International Journal of Chemical and Biochemical Sciences
(ISSN 2226-9614)

Journal Home page: www.iscientific.org/Journal.html

An Innovative Machine Learning Framework to Enhance the Design of the Bioleaching Process for Efficient Metal Recovery

Rohaila Naaz[1], Vijay Upadhye[2], Suhas Ballal[3], Swati Singh[4]

[1]*College of Computing Science and Information Technology, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India*

[2]*Department of Microbiology, Parul University, PO Limda, Vadodara, Gujarat, India*

[3]*Department of Chemistry, School of Sciences, JAIN (Deemed-to-be University), Bangalore, India*

[4]*Maharishi School of Engineering & Technology, Maharishi University of Information Technology, Uttar Pradesh, India*

Abstract

Metal recovery through bioleaching is a crucial aspect of sustainable resource management. The bioleaching process is acknowledged for its eco-friendly nature, but its efficiency is contingent upon intricate interdependencies among numerous variables. Existing bioleaching design approaches need more precision to optimize metal recovery efficiently. To develop a machine learning framework for the design of the bioleaching process by using Gradient Butterfly fused conditional Random Forest Regression (GBF-CRFR) integrates gradient boosting with dependent random forest techniques to create a robust predictive model. This approach leverages the strengths of both methods, allowing for the capture of complex dependencies and non-linear relationships in the bioleaching process variables. Collect the metal recovery dataset that contains density (PD), incubation temperature (T), percentage of energy substrates (SC) and pH control of the bioleaching fluid. The proposed method is compared with existing methods in terms of Root Mean Squared Error (RMSE) value is 0.6722, R square (R^2) value of 1.2142 and Mean Absolute Error (MAE) value of 0.2529. The findings underscore the potential of GBF-CRFR in advancing sustainable and efficient practices in bio hydrometallurgy.

Keywords: Metal Recovery, Bioleaching Process, Machine learning, Gradient Butterfly fused conditional Random Forest Regression (GBF-CRFR)

Full-length article **Corresponding Author*, e-mail: rohailanaaz2@gmail.com

1. Introduction

The growing need for effective and sustainable methods of recovering metals has generated curiosity about utilizing cutting-edge technology like machine learning (ML) [1]. To improve the bioleaching process, a bio hydrometallurgical technique recognized for its promise in removing metals from ores, the article presents a unique machine learning framework [2]. To improve and expedite metal extraction, lead to more resource and environmentally efficient methods, by incorporating ML into the conceptualization of bioleaching processes [3].

1.1. ML-Enhanced bioleaching

Microorganisms are used in the microbially aided process of bioleaching, which recovers metal from ores and waste from industries. Bioleaching process design has relied on manual optimization and empirical methods [4]. The constant change of microbial interactions and the intricacy of biological systems pose difficulties in attaining optimal circumstances for metal recovery. Acknowledging these difficulties, the use of machine learning offers a data-driven strategy for simulating complex interactions, resulting in enhanced process performance and design [5].

1.2. Importance of Bioleaching

Because bioleaching was less damaging to the environment, it can be used to extract metals from low-grade ores, which has made it more widespread [6]. Improvements in bioleaching procedures can have a significant impact on environmentally friendly mining methods. Thus, it is critical to investigate cutting-edge technologies like machine learning for accuracy and productivity in design [7]. To provide a machine-learning architecture specifically designed for the biological leaching process, highlight its potential to transform metal recovery. The research [8] used hybrid models to recover metal from coal fly ash, therefore examining the depletion of metal resources and environmental concerns associated with solid waste management. The proposal of a unique Metal Recovery Potential indicator highlighted the DAT sample's potential. Certain drawbacks underline the necessity of sustainable development of specific datasets, computational complexity, possible oversight and a need for more environmental thinking. With the use of artificial intelligence models and response surface methods, the research [9] optimized enzyme metal bioleaching from abandoned smartphone PCBs. The findings indicated high rates of metal extraction and the possibility of managing e-waste sustainably. Constraints include the bio-Fenton method's intricacy and its specificity to PCBs found in cell phones. The research [10] addressed bioleaching as an environmentally friendly technique for extracting precious metals from recyclable materials, assessing their efficacy and emphasizing elements like less testing and valuable designs. It identifies areas that require more investigation by highlighting limits, such as waste-specific issues and the intricate nature of solid waste mixtures. To enhance recovery of metal from discarded circuit boards are printed from cell phones, the study [11] used Penicillium as a simplicissimumin an inflatable column bioreactor. Outstanding recoveries of Cu and Ni were obtained using the approach; however, scaling up presents issues as well as microbe specificity.

The study [12] maximizes the gold recovery and silver from electronic waste by using Minitab software and response surface methods. The highest gold recovery (62.40%) was obtained by Bioleaching with C. violaceum, emphasizing the importance of glycine content, pulp density and oxygen concentration. There are several drawbacks, such as organism specificity and the complexity of e-waste composition. The research [13] created an environmentally friendly biotechnological method for separating copper and silver from waste deposited by electrolysis components while attaining high recovery rates. However, issues with scalability, metal-specific attention and financial viability need to be resolved. The article [14] states that non-biological reasons are the primary hindrance to the industrial biomining process's underutilization of bioleaching. For operations to remain economically viable, it emphasizes the necessity of process design enhancements that focus on sluggish subprocesses. To extract value from waste materials and mineral deposits using heap bioleaching, they proposed both active and passive procedures; nevertheless, more research and validation are required to address practical issues. With the goal of lowering the number of hazardous metals and lessening the effects on the environment, the research [15] proposed a bio-hydrometallurgical approach for extracting gold and nonferrous metals from old pyritic floatation tailings and metallurgical slag. The method, which combines ferric leaching with microbial consortia bioleaching, has the potential to be advantageous for the economy and the environment. With an emphasis on low-grade PCBs, the research [16] assessed the viability of bioleaching PCBs utilizing acidophilic bacteria for base metal recovery. The results demonstrated the benefits of colonization, commercial application and microbial adaptability; nonetheless, issues such as acid consumption require resolution. Optimization, thorough assessments and scaling up are essential for economic viability. The research [17] comparing the suggested technique to competing methods, it was more favourable for precise and dependable forecasts since it has a reduced RMSE, suggesting higher accuracy in forecasting maximum metal recovery. It focused on difficulties with tank biological leaching at Mondo Mineral and the Terrafamemine, addressing concerns with the amount of arsenic in sulphide concentrate. The results demonstrate the flexibility and feasibility of bioleaching techniques for complicated mineral deposits and concentrates, but they call for more research into the financial implications and scaling issues. The study [18] demonstrated the sustainability of the bioleaching potential of native Bacillus sp. ISO1 for metal recovery. The best precursor was found to be glycine and the addition of methionine increases the generation of cyanide lixiviant, suggesting the possibility of large-scale industrial operations. They highlighted the flexibility and viability of tank and heap biological leaching methods for complicated ore bodies and concentrates. It examined at the practical implementation of these methods for nickel and cobalt recovery from Selphie ores and concentrate in Finland.

The article was divided into four sections: Materials and Methods; Results as well as Discussions; and Conclusion. It aims to provide a systematic exploration of bioleaching process design using the GBF-CRFR framework.

2. Materials and methods

For accurate metal recovery prediction, GBF-CRFR combines probabilistic random forest regression with gradient butterfly optimization. By combining the benefits of conditional tree structures and gradient information, the hybrid model improves accuracy in single-paragraph scenarios. The suggested approach's framework is depicted in Figure 1.

2.1. Gradient Butterfly fused conditional Random Forest Regression (GBF-CRFR)

By combining predictions from many ML algorithms, Conditional Random Forest Regression (CRFR), a ML supervised technique, use collaborative modelling to provide forecasts that are more reliable. When it comes to Optimal Metal Recovery prediction, CRFR builds an extensive number of decision trees during training. The average of these trained trees' forecasts determines the Metal Recovery output. A CRFR exhibits high prediction accuracy in a short training period. The primary advantage is an increase in test accuracy for metal recovery forecasts, together with a decrease in the expenses related to training, storing and retrieving conclusions from various models. This method creates numerous decision trees, each of which runs independently of the others, from concurrent datasets. For forecasting maximum metal recovery, the resultant CRFR model is reliable and accurate. The h stands for the number of independent regression trees that were built with the metal recovery data input vector y. $k_h(y)$ is the mean prediction of metal recovery obtained from h regression trees. The following equations (1) and (2) are used to get the mean squared error for out-of-bag metal recovery data.

$$\text{CRFR Prediction} = \frac{1}{4}\sum_{h=1}^{h} k_h(y) \tag{1}$$

$$MSE_{OOB} = \frac{1}{n}\sum_{j=1}^{n}(x_j - \overline{x_{jOOB}})^2 \tag{2}$$

To create forecasts and develop a model, the CRFRis utilized for a labelled training set that is intended for metal recovery predictions. To control variation, the technique combines the bagging principles with a random parameter selection process for decision tree construction. The capacity of the CRFR technique to intuitively determine the significance of variables in both regression and classification scenarios is a significant benefit. To locate possible partners and food sources, butterflies depend on their sense of smell. Every butterfly releases a fragrance (fragrance,f) that other butterflies can detect the metal recovery. When it involves metal recovery optimization, the butterfly, in its optimal posture, emits a more potent scent. The fragrance is designed to be consistent with the fitness value (J) expressed in equation (3).

$$f = dJ^b \tag{3}$$

The exponentiation value for power is represented by a, the stimulus intensity is indicated by J and the sensory component of metal recovery is represented by d. b is a significant parameter that affects how rapidly the method converges. Its values range from zero to one. Both of these factors have a considerable influence on the algorithm; therefore, the value of d is quite significant. With the position of the particle as an input for optimization, the stimulus intensity is calculated by assessing the metal recovery from the objective function. Some butterflies in the population recognize the most alluring scent during the exploitation period and change their location accordingly. An amended stance is the result of this process and it can be stated as follows: equation (4).

$$y_j^{t+1} = y_j^t + (q^2h^* - y_j^t)dJ^b \tag{4}$$

The most effective metal recovery technique found worldwide during the iteration is represented by h^* in the case, while r is a randomly generated value between zero and one. The population's surviving butterflies wander at random, making up the exploration stage. The following equation (5) characterizes the movement.

$$y_j^{t+1} = y_j^t + (q^2y_j^t - y_h^t)dJ^b \tag{5}$$

Here y_j^t and y_h^t represent the locations of two butterflies in the same swarm and q are an arbitrary value between zero and one. Based on the chance of a switch (q) value, a random decision is made to move between the search and extraction phases.

3. Result and Discussion

Examining different aspects of metal recovery data is essential when using machine learning algorithms so that statistical analysis can be performed to find trends and patterns. A statistical description of the input variables and dataset responses is shown in Table 1. A total of 29 samples comprised the dataset [19], which was used for modelling. Data sampling was made less complicated by a random sampling technique that divided the data into three categories: training, testing and validation. Due to dataset restrictions, this research employs an 80% training to 20% testing split, whereas the standard break is 60% training to 40% testing and validation. The dataset has all the data. Density (PD), incubation temperature (T), percentage of energy substrates (SC) and pH is regulator of the bioleaching fluid. The yield proportions of zinc and magnesium are the two response variables included in the output layer.

3.1. Analyzing hyperparameter tuning for GBF-CRFR

A crucial parameter for GBF-CRFR is the number of estimators, Figure 1 shows the tuning technique, with emphasis on MAE and MSE. The investigation indicates that the quantity of forecasting has a significant outcome on MSE and MAE. However, when modelling Figure 2 (a) and (b) Mn and Zn (c) and (d), respectively, the best values of the 400 and 200 estimators were found.

3.2. Predictive Capability of GBF-CRFR Modelling for Zinc and Manganese

Regarding Zn, there are discrepancies as the expected values show distribution around the diagonal line. The objective is for the predicted data points in plots to compare actual and predicted values to connect with the diagonal line strongly. Under forecasting is represented by points below a diagonal while over forecasting is indicated by points above the diagonal. As compared to the Zn model, the Mn model predictions produced by the GBF-CRFR model show superior alignment with the diagonal. The difference between Mn and Zn data can be explained by the former's more minor standard deviation and range. Compared to Zn data, linear regression is more successful in explaining the variation in the data. The accuracy of predictions of the GBF-CRFR strategy is shown in Figure 3.

3.3. Performance Analysis of GBF-CRFR

In this section, the suggested strategy is compared to the related current methods [20], such as Random Forests Firefly Algorithm (RF-FFA), Random Forests (RF) and Support Vector Machine (SVM) in matrices of Root Mean Squared Error (RMSE), R-squared(R^2) and Mean Absolute Error (MAE).

- ***R-squared***

One statistical measure that is used to evaluate a regression model's quality of fit is the R-squared (R^2) value. R^2 Measures the percentage of the variation in the variable that is dependent (maximum metal recovery) can be accounted by the independent variable or predictors in the regression model when predicting full metal recovery. As demonstrated by Figure 4, its improved performance in R-squared values shows that the suggested method outperforms current approaches in forecasting maximum metal recovery, underscoring its essential application in the setting of metal recovery prediction. Table 2 represents the comparison matrix of RMSE, MAE and R^2. This exceeds the R-squared values of the present techniques. The equation (6) for R^2 is:

$$R^2 = 1 - \frac{Sum\ of\ Squared\ Residuals}{Total\ sum\ of\ Squared} \tag{6}$$

Where,

In a regression model, the total of squared discrepancies between the actual and projected values is referred as the total of squared residuals or the total of squares.

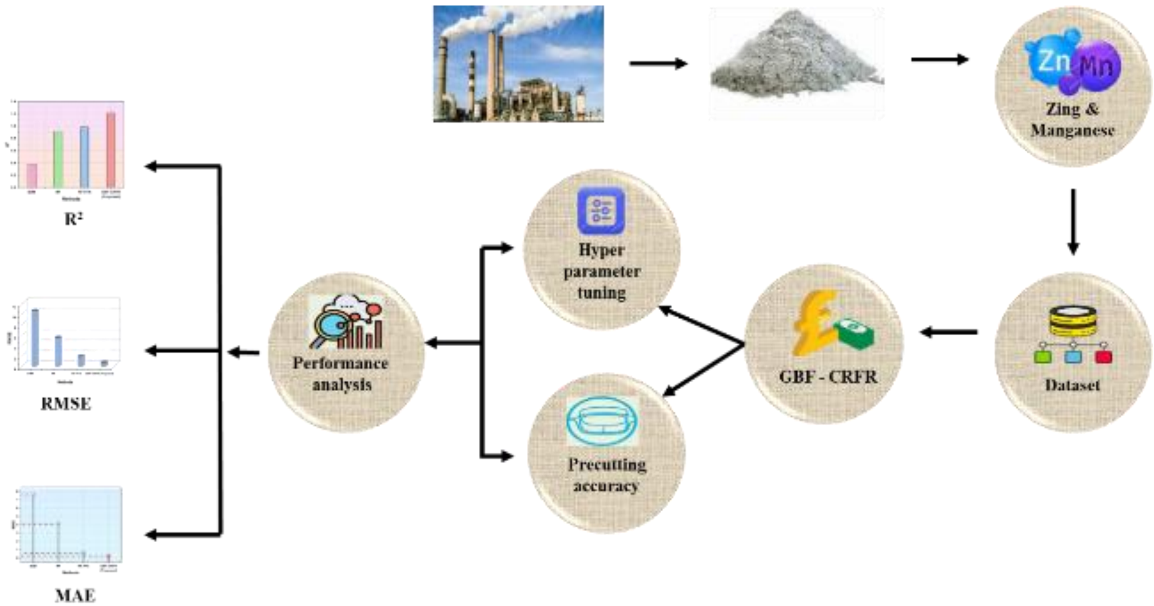

Figure 1. Framework of the suggested approach

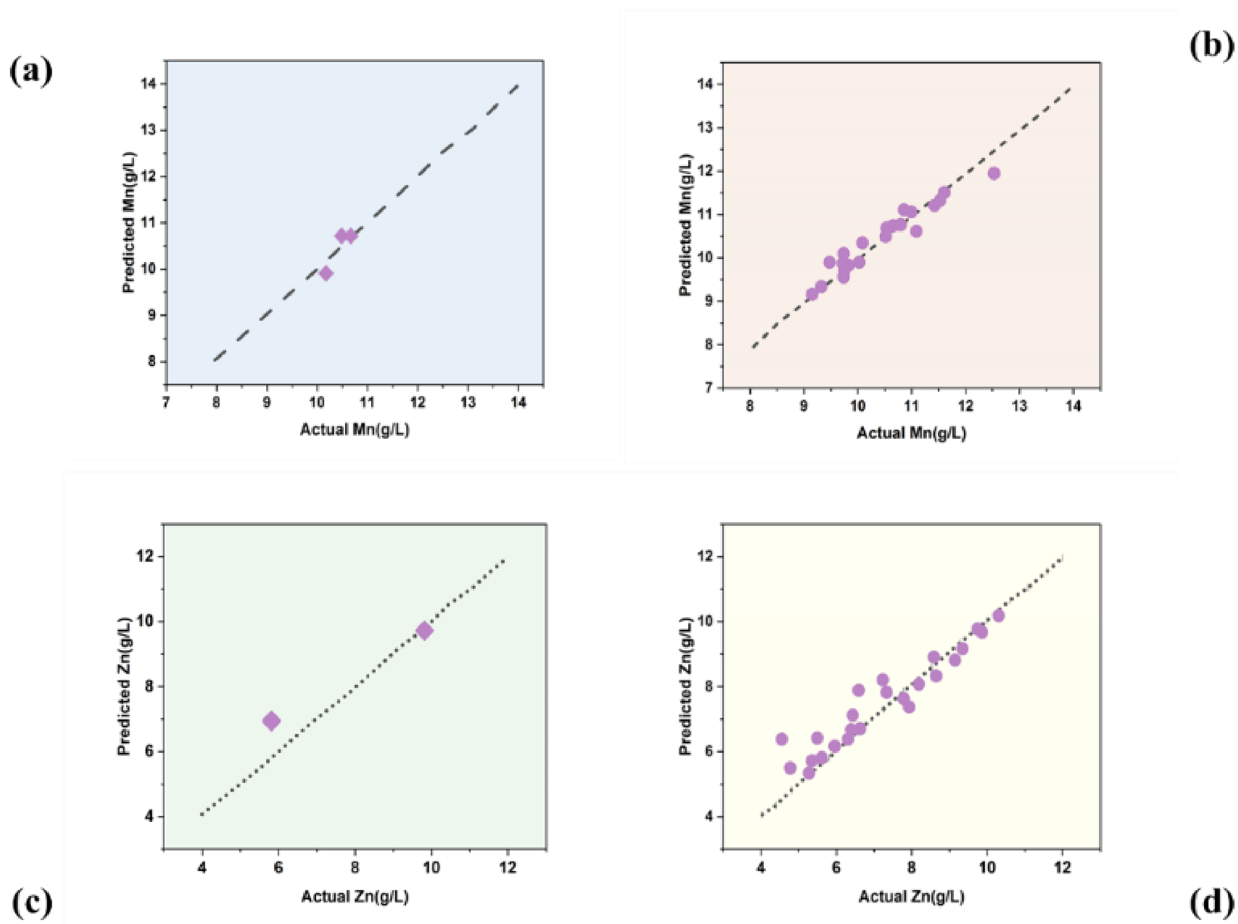

Figure 2. Hyperparameter Tuning for GBF-CRFR

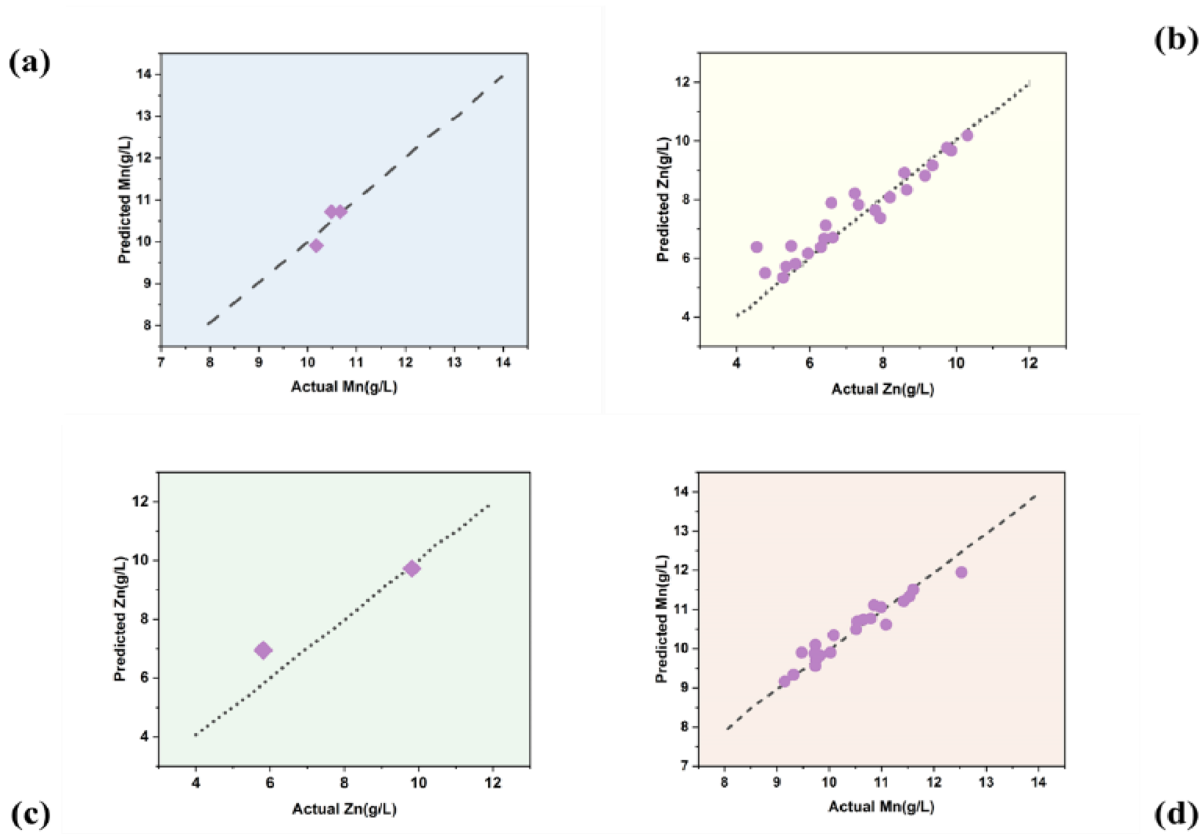

Figure 3. Predictions of accuracy GBF-CRFR

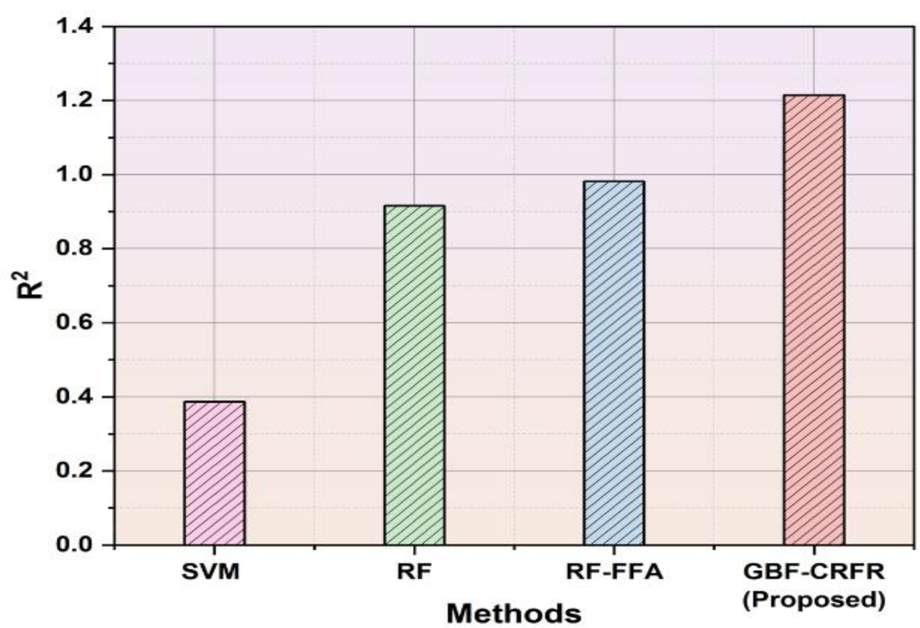

Figure 4. Result of R^2

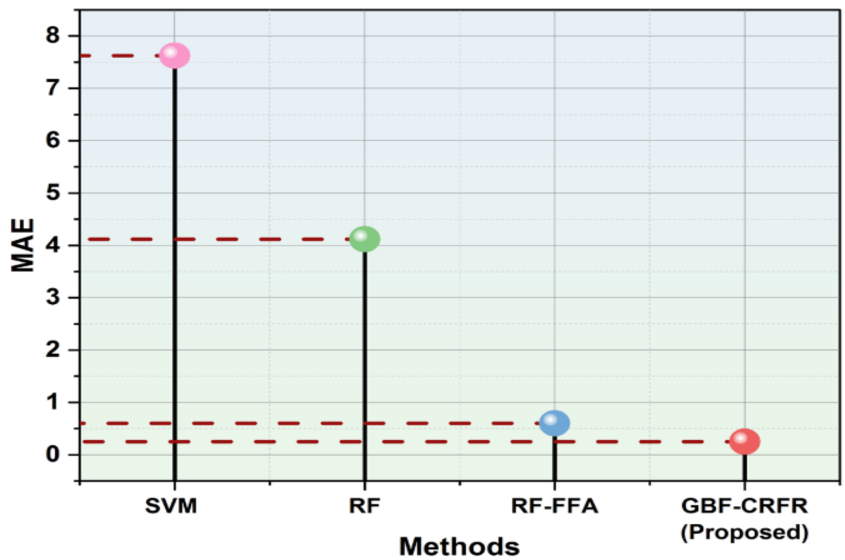

Figure 5. Graphical representation of MAE

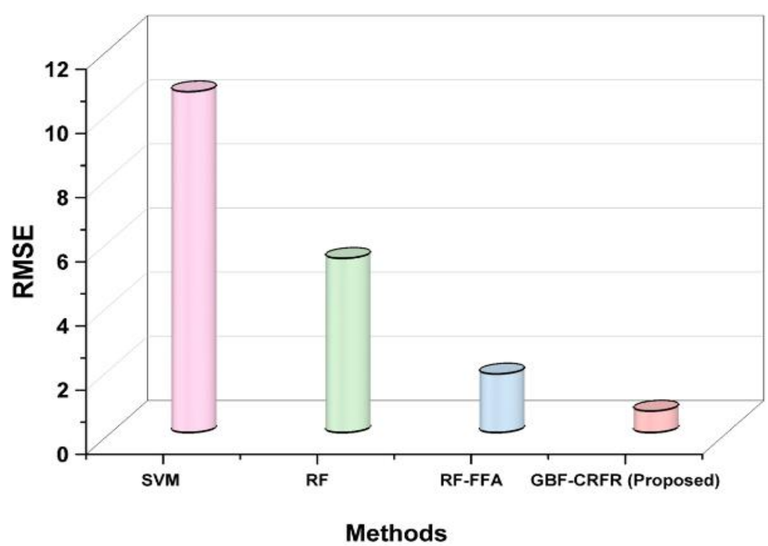

Figure 6. Outcomes of RMSE

Table 1. Description of samples

Run No.	pH	SC	PD (%)	T (°C)	Observed Mn	Observed Zn
1	1.8	28	9	32.5	11.4	9.4
2	1.8	36	9	32.5	10.9	9.1
3	2.2	28	9	32.5	10.7	7.9
4	2.2	36	9	32.5	10.6	6.4
5	1.8	28	9	37.5	11.6	8.2
…	…	…	…	…	…	…
29	2.0	32	10	35	10.8	9.9

Table 2. Outcomes of the study

Methods	R^2	MAE	RMSE
SVM	0.3869	7.623	10.6281
RF	0.9162	4.1174	5.4325
RF-FFA	0.9817	0.6015	1.8279
GBF-CRFR (Proposed)	1.2142	0.2529	0.6722

- ***Mean Absolute Error (MAE)***

The absolute difference is provided by the Mean Absolute Error (MAE) metric, which is used to quantify the average difference between expected and actual values, especially when estimating maximum metal recovery. Figure 5 represents the result of MAE. The suggested strategy achieves a lower MAE than current methods and exceeds existing approaches in forecasting maximal metal recovery. This means a greater degree of precision in the maximum metal recovery prediction, which makes the proposed system particularly advantageous for raising the accuracy of top metal recovery forecasts. The equation (7) for MAE is as follows:

$$\mathrm{MAE} = \frac{1}{n}\sum_{j=1}^{n}\left\lceil B_j - Q_j \right\rceil \quad (7)$$

The quantity of observations is n. Maximum metal recovery B_j is the actual value of the parameter that is predicted for comment. The expected value for observation j is Q_j.

- **Root Mean Squared Error (RMSE)**

A popular statistic for assessing forecast accuracy, uncommonly when projecting maximal metal recovery is RMSE. It penalizes more fantastic mistakes more severely by measuring the mean difference between the expected and real figures, squared. Figure 6 depicts the outcome of RMSE. When compared to other approaches, the suggested method is advantageous for precise and dependable predictions since it exceeds existing methods to forecast maximum metal recovery. Its lower RMSE indicates higher accuracy in predicting this recovery. The equation (8) for RMSE is as follows:

$$\boldsymbol{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left\lceil B_j - Q_j \right\rceil} \quad (8)$$

3.4. Irrigation

In complicated bioleaching datasets, RF and SVM are susceptible to overfitting, which reduces their prediction accuracy. Because firefly algorithms are stochastic, RF-FFA can have trouble converging effectively. The GBF-CRFR improves prediction performance by identifying complex linkages and non-linear dependencies, which allows it to overcome the drawbacks of conventional techniques.

4. Conclusion

Examine the metal recovery through bioleaching and utilize the GBF-CRFR machine learning architecture to improve the accuracy of the bioleaching process design. GBF-CRFR, a robust predictive model that can capture intricate dependencies and non-linear connections among bioleaching process variables, was developed by combining gradient boosting and conditional random forest approaches. The suggested method was compared with previous techniques in terms of R-squared (1.2142), MAE (0.2529) and RMSE (0.6722) utilizing a dataset that included density, incubation temperature, % of energy substrates and pH control. The outcomes showed how well GBF-CRFR performed, underscoring its potential to promote effective and sustainable biohydrometallurgical processes. Its shortcomings, such as dataset breadth, are acknowledged while appreciating its accomplishments. Subsequent investigations ought to delve into the enlargement of datasets and the enhancement of the model to ensure its broader relevance in enhancing bioleaching procedures for varied metal retrieval situations.

References

[1] M. Elahi, S.O. Afolaranmi, J.L. Martinez Lastra, J.A. Perez Garcia. (2023). A comprehensive literature review of the applications of AI techniques through the lifecycle of industrial equipment. Discover Artificial Intelligence. 3(1): 43.

[2] S. Chen, Y. Ding. (2023). Tackling Heavy Metal Pollution: Evaluating Governance Models and Frameworks. Sustainability. 15(22): 15863.

[3] X. Wei, D. Liu, W. Huang, W. Huang, Z. Lei. (2020). Simultaneously enhanced Cu bioleaching from E-wastes and recovered Cu ions by direct current electric field in a bioelectrical reactor. Bioresource Technology. 298: 122566.

[4] Y. Hua, S. Zhou, Y. Huang, X. Liu, H. Ling, X. Zhou, S. Yang. (2020). Sustainable value chain of retired lithium-ion batteries for electric vehicles. Journal of Power Sources. 478: 228753.

[5] W. Gu, J. Bai, L. Lu, X. Zhuang, J. Zhao, W. Yuan, J. Wang. (2019). Improved bioleaching efficiency of metals from waste printed circuit boards by mechanical activation. Waste Management. 98: 21-28.

[6] R. Rathna, E. Nakkeeran. (2020). Biological treatment for the recovery of minerals from low-grade ores. In Current Developments in Biotechnology and Bioengineering. 437-458.

[7] S. Vyas, S. Das, Y.P. Ting. (2020). Predictive modeling and response analysis of spent catalyst bioleaching using artificial neural network. Bioresource Technology Reports. 9: 100389.

[8] M. Wu, C. Qi, Q. Chen, H. Liu. (2023). Evaluating the metal recovery potential of coal fly ash based on sequential extraction and machine learning. Environmental Research. 224: 115546.

[9] A. Trivedi, S. Hait. (2023). Metal bioleaching from printed circuit boards by bio-Fenton process: Optimization and prediction by response surface methodology and artificial intelligence models. Journal of Environmental Management. 326: 116797.

[10] T. Naseri, V. Beiki, S. M. Mousavi, S. Farnaud. (2023). A comprehensive review of bioleaching optimization by statistical approaches: recycling mechanisms, factors affecting, challenges, and sustainability. Royal Society of Chemistry Advances. 13(34): 23570-23589.

[11] S. Nili, M. Arshadi, S. Yaghmaei. (2022). Fungal bioleaching of e-waste utilizing molasses as the carbon source in a bubble column bioreactor. Journal of Environmental Management. 307: 114524.

[12] W.N.F. Abdol-Jani, F. Suja, S.I.S. Jamaludin, N.F. Mohamad, N.H. Abdul-Rani. (2023). Optimization of precious metals recovery from electronic waste by Chromobacteriumviolaceum using response surface methodology (RSM). Bioinorganic Chemistry and Applications.

[13] E. Pakostova, A. Herath. (2023). A Bioleaching Process for Sustainable Recycling of Complex Structures with Multi-Metal Layers. Sustainability. 15(19): 14068.

[14] J. Petersen. (2023). From understanding the rate limitations of bioleaching mechanisms to improved bioleach process design. Hydrometallurgy. 221: 106148.

[15] N. Fomchenko, M. Muravyo. (2020). Sequential bioleaching of pyritic tailings and ferric leaching of nonferrous slags as a method for metal recovery from mining and metallurgical wastes. Minerals. 10(12): 1097.

[16] A. Hubau, C.G. Bryan. (2022). Metal Recovery from E-wastes. In Biomining Technologies: Extracting and Recovering Metals from Ores and Wastes. 239-259.

[17] M. Gericke, J.W. Neale, P. Määttä. (2022). Biomining in Finland: commercial application of heap and tank bioleaching technologies for nickel recovery. In Biomining Technologies: Extracting and Recovering Metals from Ores and Wastes. 209-228.

[18] P. Thakur, S. Kumar. (2023). Augmentation in bioleaching potential of indigenous Bacillus sp. ISO1 for metals recovery from waste computer-printed circuit boards. International Microbiology. 1-11.

[19] C. Ruhatiya, S. Shaosen, C.T. Wang, A.K. Jishnu, Y. Bhalerao. (2020). Optimization of process conditions for maximum metal recovery from spent zinc-manganese batteries: Illustration of statistical based automated neural network approach. Energy Storage. 2(3): e111.

[20] R. Cook, K.C. Monyake, M.B. Hayat, A. Kumar, L. Alagha. (2020). Prediction of flotation efficiency of metal sulfides using an original hybrid machine learning model. Engineering Reports. 2(6): e12167.