of Geology and Technical sciences" scientific journal has been accepted for indexing in the Emerging Sources Citation Index, a new edition of Web of Science. Content in this index is under consideration by Clarivate Analytics to be accepted in the Science Citation Index Expanded, the Social Sciences Citation Index, and the Arts & Humanities Citation Index. The quality and depth of content Web of Science offers to researchers, authors, publishers, and institutions sets it apart from other research databases. The inclusion of News of NAS RK. Series of Geology and Technical Sciences in the Emerging Sources Citation Index demonstrates our dedication to providing the most relevant and influential content of geology and engineering sciences to our community.

«Орталық Азия академиялық ғылыми орталығы» ЖШС «ҚР ҰҒА Хабарлары. Геология және техникалық ғылымдар сериясы» ғылыми журналының Web of Science-тің жаңаланған нұсқасы Emerging Sources Citation Index-те индекстелуге қабылданғанын хабарлайды. Бұл индекстелу барысында Clarivate Analytics компаниясы журналды одан әрі the Science Citation Index Expanded, the Social Sciences Citation Index және the Arts & Humanities Citation Index-ке қабылдау мәселесін қарастыруда. Web of Science зерттеушілер, авторлар, баспашылар мен мекемелерге контент тереңдігі мен сапасын ұсынады. ҚР ҰҒА Хабарлары. Геология және техникалық ғылымдар сериясы Emerging Sources Citation Index-ке енуі біздің қоғамдастық үшін ең өзекті және беделді геология және техникалық ғылымдар бойынша контентке адалдығымызды білдіреді.

ТОО «Центрально-азиатский академический научный центр» сообщает, что научный журнал "Известия НАН РК. Серия геологии и технических наук» был принят для индексирования в Emerging Sources Citation Index, обновленной версии Web of Science. Содержание в этом индексировании находится в стадии рассмотрения компанией Clarivate Analytics для дальнейшего принятия журнала в the Science Citation Index Expanded, the Social Sciences Citation Index и the Arts & Humanities Citation Index. Web of Science предлагает качество и глубину контента для исследователей, авторов, издателей и учреждений. Включение Известия НАН РК. Серия геологии и технических наук в Emerging Sources Citation Index демонстрирует нашу приверженность к наиболее актуальному и влиятельному контенту по геологии и техническим наукам для нашего сообщества.

News of the National Academy of Sciences of the Republic of Kazakhstan. Series of geology and technology sciences.
ISSN 2518-170X (Online),
ISSN 2224-5278 (Print)
Owner: «Central Asian Academic Research Center» LLP (Almaty).
The certificate of registration of a periodical printed publication in the Committee of information of the Ministry of Information and Social Development of the Republic of Kazakhstan **No. KZ39VPY00025420,** issued 29.07.2020.
Thematic scope: *geology, hydrogeology, geography, mining and chemical technologies of oil, gas and metals*
Periodicity: 6 times a year.
http://www.geolog-technical.kz/index.php/en/

«ҚР ҰҒА» РҚБ Хабарлары. Геология және техникалық ғылымдар сериясы».
ISSN 2518-170X (Online),
ISSN 2224-5278 (Print)

Меншіктеуші: «Орталық Азия академиялық ғылыми орталығы» ЖШС (Алматы қ.).

Қазақстан Республикасының Ақпарат және қоғамдық даму министрлігінің Ақпарат комитетінде 29.07.2020 ж. берілген **№ KZ39VPY00025420** мерзімдік басылым тіркеуіне қойылу туралы куәлік.

Тақырыптық бағыты: *Геология, гидрогеология, география, тау-кен ісі, мұнай, газ және металдардың химиялық технологиялары*

Мерзімділігі: жылына 6 рет.

http://www.geolog-technical.kz/index.php/en/

«Известия РОО «НАН РК». Серия геологии и технических наук».
ISSN 2518-170X (Online),
ISSN 2224-5278 (Print)
Собственник: ТОО «Центрально-азиатский академический научный центр» (г. Алматы).
Свидетельство о постановке на учет периодического печатного издания в Комитете информации Министерства информации и общественного развития Республики Казахстан **№ KZ39VPY00025420,** выданное 29.07.2020 г.
Тематическая направленность: *геология, гидрогеология, география, горное дело и химические технологии нефти, газа и металлов*
Периодичность: 6 раз в год.
http://www.geolog-technical.kz/index.php/en/

CONTENTS

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC
OF KAZAKHSTAN, SERIES OF GEOLOGY AND TECHNICAL SCIENCES
ISSN 2224–5278
Volume 5. Number 473 (2025), 146–157

https://doi.org/10.32014/2025.2518-170X.555

UDC 552.5:519.87

© L.V. Krasovskaya[1], V.S. Tynchenko[2*], O.A. Antamoshkin[3], S.V. Pchelintseva[1], M.S. Nikanorov[1], 2025.
[1]RSAU-MAA Named after K.A. Timiryazev, Moscow, Russia;
[2]Bauman Moscow State Technical University, Moscow, Russia;
[3]Siberian Federal University, Krasnoyarsk, Russia.
E-mail: vadimond@mail.ru

L.V. Krasovskaya — Cand. Tech. Sc., Associate Professor, Russian State Agrarian University - Moscow Timiryazev Agricultural Academy, Moscow, Russia,
E-mail: kraslud@yandex.ru, ORCID: 0000-0002-9674-8384;
Vadim S. Tynchenko — Doctor of Technical Sciences, Bauman Moscow State Technical University, Moscow, Russia,
E-mail: vadimond@mail.ru, ORCID: 0000-0002-3959-2969;
Oleslav A. Antamoshkin — Doctor of Technical Sciences, Siberian Federal University, Krasnoyarsk, Russia; Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, Russia,
E-mail: oleslav@mail.ru, ORCID: 0000-0002-5976-5847;
Svetlana V. Pchelintseva — Cand. Tech. Sc., Associate Professor, Russian State Agrarian University - Moscow Timiryazev Agricultural Academy, Moscow, Russia,
E-mail: pchelintseva@inbox.ru, ORCID: https://orcid.org/0000-0003-0230-780X;
Mikhail S. Nikanorov - Senior Lecturer, Russian State Agrarian University - Moscow Timiryazev Agricultural Academy, Moscow, Russia,
E-mail: nikanorov@rgau-msha.ru, ORCID: 0000-0001-7487-6273.

APPLICATION OF MACHINE LEARNING METHODS AS A MODERN APPROACH TO ROCK ANALYSIS

Abstract. *Relevance.* The classification of geological rocks traditionally requires significant time, expert knowledge, and manual analysis of chemical composition data. Automating this process through machine learning methods can substantially improve efficiency and accuracy in geological and mining studies. *Objective.* The study aims to evaluate the effectiveness of the Random Forest algorithm for classifying geological rocks based on their chemical composition and to assess its potential for improving the accuracy and automation of rock type identification. *Methods.* A dataset containing concentrations of major oxides (SiO_2, K_2O, FeO, TiO_2, CaO, MgO) in samples of different rock types (basalt, basaltic andesite, granite, quartz diorite, rhyolite) was used. Data preprocessing, feature correlation analysis, and division into training and test subsets were performed. The Random

Forest model's performance was evaluated using precision, recall, and F1-measure metrics. *Results and Conclusions.* The Random Forest model demonstrated an overall classification accuracy of 84%, indicating its strong potential for geological applications. Correlation analysis revealed clear relationships between specific oxide concentrations and rock types. However, reduced recall values for certain classes highlighted limitations in distinguishing similar rock types. The findings confirm the applicability of machine learning for geological rock classification and suggest that further improvements – such as hyperparameter tuning, expansion of training datasets, and adoption of advanced ensemble techniques – can enhance model performance and reliability.

Keywords: classification of geological rocks, chemical composition of rocks, machine learning, random forest, correlation analysis

© Л.В. Красовская[1], В.С. Тынченко[2*], О.А. Антамошкин[3], С.В. Пчелинцева[1], М.С. Никаноров[1], 2025.

[1] К.А. Тимирязев атындағы Ресей мемлекеттік аграрлық университеті – АШМ, Мәскеу, Ресей;

[2] Н.Е. Бауман атындағы МТУ, Мәскеу, Ресей;

[3]Сібір федералды университеті, Красноярск, Ресей.

E-mail: vadimond@mail.ru

ТАУ ЖЫНЫСТАРЫН ТАЛДАУДЫҢ ЗАМАНАУИ ТӘСІЛІ РЕТІНДЕ МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІН ҚОЛДАНУ

Людмила Владимировна Красовская — т.ғ.к., доцент, К.А. Тимирязев атындағы Ресей мемлекеттік аграрлық университеті – АШМ, Мәскеу, Ресей,
E-mail: kraslud@yandex.ru, ORCID: 0000-0002-9674-8384;
Вадим Сергеевич Тынченко — техника ғылымдарының докторы, Н.Е. Бауман атындағы МТУ, Мәскеу, Ресей; Сібір федералды университеті, Красноярск, Ресей,
E-mail: vadimond@mail.ru, ORCID: 0000-0002-3959-2969;
Олеслав А. Антамошкин — техника ғылымдарының докторы, доцент, Сібір федералды университеті, Красноярск, Ресей; М.Ф. Решетнева атындағы Сібір мемлекеттік Ғылым және технологиялар университеті, Красноярск, Ресей,
E-mail: oleslav@mail.ru, ORCID: 0000-0002-5976-5847;
Светлана Вячеславовна Пчелинцева — т.ғ.к., доцент, К.А. Тимирязев атындағы Ресей мемлекеттік аграрлық университеті – АШМ, Мәскеу, Ресей,
E-mail: pchelintseva@inbox.ru, ORCID: https://orcid.org/0000-0003-0230-780X;
Михаил Сергеевич Никаноров — аға оқытушы, К.А. Тимирязев атындағы Ресей мемлекеттік аграрлық университеті – АШМ, Мәскеу, Ресей,
E-mail: nikanorov@rgau-msha.ru, ORCID: 0000-0001-7487-6273.

Аннотация. *Өзектілігі.* Геологиялық жыныстардың жіктелуі дәстүрлі түрде айтарлықтай уақыт пен сараптамалық ресурстарды қажет етеді, өйткені ол химиялық құрамды көп уақытты қажет ететін талдауға негізделген. Машиналық оқыту әдістерін қолдану бұл процесті автоматтандыруға және оның дәлдігін арттыруға мүмкіндік береді, бұл әсіресе геология мен тау-кен жұмыстарына қатысты. *Мақсат.* Геологиялық жыныстарды химиялық құрамы бойынша жіктеу үшін кездейсоқ орман әдісін (Random Forest) қолданудың тиімділігін бағалау және осы процестің дәлдігі мен автоматтандыру мүмкіндіктерін анықтау. *Әдістері.* Зерттеу әртүрлі тау жыныстарының үлгілерінде (базальт, базальт андезиті, гранит, кварц диориті, риолит) негізгі оксидтердің (SiO_2, K_2O, FeO, TiO_2, CaO, MgO) концентрациялары бар деректер жинағын пайдаланды. Деректер алдын ала өңделіп, оқыту және сынақ үлгілеріне бөлінді. Модельдің өнімділігі дәлдік (дәлдік), толықтық (recall) және F1 өлшемдері бойынша бағаланды. Сонымен қатар, оксидтердің құрамы мен тау жыныстарының арасындағы байланысты анықтау үшін корреляциялық талдау жасалады. *Нәтижелер мен қорытындылар.* Кездейсоқ орман моделі 84% жіктеудің жалпы дәлдігін қамтамасыз ете отырып, жақсы тиімділік көрсетті. Талдау жеке оксидтердің концентрациясы мен тау жынысының түрі арасында тұрақты корреляцияның болуын көрсетті. Сонымен қатар, толықтығы төмен сипатталатын тұқымдардың кейбір кластарын жіктеудегі қиындықтар анықталды. Алынған нәтижелер геологиялық жыныстарды жіктеу міндеттері үшін кездейсоқ орман әдісін қолдану перспективасын растайды және гиперпараметрлерді одан әрі оңтайландыру, оқыту деректерінің көлемін ұлғайту және жыныстардың жекелеген түрлерін тану дәлдігін арттыру үшін ансамбльдің неғұрлым жетілдірілген тәсілдерін енгізу қажеттілігін көрсетеді.

Түйін сөздер: геологиялық жыныстардың жіктелуі, тау жыныстарының химиялық құрамы, Машиналық оқыту, кездейсоқ орман, корреляциялық талдау

© Л.В. Красовская[1], В.С. Тынченко[2]*, О.А. Антамошкин[3], С.В. Пчелинцева[1], М.С. Никаноров[1], 2025.

[1]Российский государственный аграрный университет – МСХА имени К.А. Тимирязева, Москва, Россия;
[2]МГТУ им. Н.Э. Баумана, Москва, Россия;
[3]Сибирский федеральный университет, Красноярск, Россия.
E-mail: vadimond@mail.ru

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ КАК СОВРЕМЕННЫЙ ПОДХОД К АНАЛИЗУ ГОРНЫХ ПОРОД

Людмила Владимировна Красовская — к.т.н., доцент, Российский государственный аграрный университет – МСХА имени К.А. Тимирязева, Москва, Россия,
E-mail: kraslud@yandex.ru, ORCID: 0000-0002-9674-8384;
Вадим Сергеевич Тынченко — д.т.н., г.н.с, МГТУ им. Н.Э. Баумана, Москва, Россия,
E-mail: vadimond@mail.ru, ORCID: 0000-0002-3959-2969;
Олеслав Александрович Антамошкин — доктор технических наук, доцент, Сибирский федеральный университет, Красноярск, Россия,
E-mail: oleslav@mail.ru, ORCID: 0000-0002-5976-5847;

Светлана Вячеславовна Пчелинцева — к.т.н., доцент, Российский государственный аграрный университет – МСХА имени К.А. Тимирязева, Москва, Россия,
E-mail: pchelintseva@inbox.ru, ORCID: https://orcid.org/0000-0003-0230-780X;
Михаил Сергеевич Никаноров — старший преподаватель, Российский государственный аграрный университет – МСХА имени К.А. Тимирязева, Москва, Россия,
E-mail: nikanorov@rgau-msha.ru, ORCID: 0000-0001-7487-6273.

Аннотация: *Актуальность.* Классификация геологических пород традиционно требует значительных временных и экспертных ресурсов, поскольку основывается на трудоёмком анализе химического состава. Применение методов машинного обучения позволяет автоматизировать данный процесс и повысить его точность, что особенно актуально для геологии и горного дела. *Цель***.** Оценить эффективность метода случайного леса (Random Forest) при классификации геологических пород по химическому составу и определить пути повышения точности и автоматизации этого процесса. *Методы.* В исследовании использован набор данных, содержащий концентрации основных оксидов (SiO_2, K_2O, FeO, TiO_2, CaO, MgO) в образцах различных типов пород (базальт, базальтовый андезит, гранит, кварцевый диорит, риолит). Данные были предварительно обработаны и разделены на обучающую и тестовую выборки. Производительность модели оценивалась по метрикам точности (precision), полноты (recall) и F1-меры. Дополнительно выполнен корреляционный анализ для выявления взаимосвязей между содержанием оксидов и типом породы. *Результаты и выводы.* Модель случайного леса продемонстрировала хорошую эффективность, обеспечив общую точность классификации 84 %. Анализ показал наличие устойчивых корреляций между концентрациями отдельных оксидов и типом породы. При этом выявлены сложности в классификации некоторых классов пород, характеризующихся низким значением полноты. Полученные результаты подтверждают перспективность применения метода случайного леса для задач классификации геологических пород и свидетельствуют о необходимости дальнейшей оптимизации гиперпараметров, увеличения объёма обучающих данных и внедрения более продвинутых ансамблевых подходов для повышения точности распознавания отдельных типов пород.

Ключевые слова: классификация геологических пород, химический состав пород, машинное обучение, случайный лес, корреляционный анализ

Introduction. Geology plays a key role in understanding the composition and structure of the Earth. One of the important tasks in this science is the classification of types of geological rocks, which has wide application in various fields such as mining, construction, ecology and many others (Epikhin et al., 2024; Bosikov et al., 2024). Traditional classification methods require considerable time and effort, as well as deep knowledge of experts in the field of geology. In recent years, with the development of technology and the advent of large amounts of data, machine learning methods have become an important tool for automating and improving classification processes (Olentsova et al., 2024).

Machine learning makes it possible to effectively analyze large datasets, identify hidden dependencies and patterns, and predict results with high accuracy. In this paper, we consider a random forest machine learning model that is used to classify types of geological rocks based on chemical composition.

The purpose of this work is to analyze the effectiveness of the random forest method for classifying geological rocks.

Research methods and materials. To conduct this study, we used data containing the chemical composition of various types of geological rocks. The data set included concentrations of major oxides such as SiO, KO, FeO, TiO, CaO, and MgO. Various types of rocks were used as target markers, including basalt, basalt andesite, granite, quartz diorite (Bosikov et al., 2023), and rhyolite. Each sample was a feature vector reflecting the concentrations of chemical components that were used for classification.

The data was preprocessed before building the models. The data was scaled using a standard normalization method to bring the feature values to a single scale. The data set was divided into training and test samples in a ratio of 80% to 20%. To classify the types of rocks, a random forest model was used, which is an ensemble method that builds a set of decision trees. Each classifier was trained on a random subset of data and features, which increased the model's resistance to overfitting (Borodulin et al., 2023). 100 trees with a maximum tree depth set at 15 levels were used as key hyperparameters of a random forest. To split the nodes of the tree, the "gini" criterion was used, which measures the purity of the nodes.

The model was evaluated using accuracy, completeness, and F1-measure metrics. Accuracy indicated the proportion of correct positive predictions, completeness indicated the proportion of correctly recognized positive examples, and the F1 measure combined these two metrics to provide a balanced assessment of model performance (Josephine, Khan, Perras, 2019). Visualization of the results included an analysis of the error matrix and Precision-Recall curves, which made it possible to evaluate the model's performance on a test sample. In addition, correlation analysis was performed using Pearson coefficients to study the relationships between chemical components and rock types. This analysis revealed linear dependencies, which contributed to a deeper interpretation of the classification results (et al., 2024).

Machine learning methods are widely used in rock analysis, making it possible to improve the accuracy of their classification and prediction of mechanical properties. Hamada and others are considering advanced machine learning methods for classifying carbonate rocks. The authors use hyperparameter optimization strategies such as Bayesian optimization and SMOTE, which improves classification accuracy by up to 99.2%. This study demonstrates the effectiveness of ensemble methods and deep learning for geological tasks.

Morgenroth (2019) presents a roadmap for the introduction of machine learning in mining (Tynchenko, Kukartsev, Gladkov, Panfilova, 2024). The main focus is on the problems of implementing machine learning models, including lack of data, model bias, and difficulties in interpreting results. The paper proposes strategies for standardizing machine learning processes, which contributes to their wider application in geoengineering.

Zhu and others are exploring the fusion of finite elements and machine learning to predict rock strength parameters (Hamada et al., 2024). Using the method of supporting vectors, gradient boosting, and neural networks allows for high prediction accuracy. The results of the study confirm the prospects of integrating numerical methods and machine learning for the mechanical analysis of rocks.

Zemlyanoy (2024) considers the use of deep learning to determine the properties of rocks (Zemlyanoy et al., 2024). Convolutional neural networks are used to predict the porosity and permeability of samples, which significantly speeds up the analysis process. This study demonstrates the potential of neural networks in processing geological survey data.

Wang and others propose a group machine learning method for predicting the parameters of rock formations during drilling (Morgenroth, 2019). The use of Fuzzy C-Means clustering in conjunction with machine learning methods makes it possible to improve the classification accuracy of rock arrays, which is especially important for planning tunneling and drilling operations.

Lee and others are analyzing the use of deep learning to accurately identify the lithological features of sandstones (Zhu et al., 2024). The authors have developed the Rock-ViT method based on visual transformers, which demonstrates high accuracy in difficult conditions. This study highlights the promise of modern image processing techniques in geology. Grant and others propose a new methodology for using spatial metadata in machine learning to analyze geological cores. The integration of spatial information makes it possible to increase classification accuracy by up to 25%, which makes this approach promising for automated analysis of core samples.

Rudianto and others are conducting a comparative study of SVM and random forest algorithms for classifying rock types. The results show that the random forest demonstrates higher accuracy, confirming its effectiveness for geological tasks. In this paper, the authors consider the classification of rock drillability using machine learning (Koteleva et al., 2024; Morgoeva et al., 2024). The developed models based on machine learning make it possible to accurately assess the properties of rocks, which is useful for optimizing drilling. In this paper, the authors present a new method for classifying mountain impacts based on cross-reconstructive learning (Setyanto, Sunyoto, 2024). The proposed model, using XGBoost and random forest ensembles, demonstrates improved prediction performance compared to traditional methods.

Nguyen and others are conducting a comparative assessment of self-learning methods for classifying rock images. The authors demonstrate that the MoCo-v2 method is superior to traditional algorithms with a limited amount of labeled data. Lee and others are investigating the classification of mountain ranges based on sensors and ensemble machine learning (Cai et al., 2024). The use of several machine learning methods combined with vibration data analysis makes it possible to achieve high prediction accuracy. Hansen and others analyze machine learning to predict the quality of mountain ranges during tunneling. The use of KNN-based algorithms and decision trees makes it possible to automate the analysis of drilling data and improve the accuracy of rock stability assessment. Others consider machine learning to classify mining impacts in coal mines. Ensamble methods such as gradient boosting and KNN can effectively distinguish between destructive and non-destructive tremors, increasing mining safety (Klyuev et al., 2024; Zykova et al., 2019; Malozyomov et al., 2024). The presented literature review demonstrates significant progress in the application of machine learning methods in geology and mining. The use of ensemble methods, deep learning, and integration with numerical models can significantly improve the accuracy of rock analysis.

Results and discussion. In the course of the study, a correlation analysis was performed to identify possible linear relationships between the input chemical data and the output parameter – the type of rock. Figure 1 shows a correlation matrix showing the Pearson correlation coefficients between all chemical parameters and the output parameter (rock type).

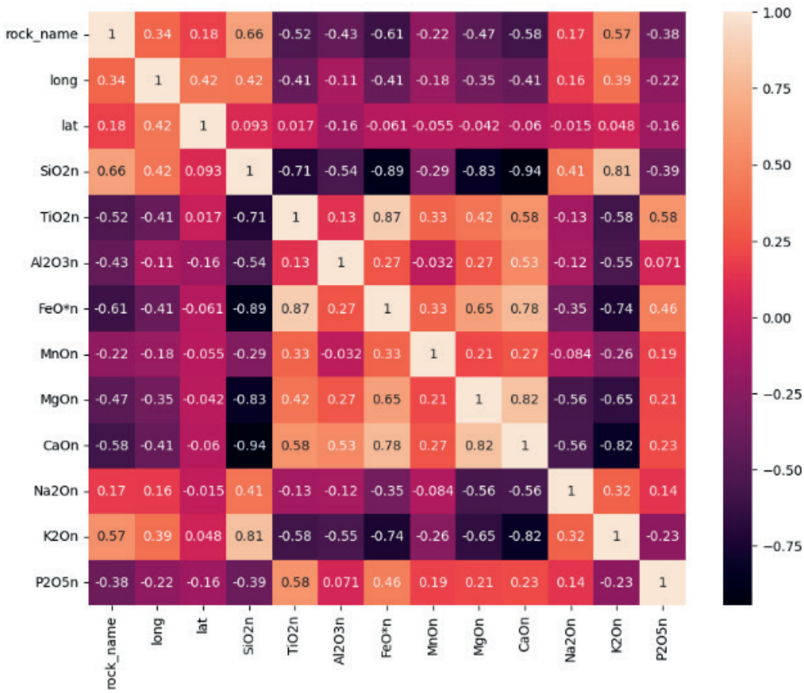

Fig. 1. Correlation matrix

The following strong positive correlations have been identified. SiO_{2n}-K_2O_n: 0.81. This means that the concentrations of SiO_{2n} (silica) and K_2O_n (potassium oxide) in the samples are strongly positively correlated. Mathematically, this means that as the SiO2n value increases, the K_2O_n value in the samples is likely to increase, and vice versa (Hansen et al., 2024). Rocks such as rhyolites and basaltic andesites often contain high concentrations of silica (SiO_{2n}) and potassium oxide (K_2O_n). For example, the high content of SiO_{2n} in rhyolites is due to their silica composition, and potassium may be a part of potassium feldspaths, which are also present in these rocks. FeO_{*n}-TiO_{2n}: 0.87. Concentrations of FeO_n (iron oxide) and TiO_{2n} (titanium oxide) are also strongly positively correlated. An increase in the FeOn content is often accompanied by an increase in the TiO2 content and vice versa (Gao, 2024). Basalts and diorites, as well as basaltic andesites and diorite rocks, often contain high concentrations of iron oxide (FeO_{*n}) and titanium oxide (TiO_2N). This is due to their magmatic origin, where these elements can form minerals such as pyroxenes (containing iron) and aluminum (containing titanium). CaO_n-MgO_n: 0.82. This indicates a strong positive correlation between the content of CaO_n (calcium oxide) and MgO_n (magnesium oxide). An increase in the content of one of these elements often corresponds to an increase in the content of the other. Granites, granodiorites, and diorites typically contain high concentrations of calcium (CaO_n) and magnesium (MgO_n). For example, calcite and dolomite containing calcium and magnesium, respectively, are often found in these rocks, leading to a positive correlation between them (Tynchenko et al., 2024).

The following strong negative correlations have been identified. SiO_{2n}-FeO_{*n}: -0.89. The negative correlation between SiO_{2n} and FeO_n indicates that an increase in $SiO2_n$ content is usually accompanied by a decrease in FeO_n content and vice versa. Diorites and granites, characterized by a high content of silica ($SiO2_n$), often have a low content of iron oxide (FeO_{*n}). This is due to the fact that silica-rich minerals such as quartz, which do not contain significant amounts of iron, are often found in granite and diorite rocks. SiO_{2n}-CaO_n: -0.94. The high negative correlation between SiO_{2n} and CaO_n indicates that an increase in SiO2n content is usually accompanied by a decrease in CaO_n content and vice versa. Rhyolites and rhyolitic tuffs rich in silica (SiO2n) usually have a low content of calcium (CaO_n). This may be due to the fact that rhyolites are often formed in conditions that are not conducive to the enrichment of rocks with calcium, which is more common in carbonate rocks, for example, in limestones. SiO_{2n}-MgO_n: -0.83. This means that the concentrations of SiO_{2n} and MgO_n have a strong negative correlation. An increase in SiO2n content is often accompanied by a decrease in MgO_n content and vice versa. Diorites and quartz diorites rich in silica (SiO_{2n}) often have low magnesium (MgO_n) content. This is due to the fact that such rocks are dominated by silica minerals, which usually do not contain significant amounts of magnesium. CaOn-K_2O_n: -0.82. The negative correlation between CaO_n and K_2O_n indicates that an increase in CaO_n content is often accompanied by a decrease in K_2O_n content and vice versa. Basalt andesites and basalt andesite inclusions often contain low concentrations of calcium (CaO_n) and high concentrations of potassium oxide (K_2O_n). This may be due to the fact that magma-rich basalts are often formed under conditions conducive to the enrichment of rocks with potassium due to various minerals such as olivine.

The overall accuracy of the random forest model is 84%, which is a good indicator. Figure 2 shows the classification report for this model.

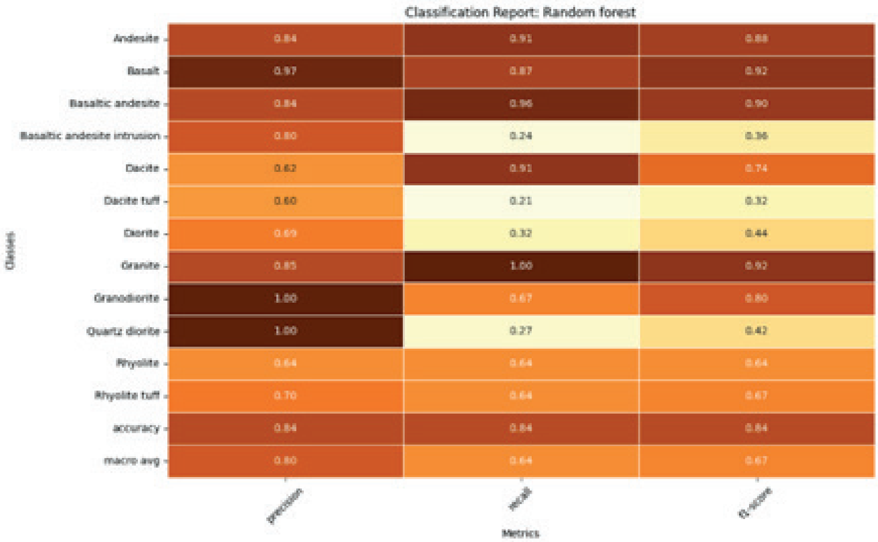

Fig. 2. Classification report

The random forest exhibits high Precision values for most classes, for example, for Basalt (0.97), Basaltic andesite (0.84), and Granite (0.85). This means that most of the predictions of the positive class were correct.

For classes with high Recall (for example, Granite with a completeness of 1.00), the model successfully detects almost all true positive examples of this class. This means that the model rarely skips breeds of this type, which is important for accurate classification. Low Recall values (for example, Basaltic andesite intrusion with a Recall of 0.24 and Quartz diorite with a Recall of 0.27) indicate the difficulty of the model in detecting true positive cases for these classes. This may be caused by an insufficient representation of these classes in the training dataset or their similarity to other classes, which makes it difficult to distinguish them.

The F1 measure also shows a variety of results from low values (for example, Basaltic andesite intrusion 0.36 and Quartz diorite 0.42) to high values (for example, Basalt 0.92 and Granite 0.92). This confirms the balance between accuracy and completeness for different classes, where high values indicate the effectiveness of the model in correct prediction and the balance between these two metrics.

Figure 3-4 shows the accuracy and recall curve and the error matrix for the random forest model.

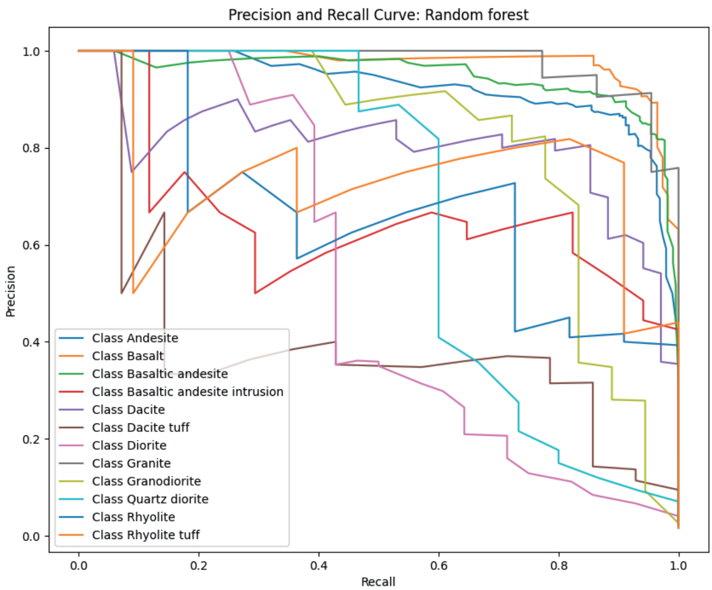

Fig. 3. Precision and Recall Curve.

The random forest shows stable and smooth curves for many classes, indicating high accuracy and recall.

The model achieves high accuracy for several classes such as Basalt, Granite and Basalt andesite. The curves indicate the model's good data generalization ability.

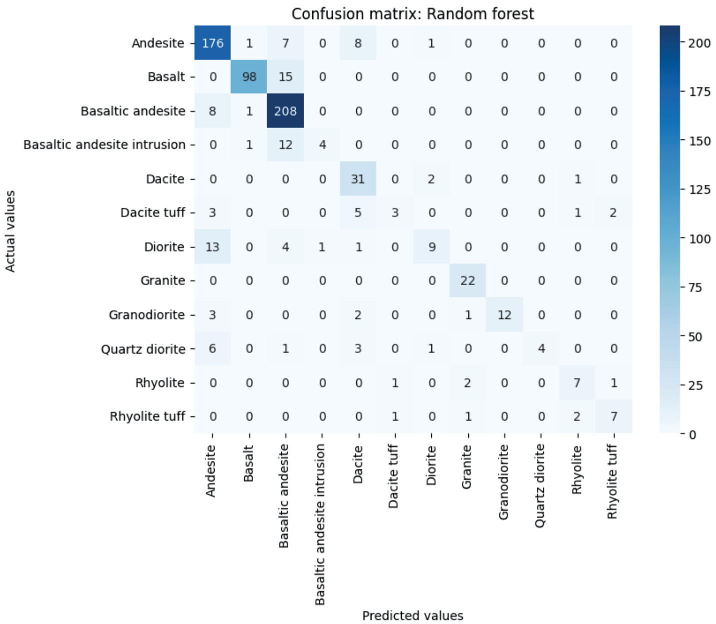

Fig. 4. Confusion matrix.

The random forest model has shown good results. The error matrix demonstrates that a random forest has a small number of errors, especially for rocks such as Basaltic andesite and Granite. The Precision-Recall graphs confirm the high accuracy and reliability of this model, with high precision and recall values for most classes.

In future work and research, it is necessary to work on improving the accuracy of the random forest model. This can be done using various methods, such as optimizing hyperparameters, using ensemble methods, and increasing the amount of data for training, since the constant updating and expansion of the dataset with new data on geological rocks can also improve classification accuracy.

Conclusion. The study examined the use of one of the machine learning methods for classifying types of geological rocks based on their chemical composition. The developed random forest model showed an accuracy of 84%.

The analysis of error matrices and Precision-Recall graphs allowed us to confirm that the random forest model demonstrates good performance.

Based on the analysis, it is recommended to focus on further improving this model by optimizing hyperparameters, increasing the amount of data for training, as well as using ensemble methods and deep learning.

Thus, this work demonstrates that the use of modern machine learning methods for classifying geological rocks is an effective approach that can significantly improve the quality and speed of tasks in geology.

References

Borodulin A.S., Kukartsev V.V., Glinscaya A.R., Gantimurov A.P., Nizameeva A.V. Analyzing Data by Applying Neural Networks to Identify Patterns in the Data. In: Proceedings of the Computational Methods in Systems and Software. Cham: Springer Nature Switzerland, 2023. — P. 99–108. (in Eng.)

Bosikov I.I., Klyuev R.V., Martyushev N.V., Modina M.A., Khekert E.V. Analysis of the quality of underground mineral waters of terrigenous deposits of the hauteriv barremian aquifer of the lower cretaceous. News of the National Academy of Sciences of the Republic of Kazakhstan, Series of Geology and Technical Sciences, 2024. — T. 2. — № 464. — P. 36-47. DOI: 10.32014/2024.2518-170x.392 (in Eng.)

Bosikov I.I., Klyuev R.V., Revazov V.Ch., Martyushev N.V. Analiz i otsenka perspektiv vysokokachestvennykh kvartsevykh resursov na Severnom Kavkaze [Analysis and evaluation of prospects for highquality quartz resources in the North Caucasus]. Mining Science and Technology, 2023. — №8(4). — P. 278–289. https://doi.org/10.17073/2500-0632-2023-10-165 (In Russ.).

Cai X., Chen L., Zhou Z., Yuan H., Wang P., & Zheng A. (2024) Rockburst classification based on cross reconstruction learning under small-sample condition. Journal of Physics: Conference Series, 2738. (in Eng.)

Epikhin A.I., Stupina A.A., Panfilov I.A., Bukhtoyarov V.V., Shepeta N.A. Determinants for assessing the energy efficiency of a coal mining enterprise. News of the National Academy of Sciences of the Republic of Kazakhstan, Series of Geology and Technical Sciences, 2024. — 2024(4). — P. 61–71. (in Eng.)

Hamada G., Al-Khudafi A., Al-Yazidi A., Al-Sharifi H., Al-Qadhi T., Al-Gathe A. Enhanced Machine Learning Modelling Techniques for Better Classification of Carbonate Reservoir Rocks. Mediterranean Offshore Conference (Alexandria, Egypt, 20–22 October 2024), 2024. (in Eng.)

Hansen T., Erharter G., Liu Z., Torresen J. A comparative study on machine learning approaches for rock mass classification using drilling data. ArXiv, 2024. — abs/2403.10404. (in Eng.)

Klyuev R.V., Martyushev N.V., Kukartsev V.V., Kukartsev V.A., Brigida V. Analiz geologicheskoy informatsii dlya obespecheniya ustoychivogo funktsionirovaniya gornotekhnicheskikh sistem [Analysis of geological information toward sustainable performance of geotechnical systems]. Mining Informational and Analytical Bulletin [Gornyy informatsionno-analiticheskiy byulleten'], 2024. — No. 5. DOI: 10.25018/0236-1493-2024-5-0-144. (In Russ.).

Koteleva N.I., Korolev N.A., Revin I.E. Primeneniye algoritmov avtomaticheskogo mashinnogo obucheniya pri diagnostike neispravnostey elektrodvigateley peremennogo toka na predpriyatiyakh gornodobyvayushchey promyshlennosti [Application of automatic machine learning algorithms in fault diagnostic of ac electric motors in mineral resource industry enterprises]. Sustainable Development of Mountain Territories [Ustoychivoye razvitiye gornykh territoriy]. – 2024. – vol. 16. — no. 4. — P. 1671–1680. https://doi.org/10.21177/1998- 4502-2024-16-4-1671-1680. (In Russ.).

Malozyomov B.V., Martyushev N.V., Sorokova S.N., Efremenkov E.A., Valuev D.V., Qi M. Analysis of a Predictive Mathematical Model of Weather Changes Based on Neural Networks. Mathematics, 2024. — Vol. 12, No. 3. — Art. 480. DOI: 10.3390/math12030480. (in Eng.)

Morgenroth J., Khan U.T., Perras M.A. An overview of opportunities for machine learning methods in underground rock engineering design. Geosciences, 2019. — 9.12. — P. 504. (in Eng.)

Morgoeva A.D., Mandzhieva S.S., Kirichkov M.V., Sokolov A.A. Issledovaniye modeley mashinnogo obucheniya dlya otsenki vliyaniya predpriyatiy ugledobychi i energetiki na ekosistemy [Machine learning models study for assessing the effect of coal mining and energy enterprises on ecosystems]. Sustainable Development of Mountain Territories [Ustoychivoye razvitiye gornykh territoriy], 2024. — vol. 16. —no. 3. — P. 1130–1143. https://doi.org/10.21177/1998- 4502-2024-16-3-1130-1143. (In Russ.).

Olentsova J., Kukartsev V., Orlov V., Semenova E., Pinchuk I. Predictive modelling of post-monsoon groundwater quality in Telangana using machine learning techniques. BIO Web of Conferences, 2024. — 116. — P. 03021. EDP Sciences. (in Eng.)

Setyanto A., K., Sunyoto A. Comparative Study SVM and Random Forest Algorithms for the Classification of Terrestrial Visual Rock Types. IOP Conference Series: Earth and Environmental Science, 2024. — 1357. — P. 012036. (in Eng.)

Tynchenko Y.A., Kukartsev V.V., Xiaogang W., Kravtsov K.I. Modelirovaniye intensivnosti zasukhi v gornykh rayonakh s ispol'zovaniyem meteorologicheskikh parametrov [Modeling the drought intensity in mountainous areas using meteorological parameters]. Sustainable Development of Mountain Territories [Ustoychivoye razvitiye gornykh territoriy], 2024. — 16(2). — P. 655–668. (In Russ.).

Tynchenko Ya.A., Kukartsev V.V., Gladkov A.A., Panfilova T.A. Otsenka kachestva tekhnicheskoy vody v gornodobyvayushchey promyshlennosti na osnove metodov mashinnogo obucheniya [Assessment of technical water quality in mining based on machine learning methods]. Sustainable Development of Mountain Territories [Ustoychivoye razvitiye gornykh territoriy], 2024. — vol. 16. — no. 1. — P. 56–69. https://doi.org/10.21177/1998-4502-2024-16-1-56-69. (In Russ.).

Zemlyanoy M.A., Gabaraev O.Z., Mulukhov K.K., Vyskrebenets A.S. Substantiation of model for prediction of ore body contours and useful components content in reserves based on artificial neural networks. Sustainable Development of Mountain Territories, 2024. — vol. 16. — no. 2. — P. 780–788. https://doi.org/10.21177/1998-4502-2024-16-2-780-788. (in Eng.)

Zhu D., Yu B., Wang D., Zhang Y. Fusion of finite element and machine learning methods to predict rock shear strength parameters. Journal of Geophysics and Engineering, 2024. — 21(4). — P. 1183–1193. (in Eng.)

Zykova A., Martyushev N., Skeeba V., Zadkov D., Kuzkin A. Influence of W addition on microstructure and mechanical properties of Al-12%Si alloys. Materials, 2019. — Vol. 12, No. 6. — Art. 981. DOI: 10.3390/ma12060981. (in Eng.)

Publication Ethics and Publication Malpractice in the journals of the Central Asian Academic Research Center LLP

For information on Ethics in publishing and Ethical guidelines for journal publication see http://www.elsevier.com/publishingethics and http://www.elsevier.com/journal-authors/ethics.

Submission of an article to the journals of the Central Asian Academic Research Center LLP implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see http://www.elsevier. com/postingpolicy), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The Central Asian Academic Research Center LLP follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct (http://publicationethics. org/files/ u2/New_Code.pdf). To verify originality, your article may be checked by the Cross Check originality detection service http://www.elsevier.com/editors/plagdetect.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/ or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the Central Asian Academic Research Center LLP.

The Editorial Board of the Central Asian Academic Research Center LLP will monitor and safeguard publishing ethics.

Правила оформления статьи для публикации в журнале смотреть на сайтах:

www:nauka-nanrk.kz
http://www.geolog-technical.kz/index.php/en/
ISSN 2518-170X (Online),
ISSN 2224-5278 (Print)

Ответственный редактор *А. Ботанқызы*
Редакторы: *Д.С. Аленов, Т. Апендиев*
Верстка на компьютере: *Г.Д. Жадырановой*

Подписано в печать 15.10.2025.
Формат $70x90^1/_{16}$. 20,5 п.л.
Заказ 5.

«Central Asian Academic Research Center» LLP
Алматы, Қонаев к-сі, 142