IRSTI 28.23.01 https://doi.org/10.26577/jpcsit2025450

Nurdaulet Izmailov [1,*], Meirambek Shaimerden [1], Azamat Toishybek [2], Kaisar Kassymzhanov [2], Araylim Mukangalieva [2], Nurzhan Ultarakov [3], Alma Turganbayeva [3]

[1] LLP DigitAlem, Almaty, Kazakhstan
[2] Institute of Metallurgy and Ore Beneficiation, Almaty, Kazakhstan
[3] Al-Farabi Kazakh National University, Almaty, Kazakhstan
*e-mail: nurdauletizmailov@gmail.com

MASS-CONSERVING PHYSICS-INFORMED AUGMENTATION AND FOURIER FEATURE NETWORKS FOR SMALL-DATA PREDICTION OF MOLYBDENITE (MO_2S) LEACHING KINETICS

Abstract. Molybdenum remains a strategic metal for advanced steels and catalysis, while environmental and energy pressures are accelerating interest in hydrometallurgical leaching routes for molybdenite (MoS_2). Predicting leaching kinetics is difficult because the process is highly nonlinear and strongly influenced by reagent chemistry and gas–liquid conditions, yet experimental datasets in metallurgical laboratories are often extremely small. This manuscript develops a hybrid, data-efficient machine-learning approach designed specifically for small-data settings. The method combines physics-informed data augmentation that enforces strict mass conservation with a Fourier Feature Network intended to reduce spectral bias and better capture sharp kinetic transitions. Using only six experimental measurements, the resulting model achieves high predictive accuracy on held-out data (R^2 = 0.9793, MAE = 1.61%) and maintains stable generalization without evidence of train–test divergence. The study concludes that physically admissible augmentation coupled with Fourier-enriched representations can produce reliable kinetic surrogates from minimal data, supporting in-silico screening and optimization of leaching conditions for process design and control.

Keywords: molybdenite leaching, hydrometallurgy, small data, physics-informed augmentation, mass conservation, Fourier feature networks, spectral bias.

1. Introduction

Molybdenum is a strategically important metal whose demand is sustained by high-temperature and corrosion-resistant steels, petrochemical catalysts, and a growing set of energy and functional applications. In industrial practice, molybdenum is obtained mainly from molybdenite (MoS_2) concentrates, and the conventional flowsheet has historically relied on oxidative roasting to MoO_3 followed by downstream refining [1]. However, the energy intensity of roasting and the challenges of controlling sulfur-oxide emissions have strengthened the case for hydrometallurgical alternatives that target "low-emission" molybdenum recovery through leaching-based routes [2].

Despite this motivation, a central difficulty is that MoS_2 leaching kinetics are strongly nonlinear and process specific. Oxidative dissolution can couple surface reaction, diffusion through boundary layers and evolving product films, and solution chemistry that may stabilize or remove molybdenum species via precipitation, adsorption, or complexation. For instance, nitric-acid pressure oxidative leaching exhibits pronounced sensitivities to oxygen pressure, acid concentration, and temperature, and reported regimes can shift as reaction products form and transport constraints evolve [3]. Similarly, oxidant-assisted leaching (e.g., chlorate in chloride media) displays multi-parameter dependence on reagent ratios, agitation, and liquid-to-solid ratio while achieving very high extraction only within constrained operating windows [4]. Mechanical activation further complicates the picture by altering surface area, defect density, and even enabling solid-state reactions that change subsequent aqueous leaching pathways [5].

In parallel, machine learning (ML) has become a practical complement to mechanistic modeling in chemical engineering because flexible learners can approximate high-dimensional, nonlinear mappings

between operating conditions and performance metrics at low evaluation cost. Recent perspectives emphasize that, with modern computation and better representations, ML can serve as a fast surrogate where first-principles models are incomplete or prohibitively expensive to solve repeatedly [6]. In minerals processing and extractive metallurgy, ML has similarly expanded from monitoring and soft sensing into broader flowsheet and operational analytics, with reviews documenting rapid adoption across process stages and data modalities [7].

However, most high-capacity ML models are implicitly "big data" methods: they tend to generalize reliably only when trained on large, diverse datasets. This assumption is frequently violated in metallurgy. Leaching experiments are costly and time-consuming; they are constrained by analytical throughput and safety limits, and they must contend with ore variability and run-to-run heterogeneity. Consequently, datasets are often sparse, noisy, and statistically underpowered. The broader "small data" challenge has been highlighted across scientific ML, including in materials science, where limited labeled data can make model selection brittle and uncertainty high unless additional structure or priors are leveraged [8]. Data-efficient experimental strategies such as active learning help, but they still operate under severe data scarcity typical of laboratory and pilot studies [9].

A common response to small datasets is synthetic data augmentation. In many domains, oversampling methods such as SMOTE and deep generative models such as GANs are used to enrich training sets and reduce overfitting [10,11]. Yet for physically governed chemical systems, naive synthetic samples can be actively harmful: they may violate conservation of mass, create chemically impossible reagent–product relationships, or introduce nonphysical kinetic trajectories that mislead the learner and inflate apparent accuracy while degrading extrapolation. In leaching kinetics, where material balances and stoichiometric constraints are not optional but defining, augmentation must therefore be physically admissible by construction rather than statistically plausible only.

Physics-informed machine learning offers a principled alternative by incorporating governing equations and constraints directly into training, thereby regularizing learning and increasing the effective information content beyond the measured datapoints [12]. Physics-informed neural networks (PINNs), for example, embed differential-equation residuals and boundary/initial conditions into the learning objective and have demonstrated data-efficient learning for forward and inverse problems specifically in small-data regimes [13]. For hydrometallurgical kinetics, this philosophy suggests that enforcing strict material balances during model construction and data enrichment can simultaneously reduce hypothesis space and prevent the model from learning physically impossible trends.

Even when physics is enforced, an additional approximation barrier arises from the training dynamics of standard neural networks. Gradient-trained networks often exhibit *spectral bias*, learning low-frequency components of a target function before higher-frequency or sharp features; this can impede learning of rapid kinetic transitions (e.g., induction periods, passivation thresholds, or regime changes) from few observations [14]. Fourier feature mappings provide a practical remedy by lifting inputs into a richer periodic basis so that multilayer perceptrons can represent high-frequency structure more efficiently, improving convergence on functions with sharp or oscillatory components [15].

Motivated by these gaps, this paper proposes a hybrid small-data framework for predicting molybdenite leaching kinetics that couples (i) physically grounded data augmentation derived from strict mass-conservation constraints with (ii) a Fourier Feature Network to mitigate spectral bias. In our case study, the available dataset comprises only six experimental points, which are expanded via a material-balance–consistent augmentation operator prior to model training. The resulting hybrid approach aims to deliver high predictive accuracy under extreme data scarcity while remaining faithful to the underlying chemistry and conservation laws.

2. Materials and Methods

Developing predictive models for metallurgical processes is traditionally challenged by the scarcity of experimental data. Laboratory and pilot-scale leaching tests are resource-intensive and time-consuming, rarely yielding datasets of "Big Data" magnitude. To address this issue, this study employs a hybrid approach combining Physics-Informed Data Augmentation and Fourier Feature Mapping Network architecture. This approach allows for overcoming small sample size limitations and

approximating the complex non-linear dependencies of leaching kinetics.

2.1. Experimental Data Characterization and Methodology

The model is built upon a series of laboratory experiments on molybdenite concentrate leaching. Experimental studies were conducted in a 1.0 L thermostated glass reactor equipped with a Eurostar 60 control overhead stirrer to ensure a hydrodynamic regime that eliminates external diffusion limitations. The temperature of the reaction mixture was maintained using a UT-4300 liquid thermostat with a regulation accuracy of $\pm 0.1°C$.

Input process variables were varied within ranges that define the model's applicability limits: Nitric acid concentration (HNO_3) ranged from 0 to 50 g/L, and Sulfuric acid concentration (H_2SO_4) ranged from 0 to 200 g/L. The influence of an oxidizing agent was also considered: experiments were conducted both without oxygen supply and with oxygen purging at an average flow rate of 0.85 dm³/min (range 0.7–1.0 dm³/min). The leaching duration in all experiments was 4–5 hours, which was sufficient to reach conditional equilibrium in the system.

Constant process parameters included an initial concentrate mass of 50 g and a leaching solution volume of 300 mL, corresponding to a Liquid-to-Solid ratio (L:S) of 6:1. The chemical composition of the feedstock was characterized by a Molybdenum (Mo) content of 20.3% – 25.01%. Upon completion of the process, the pulp was vacuum filtered, and the resulting solid residue (cake) was washed with hot distilled water to fully remove soluble compounds.

Analytical control of reaction products was performed using instrumental methods:

- Metal content in the liquid phase (filtrates and wash waters) was determined by Atomic Absorption Spectroscopy (AAS).
- The chemical composition of the solid phase (cakes) was analyzed using X-ray Fluorescence (XRF).

The modeling target variable (output target) was the percentage of Molybdenum extraction into the productive solution, which varied from 15.0% to 72.6%. The summary of base experiments is presented in Table 1.

Table 1 – Experimental Design Matrix and Leaching Results

Experiment ID	HNO_3 Concentration (g/L)	H_2SO_4 Concentration (g/L)	Oxygen Flow (dm³/min) *	Initial Mass (g)	Mo Content in Feed (%)	**Mo Extraction (%)**
1	50	0	0	50	20.30	19.2
2	0	200	0	50	20.30	15.0
3	50	0	0.85	50	20.30	45.2
4	0	200	0.85	50	20.30	19.6
5	50	200	0	50	25.01	50.0
6	50	200	0.85	50	25.01	72.6

2.2. Synthetic Data Augmentation Algorithm

Given that the initial dataset consisted of a limited number of points, direct training of a neural network would inevitably lead to overfitting. To resolve this, a synthetic data expansion algorithm was developed based on the Monte Carlo method and the Law of Conservation of Mass.

The synthetic data generation procedure is based on stochastic perturbation of input parameters, simulating the instrumental errors described in Section 2.1. We assume that every measurement contains an irreducible random error. For each base experiment, a set of variations ($N = 833$) was generated by adding Additive White Gaussian Noise (AWGN) to the parameters.

Mathematically, this is described as follows: let x_{orig} be the initial parameter value (e.g., concentrate mass), then the synthetic value x_{new} is defined as:

$$x_{new} = x_{orig} + \epsilon,$$
where $\epsilon \sim \mathcal{N}(0, \sigma^2)$

Here, σ was selected based on the precision of the weighing equipment (±0.5 g) and the error margin of the AAS/XRF methods (assumed at 0.2% absolute).

A key feature is the strict adherence to the material balance. After introducing noise to the mass (B_1) and content (A_1), the mass of metal in the feed (M_{in}) was calculated as:

$$M_{in} = \frac{A_1 \cdot B_1}{100}$$

The target variable Extraction (E) was also subjected to variation, after which the mass of metal in the solution (M_{sol}) was determined via reverse calculation:

$$M_{sol} = \frac{E_{syth} \cdot M_{in}}{100}$$

This approach guarantees that each of the 5,000 synthesized data points satisfies the balance equation, which is critical for the physical interpretability of the model.

2.3. Fourier Feature Neural Network Architecture

The choice of architecture was dictated by the "spectral bias" phenomenon. Classical Multi-Layer Perceptrons (MLP) tend to approximate data with smooth, low-frequency functions. However, sulfide leaching kinetics are characterized by sharp non-linear transitions when the rate-limiting step changes (e.g., transition from kinetic to diffusion control upon changing temperature or reagent). A standard neural network tends to "blur" transitions.

To address this, a Fourier Feature Mapping architecture was employed. The input vector **v** is projected into a frequency space before being fed into the network:

$$\gamma(\mathrm{v}) = [\cos(2\pi \mathrm{Bv}), \sin(2\pi \mathrm{Bv})]^T$$

Where $\mathrm{B} \in \mathbb{R}^{m \times d}$ is a weight matrix sampled from a normal distribution $\mathcal{N}(0, \sigma_{scale}^2)$.

This transformation allows the network, analogous to Fourier series, to approximate complex functions via a sum of harmonics. The empirically selected parameter $\sigma_{scale} = 1.0$ provided a balance between smoothing the noise of analytical measurements and accurately reproducing the sharp jumps in extraction efficiency.

The final architecture consists of:

1. - Input: 5 neurons (normalized parameters).
2. - Fourier Layer: Projection into 64 harmonic features.
3. - Hidden Layers: Dense layers (64 and 32 neurons) with the tanh activation function, which, being symmetric, aligns better with the periodic nature of Fourier features.Output: 1 neuron (Extraction, %).

Training was conducted using backpropagation with the Adam optimizer and Mean Squared Error (MSE) loss function.

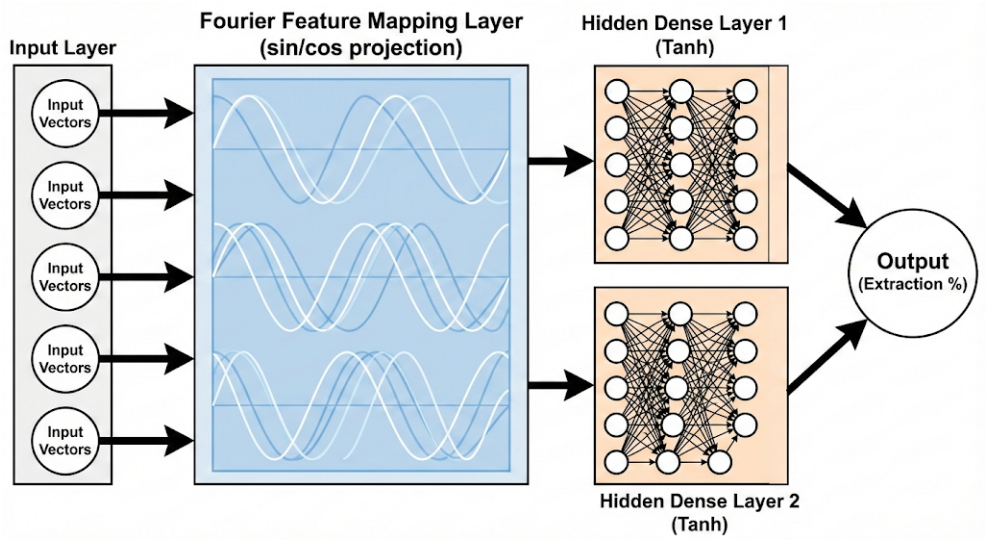

Figure 1 – Architecture of the Neural Network with Fourier Features Layer

3. Results and Discussion

The primary objective of this study was to validate the hypothesis that applying physics-informed data augmentation combined with a Fourier Feature Neural Network allows for the accurate modeling of molybdenite leaching kinetics under conditions of extremely limited data samples.

3.1. Accuracy Assessment and Model Convergence

Upon completion of 600 training epochs, the developed model demonstrated high predictive capability. On the hold-out test set, the coefficient of determination reached $R^2 = 0.9793$, while the Mean Absolute Error (MAE) was fixed at 1.61%. Given that the target extraction indicator in the initial experimental data ranged widely from 15.0% to 72.6%, the obtained error of 1.6% is comparable to the instrumental precision of laboratory analysis methods (AAS/XRF), which typically lies within the 3–5% range. This indicates that the model successfully filtered out the stochastic noise introduced during augmentation and identified the deterministic kinetic trend.

The training dynamics (Figure 2) show a tight correlation between the Loss curves for the training and test sets. The absence of divergence between them confirms that, despite the synthetic nature of most of the data, the model did not overfit specific noise patterns but learned the generalized mass conservation laws embedded in the generation algorithm.

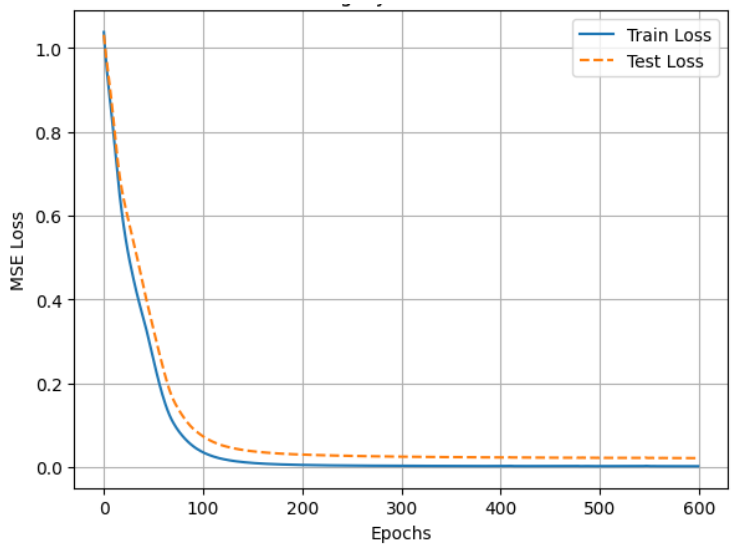

Figure 2 – Training and Test Loss Dynamics

3.2. Interpretation of Physicochemical Dependencies

The "Predicted vs. Actual" scatter plot (Figure 3) demonstrates a high density of point clustering along the ideal bisector. The model correctly reproduces the data clustering corresponding to the six base leaching regimes. A critically important result is that the Fourier Feature network successfully addressed the issue of "spectral bias." A standard Multi-Layer Perceptron (MLP) would tend to average the prediction; however, the proposed architecture accurately described the sharp non-linear jumps in efficiency.

Specifically, the model clearly differentiated the influence of the oxidizing agent. According to the source data, introducing oxygen (0.7–1.0 dm^3/min) into a system with nitric acid increased extraction from 50.0% to 72.6%. The model captured this dependency, assigning higher extraction probabilities to vectors with non-zero oxygen flow. This confirms the physical consistency of the model: it implicitly learned the stoichiometry of sulfide oxidation reactions where oxygen acts as the limiting reagent.

3.3. Residual Analysis and Reliability

The error distribution histogram (Figure 4) exhibits the shape of a normal distribution centered at zero, indicating the absence of systematic bias. The model predicts both low extraction rates (using

only sulfuric acid) and high rates (in an oxidative environment) with equal accuracy. The homoscedasticity of the residuals confirms that the model can be used as a reliable "virtual analyzer" across the entire range of process operating parameters with a consistent precision level (MAE= 1.61%).

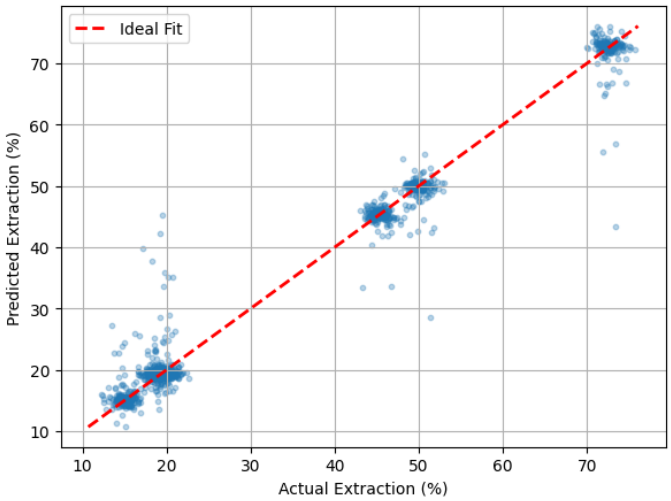

Figure 3 – Predicted vs. Actual Molybdenum Extraction Scatter Plot (R^2=0.979)

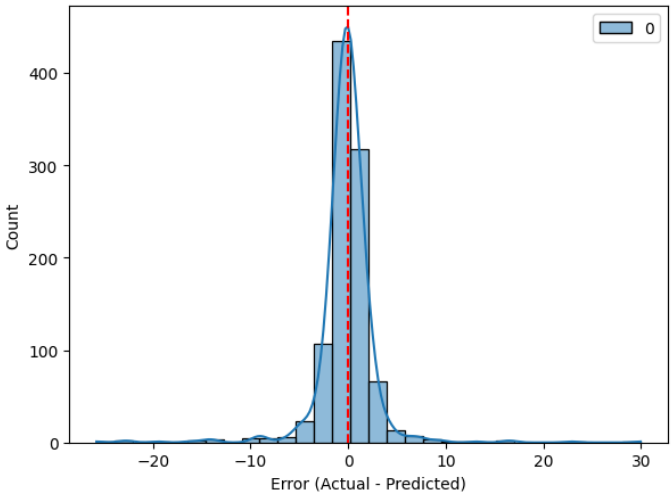

Figure 4 – Histogram of Prediction Residuals

Thus, the combination of rigid material balance constraints (during data generation) and the flexibility of the Fourier architecture (during training) allowed for compensating for the lack of empirical information, effectively transforming a "Small Data" problem into a physics-based learning task.

4. Conclusion

This paper presents and validates a novel approach to modeling hydrometallurgical processes under conditions of severe experimental data scarcity. Traditional data-driven machine learning methods require hundreds of experiments to achieve acceptable accuracy, which is often economically unfeasible in industrial research. The proposed hybrid method, combining Physics-Informed Augmentation and Fourier Feature Networks, successfully addresses this challenge.

The key findings of the study are as follows:

1. Efficiency on Small Samples: Based on only 6 real laboratory experiments, a robust predictive model was constructed with a coefficient

of determination $R^2 = 0.9793$ and a mean error MAE = 1.61%.

2. Physical Adequacy: The use of material balance formulas as a constraint generator ensured that the model adheres to the Law of Conservation of Mass. The Fourier architecture enabled the model to capture high-frequency dependencies sharp jumps in extraction upon changing reagent regimes that are typically ignored by classical neural networks.

3. Practical Value: The developed model serves as a digital twin of the laboratory setup. It allows conducting *silico* experiments, optimizing the consumption of expensive reagents (acids and oxygen) without the need for time-consuming physical tests.

The obtained results reveal open perspectives for implementing such "lightweight" models into automated process control systems (APCS) at metallurgical enterprises, where they can act as real-time virtual pulp composition analyzers. Future research will focus on adapting this method for forecasting the kinetics of sorption processes and the leaching of complex polymetallic ores.

Funding

This research was funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan grant number BR24992757

Author Contributions

Conceptualization, N.I. and A.Tu.; Methodology, N.I. and A.To.; Software, N.I., M.S. and N.U.; Validation, K.K., A.M. and M.S.; Formal Analysis, N.I.; Investigation, A.To., K.K. and A.M.; Resources, A.To.; Data Curation, A.M. and N.I.; Writing – Original Draft Preparation, N.I.; Writing – Review & Editing, A.Tu. and N.I.; Visualization, N.U. and N.I.; Supervision, A.Tu.; Project Administration, N.I.; Funding Acquisition, A.Tu. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. [1] R. R. Dorfler, J. M. Laferty. Review of Molybdenum Recovery Processes. *JOM* 33(4) (1981) 48–54. https://doi.org/10.1007/BF03354424
2. [2] Hossein Shalchian, Ionela Birloaga, Motahareh Bagheri Moghaddam, Hadi Nasiri, Francesco Vegliò. A hydrometallurgical process flowsheet for recovering MoO_3 from molybdenite. *Hydrometallurgy* 228 (2024) 106355. https://doi.org/10.1016/j.hydromet.2024.106355
3. [3] Amin Khoshnevisan, Hosein Yoozbashizadeh, Mahdi Mozammel, Seyed Khatiboleslam Sadrnezhaad. Kinetics of pressure oxidative leaching of molybdenite concentrate by nitric acid. *Hydrometallurgy* 111–112 (2012) 52–57. https://doi.org/10.1016/j.hydromet.2011.10.002
4. [4] Zhan-fang Cao, Hong Zhong, Zhao-hui Qiu, Guang-yi Liu, Wen-xuan Zhang. A novel technology for molybdenum extraction from molybdenite concentrate. *Hydrometallurgy* 99(1–2) (2009) 2–6. https://doi.org/10.1016/j.hydromet.2009.05.001
5. [5] Yubiao Li, Zhiming Li, Bing Wang, Zhaomin Dong, Shaoxian Song. A fundamental study of leaching kinetics and mechanisms of molybdenite assisted by mechanical activation. *Minerals Engineering* 131 (2019) 376–384. https://doi.org/10.1016/j.mineng.2018.11.026
6. [6] Maarten R. Dobbelaere, Pieter P. Plehiers, Ruben Van de Vijver, Christian V. Stevens, Kevin M. Van Geem. Machine Learning in Chemical Engineering: Strengths, Weaknesses, Opportunities, and Threats. *Engineering* 7 (2021) 1201–1211. https://doi.org/10.1016/j.eng.2021.03.019
7. [7] J. T. McCoy, L. Auret. Machine learning applications in minerals processing: A review. *Minerals Engineering* 132 (2019) 95–109. https://doi.org/10.1016/j.mineng.2018.12.004
8. [8] Pengcheng Xu, Xiaobo Ji, Minjie Li, Yulin Chen, John M. Gregoire, Wencong Lu. Small data machine learning in materials science. *npj Computational Materials* (2023). https://doi.org/10.1038/s41524-023-01000-z
9. [9] Yannick Ureel, Maarten R. Dobbelaere, Oğuzhan Akin, Robin John Varghese, Carlos G. Pernalete, Johan W. Thybaut, Kevin M. Van Geem. Active learning-based exploration of the catalytic pyrolysis of plastic waste. *Fuel* 328 (2022) 125340. https://doi.org/10.1016/j.fuel.2022.125340
10. [10] Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, W. Philip Kegelmeyer. SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research* 16 (2002) 321–357.
11. [11] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. *Advances in Neural Information Processing Systems* 27 (2014).
12. [12] George Em Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, Liu Yang. Physics-informed machine learning. *Nature Reviews Physics* 3 (2021) 422–440. https://doi.org/10.1038/s42254-021-00314-5

13. [13] M. Raissi, P. Perdikaris, G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics* 378 (2019) 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

14. [14] Nasim Rahaman, Aristide Baratin, Devansh Arpit, Felix Draxler, Min Lin, Fred Hamprecht, Yoshua Bengio, Aaron Courville. On the Spectral Bias of Neural Networks. *Proceedings of the 36th International Conference on Machine Learning (ICML)*, PMLR 97 (2019).

15. [15] Matthew Tancik, Pratul Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan T. Barron, Ren Ng. Fourier Features Let Networks Learn High Frequency Functions in Low Dimensional Domains. *Advances in Neural Information Processing Systems* 33 (NeurIPS 2020).

Information About Authors

Nurdaulet Izmailov – Researcher, DigitAlem LLP, Kazakhstan, Almaty, ORCID ID: 0009-0006-1417-1910

Meirambek Shaimerden – Researcher, DigitAlem LLP, Kazakhstan, Almaty, ORCID ID: 0009-0008-0336-9432

Azamat Toishybek – Lead Engineer, Institute of Metallurgy and Ore Beneficiation, master's degree, Kazakhstan, Almaty, ORCID ID: 0000-0002-7431-0103

Kaysar Kasymzhanov – Lead Engineer, Institute of Metallurgy and Mineral Processing, Kazakhstan, Almaty, ORCID ID: 0000-0001-8062-8655

Aralym Mukanaliyeva – Engineer, Institute of Metallurgy and Mineral Processing, master's degree, Kazakhstan, Almaty, ORCID ID: 0000-0001-7032-1764

Nurzhan Ultarakov – 2nd year master's student, Department of Computer Science, Al-Farabi Kazakh National University, Kazakhstan, Almaty, ORCID ID: 0009-0001-8171-5213

Alma Turganbaeva – Assistant Professor, Lecturer, Al-Farabi Kazakh National University, Department of Computer Science, Candidate of Pedagogical Sciences, Kazakhstan, Almaty, ORCID ID: 0000-0001-9723-4679

Submission received: 15 December, 2025.
Revised: 17 December, 2025.
Accepted: 17 December, 2025.