International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

ML-Based Mineral Exploration through Metallogenic Maps

Vrushika Naik[1], Akshita Raut[2], Mansi Limbad[3], Saniket Kudoo[4]

[1](Computer Engineering Department, VIVA Institute of Technology, India)
[2](Computer Engineering Department, VIVA Institute of Technology, India)
[3](Computer Engineering Department, VIVA Institute of Technology, India)
[4](Professor, Computer Engineering Department, VIVA Institute of Technology, India)

ABSTRACT:

Mineral locations become more difficult day by day to locate in geographical locations. Despite the existence of traditional means, the acquisition process is still not faster. Finding economically viable mineral reserves has become most difficult. Mineral exploration requires significant time with high investment, constant cash flow and inherent high risk. The iterative process of collecting different datasets, which are geologically interpreted, can take a long time during exploration. Large amounts of data are collected and analyzed, often without significant mineral discoveries. As a result, processes need to increase finding rates and shorten the traditional exploration life cycle, which identifies mineral locations by overlaying numerous layers of geoscientific data in GIS software.

The proposed project presents a systematic review of efforts devoted to the development of machine learning-based solutions to better use mineral data in mining and mineral studies. This project is proposed for better mineral exploration by creating ML-based mining exploration models by combining geographical data. The Geological Survey of India's BHUKOSH portal and other sources contain a variety of geological datasets. Machine Learning is a data-driven algorithm that automates clustering, prediction and classification of data. There are two types of techniques in machine learning which are supervised and unsupervised learning. This project works on three different supervised machine learning algorithms.

The proposed project tries to analyze the problems and leverage the country's accessible geological datasets by delivering an ML based solution for mineral exploration through the construction of a metallogenic model for better mineral exploration.

Keywords – *Machine Learning, Metallogenic model, Mineral exploration, Geoscientific, Geo-locations*

I. Introduction

Minerals are said to hold a huge part of India's economic growth. But even with so many advancements in technology, their extraction had always been a tedious, costly, and time-consuming process. Before mineral extraction, mineral exploration is the first important step which traditionally requires huge datasets of various geographical features to be combined together in layers to obtain or predict the presence of a mineral based on these properties. This step consumes a large amount of time in the complete process. Therefore, to reduce this time and make the process of mineral exploration efficient, this proposed project aims to deliver an ML-based solution for the same. This would work on the basis of different datasets combined together with respect to the locations. Mineral locations would be visualized as per the geographical conditions required for any mineral formation. For this project, one mineral is considered and that is Gold.

This project aims to provide a user-friendly web application that would allow users to visualize the mineral (Gold) on a map. Users will also have an option to download required datasets from BHUKOSH portal of India and give them as input to visualize the mineral of other states.

II. Review of Literature Survey

Shuai Zhang, Emmanuel John M. Carranza, et.al. [1] variable significance and partial dependence plots, which enable the interpretability of RF modeling, were introduced to examine the spatial signature between the mineralized location-based training set and the outlier-based training set in the training RF algorithms for mineral prospectivity mapping (MPM). ROC curves and correlation analysis for prediction maps by RF modeling based on different training sets. Outliers have the potential to be used as positive training samples in data-driven MPM, such as known mineralized locations in a given study area.

Yan Lv, Laijun Lu, et. al. [2] in this study, a machine learning-based study on geological and mineral energy and mineral energy classification was carried out, and a three-dimensional geological entity model was constructed. After the cube of the entity model, a three-dimensional quantitative prediction model of the study was determined by the extraction of metallogenic information. Under the guidance of the quantitative prediction model, the wireless sensor algorithm was used to perform 3D quantitative prediction based on the wireless sensor algorithm. The training accuracy of the model and the testing accuracy of the model were over 96%, which proved the accuracy of the RF model construction and achieved satisfactory results.

Xishihui Du, Kefa Zhou, Yao Cui et. al. [3] this study used a hybrid support vector machine (SVM) model with a genetic algorithm (GA) to discriminate between prospective and non-prospective areas for Au deposits in Karamai, northwest China. A key process in the implementation of the GA-SVM model was the selection of the training dataset, specifically the 'non-mineralized' locations. In complex geological environments, it is impossible to identify non-mineralized locations; Thus, point pattern analysis is a useful measure to determine the optimal distance at which non-mineralized locations can be randomly selected based on the selection criteria. The performance of the GA-SVM model was evaluated using both F1 score and spatial efficiency to distinguish potential areas in the study area. The best prospectivity model predicts 95.83% of the known Au deposits.

Nan Lin, Yongliang Chen, Haiqi Liu, et. al. [4] mineral prospectivity mapping models are created under conditions where a search space needs to be set, and bat and firefly swarm intelligence optimization algorithms are combined with various machine learning models to automatically find hyperparameters. Compared to traditional optimization algorithms, BA and FA are free to switch between global and local optimization processes and have more opportunities to find global optimal hyperparameters of machine learning models. BA and FA have different improvement effects on MLP, AdaBoost and OCSVM models. The accuracy of machine learning models is greatly improved after hyperparameter optimization, which indicates that the model hyperparameter optimization is effective and reliable for the application of machine learning methods. ROC and P-A curves are applied to quantitatively evaluate the predictive performance of mineral potential mapping models. The evaluation results show that the AUC value can effectively measure the accuracy of different models, but it is not the only index. The value of the P/A curve intersection is calculated by the P-A curve. The higher the value, the more accurate the metallogenic prediction.

Yihui Xiong, Renguang Zuo, et. al. [5] traditional deep multilayer neural networks are usually not successfully trained, mainly because the training algorithm, a standard gradient fit from random initialization, performs poorly in multilayer neural networks. Compared to traditional neural networks, deep learning algorithms use different initialization and training schemes, such as unsupervised pre-training and fine-tuning strategies. This enables better high-level extraction of complex features and data representation from large volumes of data and makes it a powerful tool for big data analytics. These advantages suggest that deep learning methods can be effectively used to identify geological anomalies and patterns.

Yongliang Chen, Wei Wu, et. al. [6] A bat-optimized one-class support vector machine is developed by combining a one-class support vector machine with the bat algorithm. The combined model can automatically optimize the parameter values of a one-class support vector machine, thereby improving the model's performance in mineral prospectivity mapping. The method proposed in this paper only needs to set the search space, and then the algorithm automatically finds the optimal parameters. Compared to the trial and error method, the proposed method has more chances to find the global optimal parameters. A bat-optimized one-class support vector machine requires a certain number of known mineral deposits in the study area, which can be used to define true positive and negative points (cells), and can be used as ground truth data for the receiver operating characteristic curve analysis. The bat-optimized one-class support vector machine can be used as a semi-supervised machine learning model to handle anomaly detection problems in other application areas.

Yan Niu, Jun Zhao, Zhiyuan Li, et. al. [7] The mining of mineral resources directly affects the national economic development, to be able to do good work in geological and mineral mining, adequate geological research and the reasonable use of mineral exploration technology are necessary to ensure that mineral exploration work can meet the needs of modern development. Remote sensing technology plays a very important role in many fields, and the use of this technology in geological research has greatly promoted the development of geological research work in China. The scientific application of modern information imaging technology can provide better conditions and more convenience for geological exploration and geological prospecting work. Remote sensing technology in geological prospecting can significantly improve the comprehensive level of geological exploration, strengthen the accuracy of geological forecasting based on advanced technology, and thus promote the smooth development of geological research and potential work in China.

Chandragiri Sandeep, Yellepeddi Srikar, Kodali Rajani, et. al. [8] convolutional Neural Networks (CNN) have achieved great results in the field of image recognition. We propose a technique that uses a convolution neural network to predict mineral precision using individual mineral images. The proposed approach has 93.6 percent accuracy. The objective is to work on the flexibility of the mineral by examining the classification of information from photos taken in different environmental conditions and at different times.

Theodoros Evgeniou, Massimiliano Pontil, et. al. [9] Support vector machines have recently been developed within the framework of statistical learning theory and have been successfully applied to a number of applications, ranging from time series prediction to face recognition to biological data processing for medical diagnosis. SVMs are trained by solving a constrained quadratic optimization problem. In others, this implies that there is a unique optimal solution for each choice of SVM parameters. Training many local SVMs instead of one global SVM can significantly improve the performance of a learning machine.

Vrushali Y Kulkarni, Dr Pradeep K Sinha, et. al. [10] Random Forest is an ensemble supervised machine learning technique. Machine learning techniques have applications in the field of data mining. Random forest has enormous potential to become a popular technique as its performance has been found to be comparable to bagging and boosting techniques. Decision trees are commonly used for supervised machine learning. Random Forest uses a decision tree as a base classifier. The randomization in random forest is present in two ways: 1) Random sampling of data for bootstrap samples as it is done in bagging and 2) Random selection of input features to build individual support decision trees. Random Forest runs efficiently on large databases, can handle thousands of input variables without variable deletion, estimates important variables, generates an internal absolute estimate of normalization error as the forest grows. This paper presents a classification of random forest algorithms and analyzes various techniques based on random forest algorithms. This analysis presented as a comparison and will serve as a guide for future research related to random forest classifiers to follow.

Rish, et. al. [11] Naive Bayes classifiers make learning very easy by assuming that the features are independent given class. In this paper, the goal is to understand the data characteristics that affect the performance of the naive bayes. This naive bayes approach uses Monte Carlo simulations that allow a systematic study of classification accuracy for many different classes of randomly generated problems. Analyze the effect of distribution entropy on classification error in this, showing that some nearly deterministic, or low-entropy, dependence performs better than a naive basis. This paper shows that Naïve Bayes works best in two cases which are completely independent features and functionally dependent features. Naive Bayes accuracy is not directly related to the degree of feature dependencies measured as class-conditional mutual information between features. Instead, the loss of information containing features about the class is a better predictor of accuracy when assuming a Naive Bayes model.

III. Analysis

The analysis table summarizes the research paper on ML-Based Mineral Exploration through Metallogenic Maps.. Below is a detailed description of the research papers that have been studied.

Table 1: Analysis Table

Title	**Summary**	**Advantage**	**Technology Used**
Mineral Prospectivity Mapping based on Isolation Forest and Random Forest: Implication for the Existence of Spatial Signature of Mineralization in Outliers.[1]	As measured by ROC curves, predictive future maps obtained by RF modeling based on outlier-based training sets are somewhat inferior to maps obtained by RF modeling based on mineralized location-based training sets, indicating a bias toward known mineralization.	Property testing of outliers not only has a major impact on traditional data-driven MPM but also provides new insights into mineral systems in brownfield or greenfield.	Isolation Forest and Random Forest.
Geological Mineral Energy and Classification Based on Machine Learning.[2]	In this study, a wireless sensor has been successfully applied to 3D mineral energy estimation. Finally, mineral resource energy prediction based on wireless sensors is an important future development trend.	Makes a positive exploration for mineral resource prediction and evaluation in the future.	Random Forest and Support Vector Machine Regression Model.
Mapping Mineral Prospectivity Using a Hybrid Genetic Algorithm–Support Vector Machine (GA–SVM) Model.[3]	The performance of the GA-SVM model was evaluated using both F1 score and spatial efficiency to distinguish potential areas in the study area. The best prospectivity model predicts 95.83% of the known Au deposits.	Displayed a strong spatial correlation between prospective areas and proximity to NE-trending faults.	Hybrid Genetic Algorithm–Support Vector Machine (GA–SVM) Model.
A Comparative Study of Machine Learning Models with Hyperparameter Optimization Algorithm for Mapping Mineral Prospectivity.[4]	The BA and FA have different improvement effects on MLP, AdaBoost, and OCSVM models. The accuracy of the machine learning models is greatly improved after hyperparameter optimization.	The P-A curve represents the prospecting benefit under limited manpower, material resources, and financial resources, and the P/A value can be used as the accuracy evaluation standard for another mineral potential prediction.	Bat algorithm and Firefly algorithm.
Mapping mineral prospectivity through big data analytics and a deep learning algorithm.[5]	The hybrid method of big data analytics and deep learning algorithms provides a novel and powerful tool for the identification of geological anomalies and integration of multi-layer geoscience datasets.	The advantages of big data analytics and deep learning algorithms are that they can deal with complex problems and identify hidden spatial patterns, which are difficult for traditional methods of mapping mineral prospectivity.	Deep autoencoder network
A Bat-Optimized One-Class Support Vector Machine for Mineral Prospectivity Mapping.[6]	The mineral targets predicted by both the common and optimized one-class support vector machine models are spatially consistent with geological and metallogenic characteristics of the study area.	The bat-optimized OCSVM is established to map mineral prospectivity. The model can also be used to solve other similar anomaly detection problems.	One-class support vector machine and bat algorithm

Optimization of Geological and Mineral Exploration by Integrating Remote Sensing Technology and Borehole Database.[7]	Mineral exploration enterprises should actively introduce advanced technologies and equipment to fundamentally improve the efficiency of geological and mineralogical exploration and mineral exploration.	Provide users with a convenient and fast data sharing channel and provide early warning service of geological disasters for coalfield exploration and development.	Remote sensing.
Comparative Analysis of Mineral Identification using CNN and Random Forest.[8]	The objective is to work on the flexibility of the mineral by examining the classification of information from photos taken in different environmental conditions and at different times.	The outcome shows that the proposed model can perceive the mineral accurately.	CNN and Random Forest.
Workshop on Support Vector Machines: Theory and Applications.[9]	There is a unique optimal solution for each choice of SVM parameters. Training many local SVMs instead of one global SVM can significantly improve the performance of a learning machine.	Training many local SVMs instead of one global SVM can significantly improve the performance of a learning machine.	Support Vector Machine algorithm.
Random Forest Classifiers: A Survey and Future Research Directions.[10]	Random forest has enormous potential to become a popular technique as its performance has been found to be comparable to bagging and boosting techniques. Decision trees are commonly used for supervised machine learning.	Random Forest runs efficiently on large databases, can handle thousands of input variables without variable deletion, estimates important variables, generates an internal absolute estimate of normalization error as the forest grows.	Random Forest algorithm.
An Empirical Study of the Naive Bayes Classifier.[11]	In this paper, Naive Bayes approach uses Monte Carlo simulations that allow a systematic study of classification accuracy for many different classes of randomly generated problems.	The loss of information containing features about the class is a better predictor of accuracy when assuming a Naive Bayes model.	Monte Carlo simulations.

IV. Conclusion

The proposed system will help to visualize the datasets present on the Geological BHUKOSH and other portals of India on a map by margining them with respect to the properties of particular location. This system will be using a hybrid model of SVM, Decision Tree and Naive Bayes for the training of models which may lead to better accuracy. The proposed project will try to analyze the problems and leverage the country's accessible geological datasets by delivering an ML based solution for mineral exploration through the construction of a metallogenic model for better mineral exploration based on the geographic factors of a particular area.

References

1. Shuai Zhang, Emmanuel John M. Carranza, Keyan Xiao, Hantao Wei, Fan Yang, Zhenghui Chen, Nan Li and Jie Xiang, "Mineral Prospectivity Mapping based on Isolation Forest and Random Forest: Implication for the Existence of Spatial Signature of Mineralization in Outliers"Natural Resources Research, May 2021,20.

2. Yan Lv and Laijun Lu, "Geological Mineral Energy and Classification Based on Machine Learning"Wireless Communications and Mobile Computing Volume 2021, Article ID 2788161, 7.

3. Xishihui Du, Kefa Zhou , Yao Cui, Jinlin Wang and Shuguang Zhou, "Mapping Mineral Prospectivity Using a Hybrid Genetic Algorithm–Support Vector Machine (GA–SVM) Model"ISPRS Int. J.Geo-Inf., 2021, 10.

4. Lin, N.; Chen, Y.; Liu, H. Liu, H., "A Comparative Study of Machine Learning Models with Hyperparameter Optimization Algorithm for Mapping Mineral Prospectivity" Minerals, 2021, 11.

5. Yihui Xiong, Renguang Zuo, Emmanuel John M. Carranza, "Mapping mineral prospectivity through big data analytics and a deep learning algorithm" Elsevier, 2018, 30.

6. Yongliang Chen , Wei Wu and Qingying Zhao, " A Bat-Optimized One-Class Support Vector Machine for Mineral Prospectivity Mapping" Institute of Mineral Resources Prognosis on Synthetic Information, 2019,23.

7. Yan Niu, Jun Zhao, Zhiyuan Li , Wenjun Xu, Dong Liu, and Meng Zhao, "Optimization of Geological and Mineral Exploration by Integrating Remote Sensing Technology and Borehole Database" Wireless Communications and Mobile Computing Volume 2022, Article ID 9717749, 11.

8. Chandragiri Sandeep, Yellepeddi Srikar, Kodali Rajani, Duvvada Rajeswara Rao, "Comparative Analysis of Mineral Identification using CNN and Random Forest" Department of Computer Science & Engineering, VR Siddhartha Engineering College,2021,10.
9. Theodoros Evgeniou and Massimiliano Pontil,"Workshop on Support Vector Machines: Theory and Applications", Conference Paper, September 2001, 7.
10. Vrushali Y Kulkarni and Dr Pradeep K Sinha,"Random Forest Classifiers :A Survey and Future Research Directions", International Journal of Advanced Computing, 2013, 10.
11. I. Rish,"An Empirical Study of the Naive Bayes Classifier", Article, 2001, 7.
12. https://www.researchgate.net/publication/327562196 , Last accessed : 01/10/22.
13. https://theintelligentminer.com/2020/02/05/machine-learning-for-mineral-exploration/ , Last accessed : 25/09/22.