

Application of neural networks to predict the quality of iron ore concentrate based on flotation data

Vladislav Kukartsev^{1,2,3} and Ksenia Degtyareva^{1,2*}

¹Bauman Moscow State Technical University, 105005 Moscow, Russia

²Reshetnev Siberian State University of Science and Technology, 660037 Krasnoyarsk, Russia

³Russian State Agrarian University - Timiryazev Moscow Agricultural Academy (RSAU-MAA Named after K.A. Timiryazev), 127434 Moscow, Russia

Abstract. This paper presents a study aimed at developing and testing a neural network model for predicting the percentage of silica in iron ore concentrate obtained during flotation. The problem of precise control of the silica content is critical for the mining industry, since the quality of the final product and, accordingly, its market value depend on it. During the study, data was collected from the flotation plant, their preliminary processing was carried out, including standardization and elimination of missing values. The developed neural network model included two hidden layers and was trained on real data. The evaluation of the model quality showed high results, which was confirmed by the metrics of mean square error (MSE), mean absolute error (MAE) and coefficient of determination (R^2). Additionally, an analysis of the visualizations of the residuals and predicted values confirmed the accuracy and stability of the model. The results of the study demonstrate that the proposed model can be effectively used in production conditions to improve process control and improve product quality in the mining industry.

1 Introduction

Flotation is one of the key processes in the mining industry used to enrich iron ore. This method makes it possible to effectively separate useful minerals from waste rock and other impurities such as silica [1, 2]. The quality of the final iron ore concentrates, which is the main product of the flotation process, critically depends on the content of impurities such as silica. It is extremely important for mining companies to minimize the silica content in the final product, as this directly affects the quality of iron ore and, consequently, its market value [3-6].

Existing methods of concentrate quality control are often based on complex chemical analyses and require significant time and material costs. In production conditions, there is a need to develop methods for predicting the quality of concentrate in real time based on operational data obtained from technological equipment. Modern machine learning methods

* Corresponding author: sofaglu2000@mail.ru

offer extensive opportunities for predicting key parameters of production processes, which can significantly increase their efficiency and accuracy [7, 8].

The purpose of this study is to develop and test a machine learning model to predict the percentage of silica in the final iron ore concentrate produced during flotation. The main objective of the research is to create a neural network capable of predicting the silica content with high accuracy based on data obtained in real time from technological installations of the flotation process [9-11].

The present study is of great practical importance, since the proposed approach makes it possible to significantly improve the efficiency of flotation process management, ensuring high quality of the final product and reducing production costs.

2 Research methods

This study used an integrated approach, including data collection and processing, development and configuration of a machine learning model, as well as an assessment of its accuracy. Initially, data was collected from the flotation plant at the mining enterprise [12-14]. The dataset contained time series representing process parameters such as the composition of the initial ore, starch and amine consumption, ore pulp flow, pH and pulp density, as well as data on air flow and level in flotation columns. The target feature was the percentage of silica in the final concentrate.

A number of preliminary operations were performed to prepare the data for analysis. After that, the rows with missing values were deleted to avoid incorrect data during model training. In conclusion, all numerical features were standardized using the StandardScaler method, which made it possible to bring them to a single scale and improve the convergence of the model [15].

To predict the silica content, a neural network model was developed using TensorFlow and Keras libraries. The architecture of the model included an input layer, two hidden layers with a ReLU activation function and an output layer with linear activation, which corresponds to the regression problem. The model was compiled using the Adam optimizer and the mean_squared_error loss function. The learning process was carried out over 50 epochs using mini-batches of size 32. The data was divided into a training sample (80%) and a test sample (20%), while a validation sample consisting of 20% of the training data was used to evaluate the performance of the model [16-18].

After completing the training, the model was tested on a test sample to assess its quality. The metrics of mean square error (MSE), mean absolute error (MAE) and coefficient of determination (R^2) were calculated. The standard error showed the average square of the difference between the actual and predicted values, which allowed us to estimate the overall accuracy of the model [19, 20]. The average absolute error gave an idea of the average deviation of predictions from the actual data, expressed in the same units as the target feature. The coefficient of determination (R^2) showed how much of the variance of the target variable is explained by the model, and demonstrated the high accuracy of the model, approaching 1.

Various visualizations have been built to deeply analyze the operation of the model and identify possible areas of improvement. The graph of losses on the training and validation samples made it possible to evaluate the learning process and identify possible retraining of the model. The scatter plot of the predicted values versus the actual ones made it possible to visually assess the accuracy of the model [21, 22]. The histogram of residuals and the graph of the density of the distribution of residuals made it possible to analyze the distribution of model errors and identify possible anomalies [23-26].

The methods used made it possible to develop and evaluate a neural network model capable of predicting the silica content in iron ore concentrate with high accuracy [27-29].

This opens up opportunities for applying the developed approach in real production conditions to improve the quality and efficiency of enrichment processes.

3 Results

As a result of the research, a neural network model was developed and trained to predict the percentage of silica in iron ore concentrate based on data obtained from a flotation plant. The model was tested on an independent test sample, and its accuracy was evaluated using several metrics.

The Mean Squared Error (MSE) was 0.1558. This value indicates a fairly low mean square deviation of the values predicted by the model from the real data. A low MSE value indicates that the model is able to make accurate predictions and minimize large errors in predictions [30].

The average absolute error (Mean Absolute Error, MAE) was 0.2923, which also confirms the high accuracy of the model. MAE shows how much, on average, the predicted values deviate from the actual ones in absolute units. A MAE value of less than 0.3 indicates small average errors, which is important for tasks related to product quality control, where the accuracy of forecasts is critical [31].

The coefficient of determination (R^2) is 0.8768, which means that about 87.7% of the variance of the target feature is explained by the model. This result demonstrates a high degree of explanatory power of the model and confirms its suitability for predicting silica content in real production conditions [32].

Additionally, various visualizations were built to analyze the operation of the model. The graph of the change in the loss function on the training and validation samples, shown in Figure 1, showed stable convergence of the model during the training process, without signs of significant retraining [33]. This confirms that the model has been correctly configured and trained based on the data.

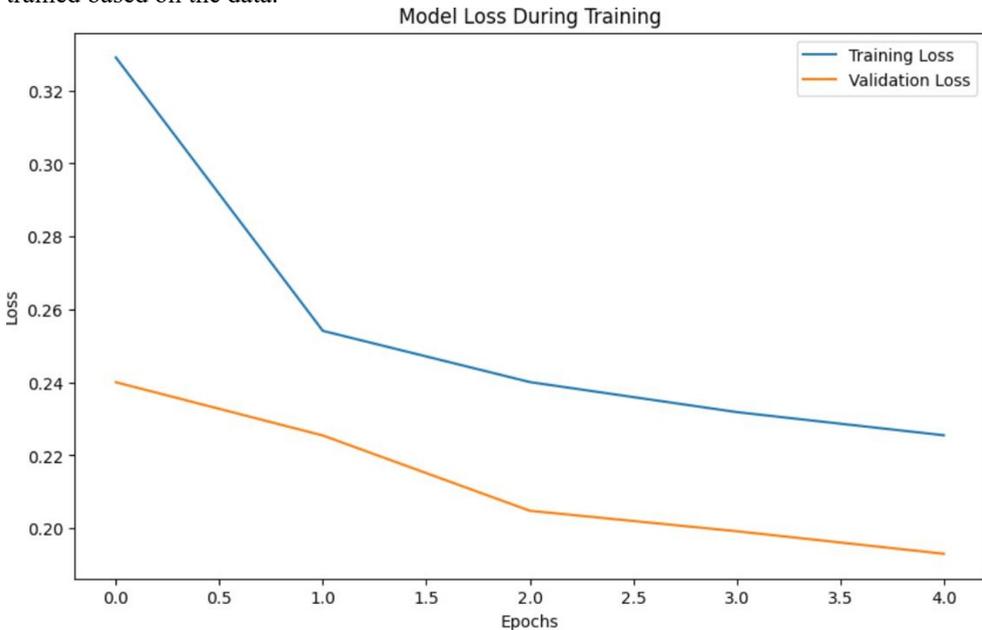


Fig. 1. Model loss.

The scattering diagram of the actual and predicted values shown in Figure 2 showed a high degree of coincidence of the model predictions with the real values. The points on the graph lie close to the perfect match line, which visually confirms the high accuracy of the model.

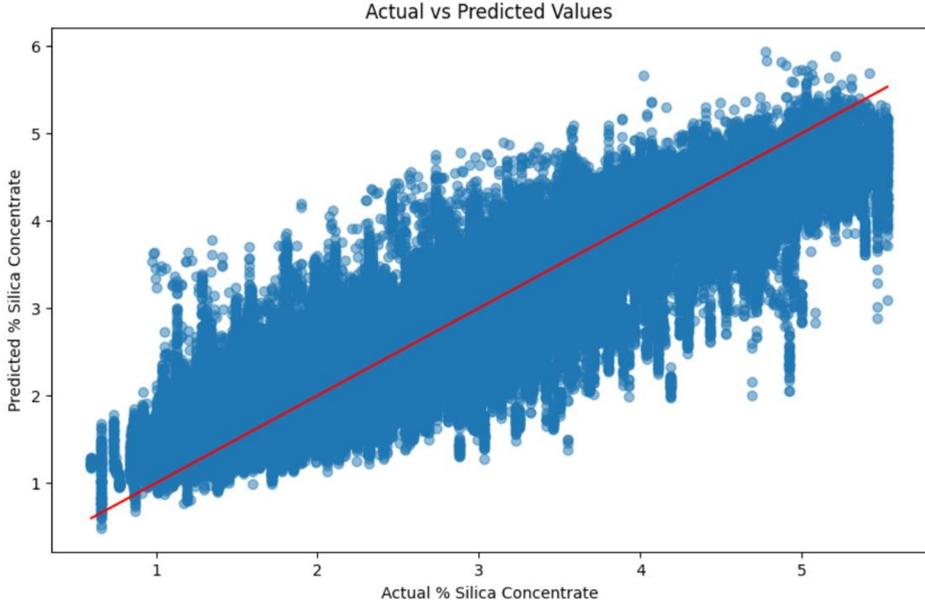


Fig. 2. The scattering diagram of the actual and predicted values.

The histogram of residues and the graph of the density of the distribution of residues, shown in Figures 3-4, demonstrated that the model errors are distributed near zero and do not have significant outliers. This indicates the absence of systematic biases in the predictions and confirms the correctness of the model.

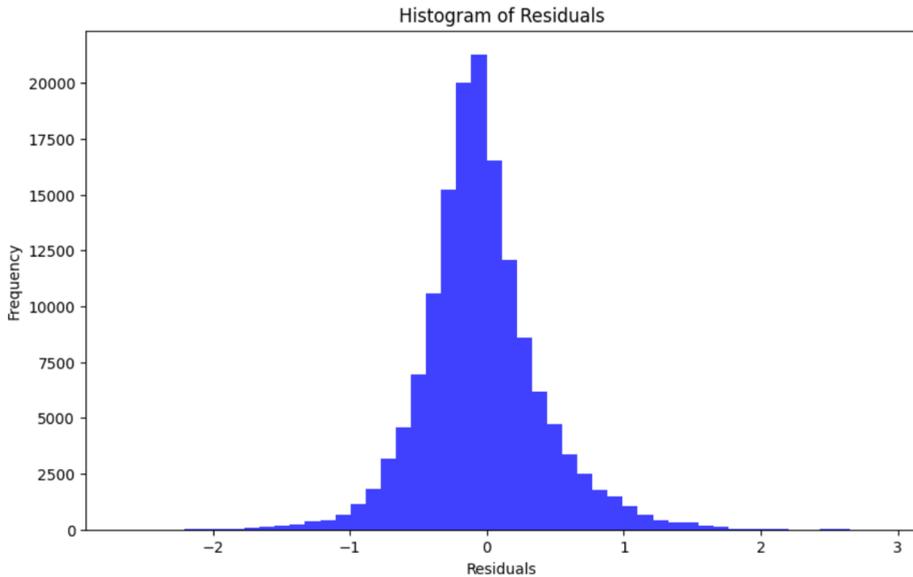


Fig. 3. The histogram of residues.

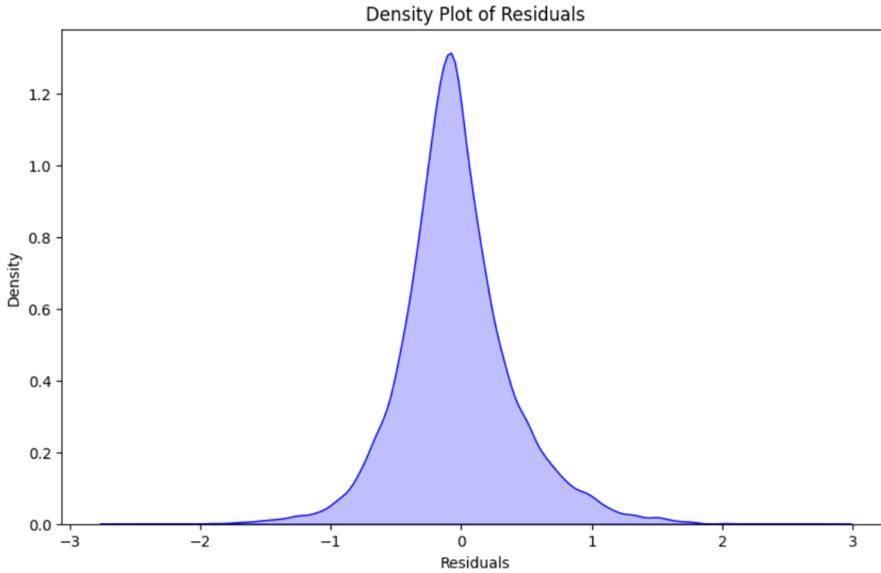


Fig. 4. The graph of the density of the distribution of residues.

Thus, the results of the study show that the developed neural network model successfully solves the problem of predicting the content of silica in iron ore concentrate with high accuracy and can be recommended for implementation in production processes to improve product quality and optimize the operation of flotation plants [34].

4 Conclusion

In the course of the study, a neural network model was successfully developed and tested to predict the silica content in the final iron ore concentrate based on data obtained from the flotation plant. The task facing the model was important for the mining industry, where accurate prediction of product quality is a critical factor for optimizing processes and increasing production profitability [35].

The analysis of the results showed that the proposed model has high accuracy. Quality assessment metrics such as mean square error (MSE), mean absolute error (MAE) and coefficient of determination (R^2) have demonstrated that the model is capable of predicting silica content with minimal error, which makes it an effective tool for industrial applications. The visualizations performed as part of the study confirmed that the model is not only accurate, but also resistant to various types of errors, such as overfitting and systematic bias [36].

One of the significant results of the work was the confirmation that the use of modern machine learning methods, such as neural networks, can significantly improve the quality of process control in the mining industry [37]. The introduction of the developed model into the production process will make it possible to more effectively control the quality of the concentrate, identify deviations in a timely manner and take the necessary measures to eliminate them, which ultimately will lead to lower costs and increased profits of the enterprise.

Further research may be aimed at optimizing the model, exploring other neural network architectures, as well as integrating the model into automated production management systems. It is also possible to expand the data set by including additional process parameters

or using methods for generating new features, which can increase the accuracy and stability of the model [38, 39].

Thus, the results of this study confirm the importance and prospects of using neural networks to solve forecasting problems in the context of iron ore processing, opening up new opportunities for improving the efficiency of mining enterprises.

References

1. Tynchenko, Vadim S., et al. AIP Conference Proceedings. Vol. **2700**. No. 1. AIP Publishing, 2023.
2. Chernykh, Nikita, et al. "*Comparative Analysis of Existing Measures to Reduce Road Accidents in Western Europe.*" 2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH). IEEE, 2023.
3. Volneikina, Ekaterina, et al. "*Simulation-Dynamic Modeling Of Supply Chains Based On Big Data.*" 2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH). IEEE, 2023.
4. Filina, Olga A., et al. *Energies* **17.1** (2023): 17.
5. Boychuk, Igor P., et al. *Energies* **16.24** (2023): 8101.
6. Golik, Vladimir I., et al. *Materials* **16.21** (2023): 7004.
7. Malozyomov, Boris V., et al. *Energies* **16.13** (2023): 5046.
8. Malozyomov, Boris V., et al. *Energies* **16.13** (2023): 4907.
9. Tynchenko, Vadim S., et al. *Journal of Physics: Conference Series*. Vol. **2373**. No. 6. IOP Publishing, 2022.
10. Martyushev, Nikita V., et al. *Materials* **16.13** (2023): 4529.
11. Kaung P. A., et al. "Principles for forming environmentally safe and economically effective sustainable development of geo resources. MIAB." *Mining Inf. Anal. Bull.* 2024;(7-1):159-175.
12. Panfilov, Iliia, et al. "*Modeling of the casting process for casting "Flywheel" of cast iron SCH20.*" 2024 12th International Conference on Smart Grid (icSmartGrid). IEEE, 2024.
13. Kravtsov, Kirill, et al. "*Creation of multi-link automatic parameter control systems at nuclear power plants.*" 2024 12th International Conference on Smart Grid (icSmartGrid). IEEE, 2024.
14. Degtyareva, Ksenia, et al. "*Manufacturing of 20XMFL Steel Bushing Casting.*" 2024 23rd International Symposium INFOTEH-JAHORINA (INFOTEH). IEEE, 2024.
15. Fedorova, Natalya, et al. *BIO Web of Conferences*. Vol. **116**. EDP Sciences, 2024.
16. Fedorova, Natalya, et al. *BIO Web of Conferences*. Vol. **113**. EDP Sciences, 2024.
17. Panfilov, Iliia, et al. *E3S Web of Conferences*. Vol. **531**. EDP Sciences, 2024.
18. Boyko, Andrei, et al. "*Simulation Model Of The Urban Population Taking Into Account The Influence Of The Manufacturing Sector Based On Big Data.*" 2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH). IEEE, 2023.
19. Lobkov, Konstantin, et al. "*Determination of the Degree of Impact of Natural Disasters on the Level of Migration of the Population by Simulation Modelling.*" 2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH). IEEE, 2023.
20. Martyushev, Nikita V., et al. *Energies* **16.2** (2023): 729.
21. Shutaleva, Anna, et al. *Sustainability* **15.4** (2023): 3011.

22. Martyshev, Nikita V., et al. *Materials* **16.9** (2023): 3490.
23. Kukartsev, Viktor Alekseevich, et al. *Metals* **13.2** (2023): 337.
24. Mikhalev, Anton S., et al. *Symmetry* **14.10** (2022): 2036.
25. Malozyomov, B. V., et al. "Investigation of properties of laminar antiferromagnetic nanostructures." *CIS IRON AND STEEL REVIEW Ore and Metals* Publishing House 27 (2024): 84-90.
26. Panfilova T. A., et al. "Treatment of wastewater from mining industrial enterprises from phenols." *MIAB. Mining Inf. Anal. Bull.* 2024;(7-1):72-82.
27. Savchenko, Makar, and Vadim Tynchenko. "*Unsupervised Production Machinery Data Labeling Method Based on Natural Language Processing.*" 2024 International Russian Smart Industry Conference (SmartIndustryCon). IEEE, 2024.
28. Tynchenko, Vadim, et al. *BIO Web of Conferences*. Vol. **113**. EDP Sciences, 2024.
29. Larichev, Pavel, Vadim Tynchenko, and Ivan Nekrasov. "*Application of Petri Nets for Modeling Ore Flows to Create Dynamic Management and Quality Control System in Mineral Resource Complexes.*" 2024 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM). IEEE, 2024.
30. Degtyareva, Ksenia, et al. "*Use of Computer Simulation Tools to Simulate Processes at the Foundry.*" 2024 23rd International Symposium INFOTEH-JAHORINA (INFOTEH). IEEE, 2024.
31. Degtyareva, Ksenia, et al. "*Automated System for Accounting of Customers and Orders.*" 2024 23rd International Symposium INFOTEH-JAHORINA (INFOTEH). IEEE, 2024.
32. Golik V. I. et al. The mechanochemical activation of leaching processes in a disintegrator. *MIAB. Mining Inf. Anal. Bull.* 2023;(11-1):175-189.
33. Panfilova, T. A. "To the concept of leaching metal-containing raw materials in the disintegrator." *MIAB* **11** (2023): 239-251.
34. Suprun, Elena, et al. *BIO Web of Conferences*. Vol. **84**. EDP Sciences, 2024.
35. Orlov, Vasiliy, et al. *E3S Web of Conferences*. Vol. **460**. EDP Sciences, 2023.
36. Kravtsov, Kirill, et al. *E3S Web of Conferences*. Vol. **458**. EDP Sciences, 2023.
37. Tynchenko, V. S., et al. *E3S Web of Conferences*. Vol. **458**. EDP Sciences, 2023.
38. Panfilov I. A., et al. "Prevention of air pollution during openpit mining of ore deposits." *MIAB. Mining Inf. Anal. Bull.* 2023;(11-1):252-264.
39. Panfilov I. A., et al. "The electrochemical technology of gold extraction from sulfide ores." *MIAB. Mining Inf. Anal. Bull.* 2023;(11-1):226-238.