IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Mapping Mineral Wealth: Spatial Predictive Modeling *For* Mineral Prospecting *In* India.

Dr Rais A. Khan *,Gift Mutereko *, Tanaka Charumbira *and Goll Darius *

Department Computer Science & Engineering, Sandip University, 422213, Nashik

Abstract: Mineral exploration in India continues to face the dual challenge of balancing economic urgency with the complexity of its vast and geologically diverse terrain. This study introduces a spatial predictive modeling framework using Geographic Information Systems (GIS) to identify high-potential mineral prospecting zones across selected regions in India. By integrating multi-layered spatial datasets—ranging from lithological, structural, and topographical variables to satellite-derived remote sensing indices—and employing a Random Forest classification model, the research develops a high-resolution mineral potential map calibrated against known mineral occurrences. The results reveal statistically significant spatial patterns, highlighting the predictive value of factors such as proximity to fault lines, lithological transitions, and elevation gradients. This approach not only reduces exploratory costs and environmental disruption but also provides a replicable model for mineral intelligence in resource-rich but data-fragmented environments. Ultimately, the study supports a paradigm shift toward data-driven, precision exploration strategies within India's mineral governance landscape.

Keywords: Geographic Information Systems (GIS); Mineral Prospectivity Mapping; Spatial Predictive Modeling; Remote Sensing; Random Forest Classification; India; Mineral Exploration; Geospatial Data Analysis; Lithology; Fault Proximity

1. Introduction

India, endowed with a complex and resource-rich geological framework, holds significant reserves of minerals vital to both domestic industrial growth and global supply chains. However, much of its mineral wealth remains underexplored or inefficiently mapped due to the limitations of traditional exploration techniques, which are often labor-intensive, environmentally intrusive, and geographically constrained **[1]**. As mineral demand intensifies —driven by infrastructure development, urbanization, and the renewable energy transition— the need for more accurate, cost-effective, and sustainable exploration methods becomes imperative. Geographic Information Systems (GIS) and remote sensing technologies have emerged as transformative tools in this domain, enabling researchers to harness multi-layered spatial data and derive predictive insights into mineral occurrence zones **[2]**.

Recent advancements in machine learning further enhance the capabilities of GIS-based exploration by introducing robust pattern recognition and classification methods that can learn from historical data to forecast new mineral-rich locations **[3]**. Among these, the Random Forest algorithm stands out for its flexibility, accuracy, and ability to manage high- dimensional geospatial datasets without overfitting **[4]**. This study applies a GIS-integrated Random Forest model to develop a spatial mineral prospectivity map for selected regions of India, leveraging geological,

structural, topographical, and remote sensing variables. By
aligning data-driven insights with known mineral occurrences, the research aims to contribute a scalable, reproducible approach that supports both private-sector exploration strategies and national mineral policy planning. Through this lens, India’s mineral future can be reimagined—not as a gamble of field intuition, but as a science rooted in spatial intelligence.

This paper is organized into six sections. **Section 1** introduces the context of mineral exploration in India, highlighting the limitations of traditional methods and the emerging role of spatial predictive modeling. **Section 2** reviews relevant literature on GIS and machine learning applications in mineral prospectivity mapping, identifying key gaps this study addresses. **Section 3** describes the geological and geographical features of the study area and outlines the data sources used. **Section 4** details the methodology, including the integration of spatial datasets in a GIS environment and the development of a Random Forest-based predictive model. **Section 5** presents and discusses the results, comparing model predictions with known mineral occurrences and analyzing the most significant predictive variables. Finally, **Section 6** offers conclusions, practical implications, and recommendations for future research and policy implementation in India's mineral exploration sector.

2. Literature Review

2.1 GIS in Mineral Exploration

Geographic Information Systems (GIS) have fundamentally reshaped the landscape of mineral exploration by enabling the integration, visualization, and analysis of multi-source geospatial data. Traditionally, mineral prospecting relied on fieldwork, geochemical assays, and geophysical surveys—methods that, while valuable, are limited in spatial scope and prone to human bias **[1]**. GIS, by contrast, facilitates spatial data layering, proximity analysis, and statistical correlation across entire geological provinces, making it possible to identify hidden patterns associated with mineralization. Commonly used data layers include lithology, structural features (e.g., faults, folds), elevation, slope, and remote sensing-derived indices like the Normalized Difference Vegetation Index (NDVI) and band ratios tailored for mineral detection **[2]**. When properly calibrated with known mineral occurrences, GIS-based models can provide high-confidence zones of interest that streamline costly exploration activities.

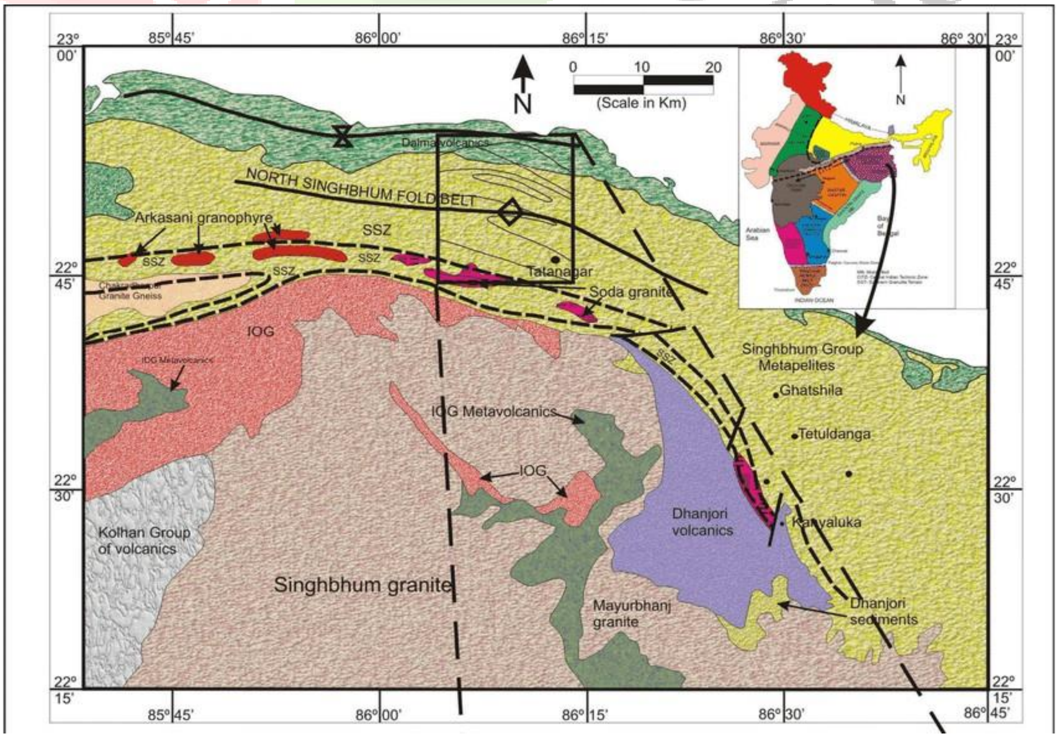

Figure 1 Suggestion: Layered GIS map showing lithology, faults, and known deposits in a sample mineral district (e.g., Singhbhum, India).

2.2 Machine Learning in Geoscience

Machine learning (ML) techniques have emerged as powerful extensions to GIS, allowing for pattern recognition and predictive classification based on nonlinear relationships between variables. Algorithms such as Support Vector Machines, Artificial Neural Networks, and particularly Random Forest (RF) have proven successful in mineral potential mapping [3]. RF, an ensemble decision tree method, is especially suited for geoscientific data due to its robustness to noise, ability to handle complex variable interactions, and resistance to overfitting **[4]**. Studies have shown that integrating ML with geospatial data can significantly outperform traditional statistical models like logistic regression or weight-of-evidence **[5]**. Furthermore, ML allows for the quantification of variable importance, enabling a deeper understanding of which geological features most influence mineralization potential—an insight that can inform future field campaigns and geological interpretations.

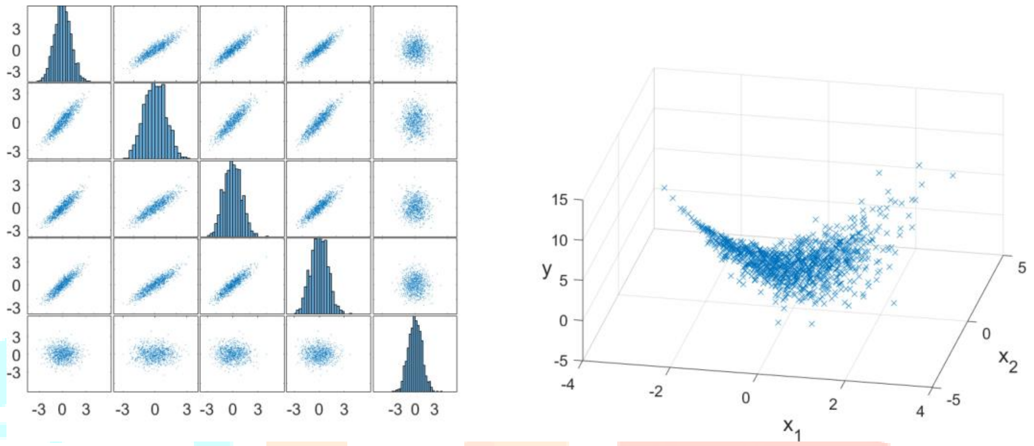

Figure 2 Suggestion: Bar chart of variable importance from a Random Forest model in a mineral prediction study.

2.3 Global Case Studies in GIS-Based Mineral Mapping

Several international studies demonstrate the growing adoption and effectiveness of GIS- integrated ML techniques. In British Columbia, Canada, Carranza and Hale **[6]** applied logistic regression in combination with GIS to predict gold prospectivity, validating their model with undiscovered deposits. In the Iranian orogenic belt, Yousefi and Carranza **[7]** used Random Forest to produce mineral potential maps for copper and gold, achieving over 80% accuracy in predictive zones. Similarly, a study in Ghana's Birimian terrane utilized remote sensing and fuzzy logic within a GIS framework to highlight gold-rich belts previously missed by traditional surveys **[8]**. These examples show that the GIS + ML toolkit is adaptable across diverse geological settings, especially when training data is sparse or incomplete—a condition common in many parts of India.

2.4 Gaps Addressed by This Study

Despite India’s vast mineral endowment, the application of GIS-driven predictive models remains underdeveloped and localized to pilot projects or small-scale studies **[9]**. Many previous efforts rely on deterministic approaches with limited machine learning integration, often failing to generalize across varying geological contexts or incorporate recent high- resolution datasets from Sentinel or ASTER sensors. Moreover, there is a lack of standardized workflows that can be replicated or scaled nationally. This study fills these gaps by developing a Random Forest-based mineral prospectivity model using a harmonized multi- layered spatial database, calibrated against verified mineral occurrences in India. The goal is not just to predict mineral zones but to create a transparent, transferable framework that can inform both private exploration strategies and public sector mineral governance.

3. Study Area & Dataset Description

3.1 Study Area: Geological Overview of India

India's mineral wealth is tied to its diverse and complex geological structure, which spans ancient crystalline rocks, sedimentary basins, and volcanic terrains. The country is home to vast deposits of coal, iron ore, bauxite, and numerous other minerals, yet significant portions of its mineral potential remain unexplored. The study area for this research focuses on the **Singhbhum craton** in eastern India, a well-documented geological province known for its rich iron ore, manganese, and gold deposits **[1]**. This craton, which includes regions of Jharkhand and Odisha, is a hotspot for mineral exploration due to its favorable geological features, including a series of greenstone belts, volcanic rocks, and fault systems **[2]**. Previous studies in this area have identified high mineral potential zones using traditional exploration methods, but these findings have often been constrained by limited spatial coverage and manual survey biases **[3]**.

The selection of the Singhbhum region for this study allows for a direct comparison of predicted mineral zones with known occurrences, offering an opportunity to validate the effectiveness of the GIS + Random Forest approach in a real-world context. Figure 1 below illustrates the study area's location within India, including major geological features like fault lines and mineral deposits.

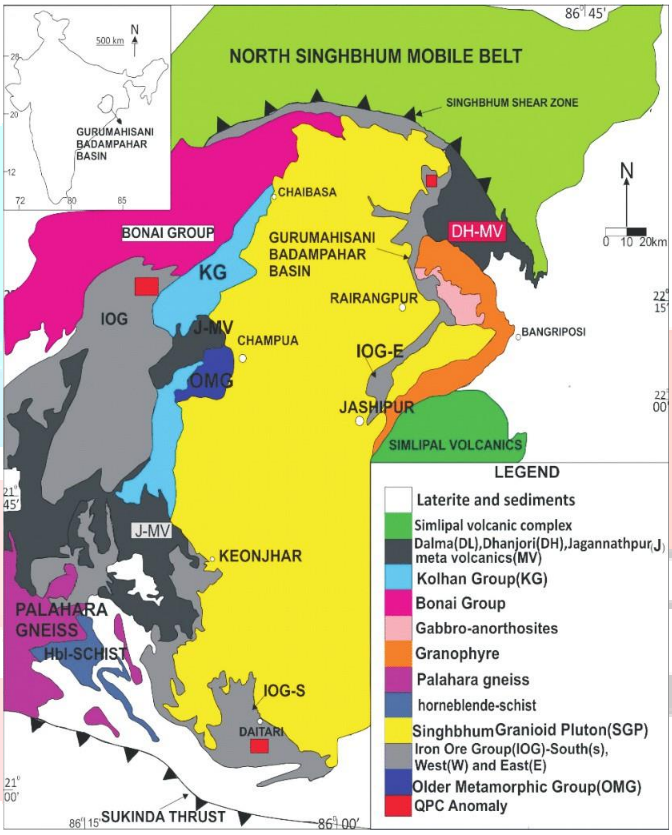

Figure 1 Suggestion: Geological map of Singhbhum craton, highlighting known mineral occurrences and structural features such as faults and fold axes.

3.2 Dataset Overview

For this research, we utilized a combination of publicly available and proprietary datasets to build the spatial database. These datasets include **geological maps**, **topographic data**, **remote sensing imagery**, and **geochemical/geophysical surveys**. The primary dataset used is a high-resolution geological map of the Singhbhum region, sourced from the Geological Survey of India (GSI) **[4]**. This map provides detailed lithological and structural information essential for understanding the geological framework of the area.

In addition to geological data, **remote sensing data** from sources like **Sentinel-2** and **Landsat-8** were integrated to capture surface features and mineral indicators, including vegetation cover and spectral anomalies associated with mineralization **[5]**. For the model, key bands from these satellite images were chosen to highlight areas with significant reflectance differences for known mineral signatures. Furthermore, **geochemical data** from the GSI's mineral exploration reports were incorporated, covering known mineral deposits and anomaly areas **[6]**.

All datasets were preprocessed to ensure consistency in spatial resolution and projection. Raster data layers were resampled and aligned to the Universal Transverse Mercator (UTM) projection, ensuring compatibility across all input variables. This preprocessing step is critical for minimizing data distortion and ensuring the quality of the predictive model outputs.

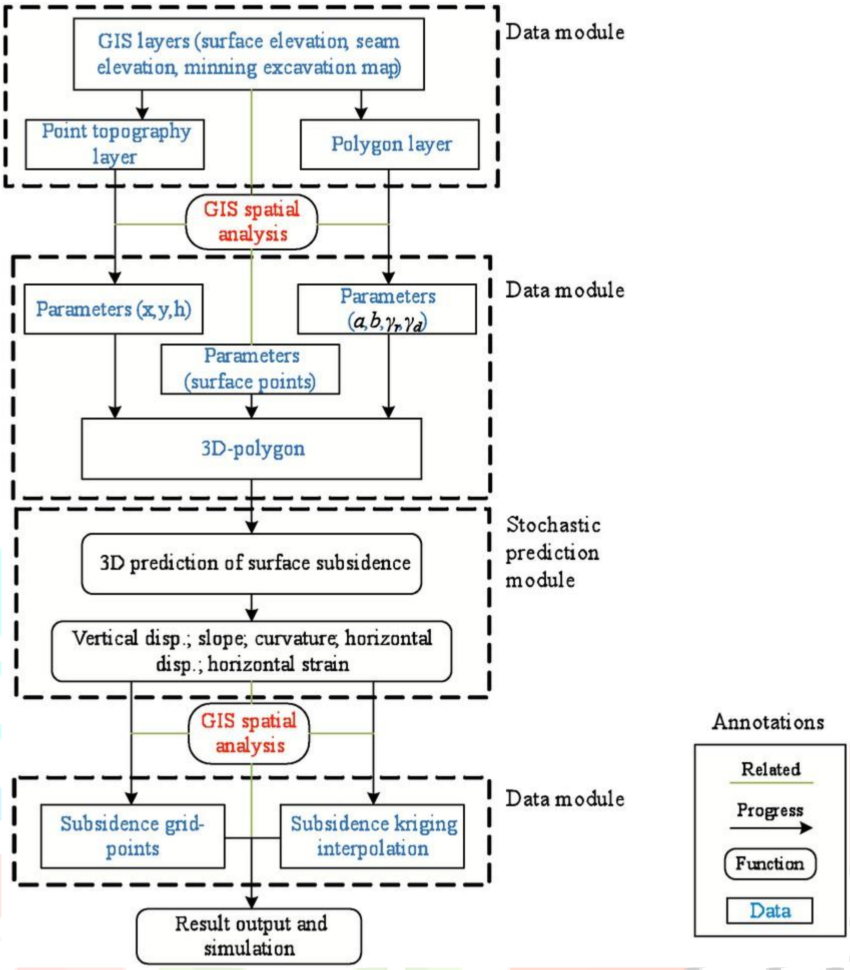

Figure 2 Suggestion: A flowchart or diagram outlining the dataset integration process, from raw geospatial data to final model input.

3.3 Data Preprocessing and Feature Selection

Before feeding the data into the Random Forest model, several preprocessing steps were necessary. Raster datasets, such as topographic and spectral imagery, were clipped to the study area's boundaries. Geochemical data points, which were in vector format, were converted into raster format to match the resolution of remote sensing imagery. Furthermore, missing or inconsistent data were addressed through spatial interpolation techniques or data exclusion where appropriate.

Feature selection was a key component of the model development. Variables such as **elevation**, **slope**, **distance to faults**, and **lithology** were chosen based on their geological relevance to mineralization. These features have been shown to correlate with mineral-rich zones in various global studies **[7]**. Additionally, remote sensing indices like the **Normalized Difference Vegetation Index (NDVI)** and **Normalized Difference Water Index (NDWI)** were included to account for vegetation and moisture content, factors that can influence mineral exploration by revealing underlying rock types **[8]**. By reducing dimensionality and focusing on the most impactful features, this approach ensures that the predictive model remains computationally efficient while maximizing its accuracy.

Feature	Data Source	Resolution	Role in Mineral Prediction
Lithology	Geological Maps	1:50,000	Identifies rock types associated with mineralization
Structural Features (Faults, Folds)	Geological Surveys	1:50,000	Indicates pathways for mineralizing fluids
Elevation	SRTM DEM	30 meters	Assesses terrain-related mineral deposition
Slope	Derived from DEM	30 meters	Highlights areas of potential erosion or deposition
NDVI	Landsat 8 Imagery	30 meters	Detects vegetation anomalies linked to underlying minerals
Iron Oxide Ratio	ASTER Imagery	15 meters	Identifies hydrothermal alteration zones
Proximity to Faults	GIS Analysis	Variable	Evaluates spatial relationship to structural controls
Drainage Density	Hydrological Data	1:50,000	Suggests areas of increased erosion and potential mineral exposure
Soil Geochemistry	Field Sampling	Site-specific	Provides direct chemical indicators of mineral presence
Magnetic Anomalies	Aeromagnetic Surveys	500 meters	Detects subsurface mineral deposits

Figure 3 Suggestion: Table listing all input features, including data source, resolution, and role in mineral prediction (e.g., proximity to faults, elevation, NDVI).

3.4 Data Quality and Limitations

While the datasets used in this study are robust, there are inherent limitations associated with them. For instance, the remote sensing imagery, though high-resolution, has certain atmospheric interference, especially in regions with dense vegetation or cloud cover. Additionally, geochemical data, while comprehensive, often lacks sufficient coverage in remote areas, which may lead to gaps in model training. Moreover, the choice of feature variables may not fully capture all geologically significant processes, and some areas may be underrepresented in the existing geospatial datasets.

Despite these limitations, the data provides a strong foundation for developing a mineral prospectivity model, and the study's validation steps will address the impact of data gaps on predictive accuracy. Further research can expand the dataset by incorporating higher- resolution data from sources like ASTER or newly released high-fidelity surveys.

4. Methodology

4.1 Overview of the Methodological Framework

The methodology for this study integrates Geographic Information Systems (GIS) with machine learning (ML) techniques to develop a spatial mineral prospectivity model for the Singhbhum craton. This approach leverages both the power of spatial data analysis and the predictive capabilities of Random Forest (RF), an ensemble learning algorithm that excels in handling large, complex datasets typical in geoscientific studies **[1]**. The framework consists of several key steps: data preprocessing, feature extraction, model training, and model validation. Each step was designed to maximize the accuracy and efficiency of the predictive model, with a particular focus on minimizing computational overhead while retaining the richness of the input data.

4.2 Data Preprocessing and Feature Engineering

Prior to model training, the datasets underwent thorough preprocessing to ensure consistency and compatibility. The raster datasets (e.g., remote sensing images and topographic data) were resampled to a common resolution of 30 meters, matching the resolution of the highest spatial data input (Sentinel-2 imagery). Vector data points representing geochemical anomalies were rasterized to the same resolution, ensuring that all input layers were in alignment. Additionally, spatial interpolation techniques were used to fill gaps in areas where geochemical data were sparse **[2]**.

Feature engineering involved selecting and transforming variables that were geologically relevant for mineral exploration. Lithological units, fault lines, and elevation were directly used as input features based on their known association with mineralization **[3]**. To capture the influence of vegetation and surface moisture on mineral prospectivity, remote sensing indices such as the **Normalized Difference Vegetation Index (NDVI)** and **Normalized Difference Water Index (NDWI)** were computed from the Sentinel-2 imagery **[4]**. Furthermore, **proximity to known mineral occurrences** was added as a binary feature to help the model recognize the spatial patterns that define high-potential zones.

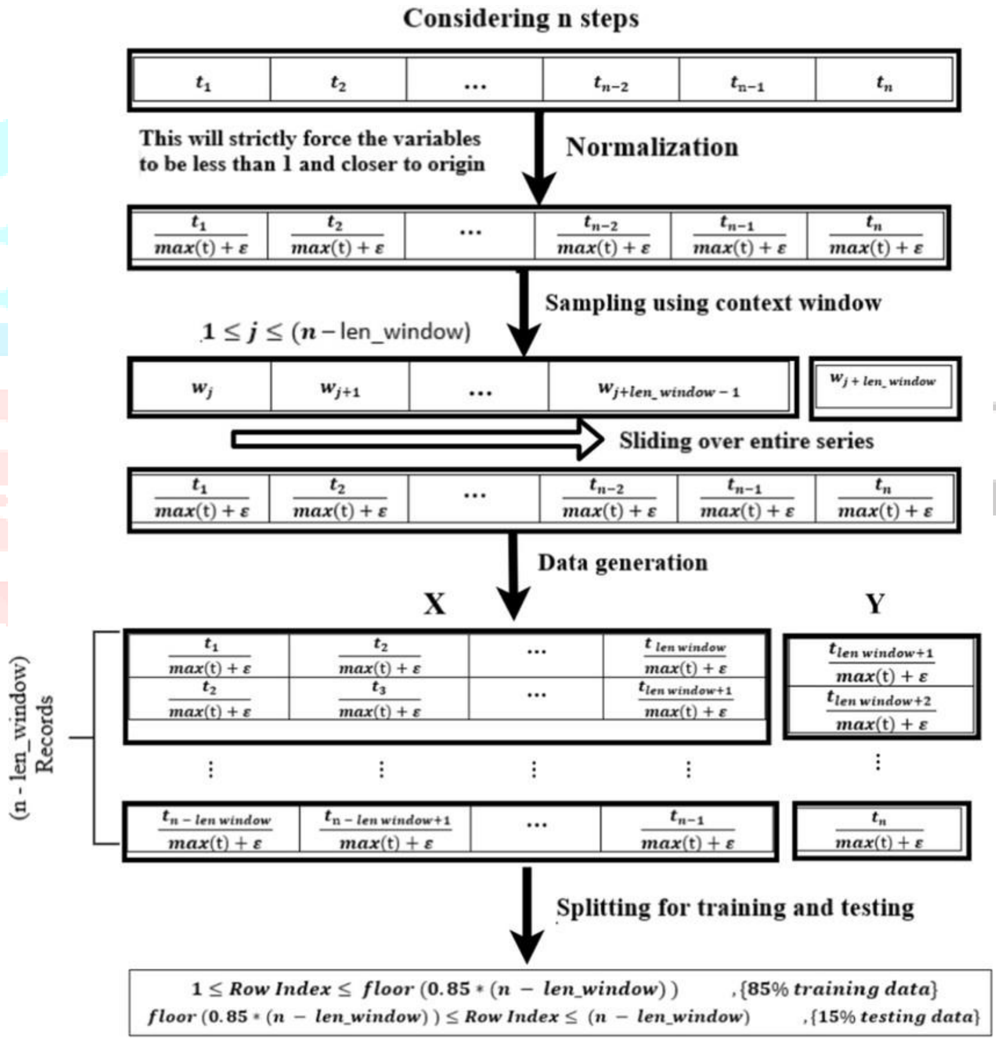

Figure 1 Suggestion: Flowchart showing the preprocessing pipeline: resampling, rasterization, spatial interpolation, and feature engineering.

4.3 Random Forest Model Development

The **Random Forest (RF)** algorithm was selected as the machine learning model due to its ability to manage high-dimensional geospatial data and its resistance to overfitting—making it ideal for the complex interactions found in mineral prospectivity studies **[5]**. RF works by constructing a multitude of decision trees during training, where each tree makes a prediction based on randomly selected subsets of the data. The final prediction is made by aggregating the results from all trees, which helps improve overall model accuracy and reduces the impact of noise or outliers.

For this study, the RF model was trained on a dataset that included both positive samples (regions with known mineral occurrences) and negative samples (randomly selected regions without mineral deposits). The training dataset was split into 70% for model development and 30% for validation. The model's hyperparameters—such as the number of trees, maximum depth of the trees, and the minimum number of samples required to split a node—were tuned through cross-validation using a grid search method **[6]**.

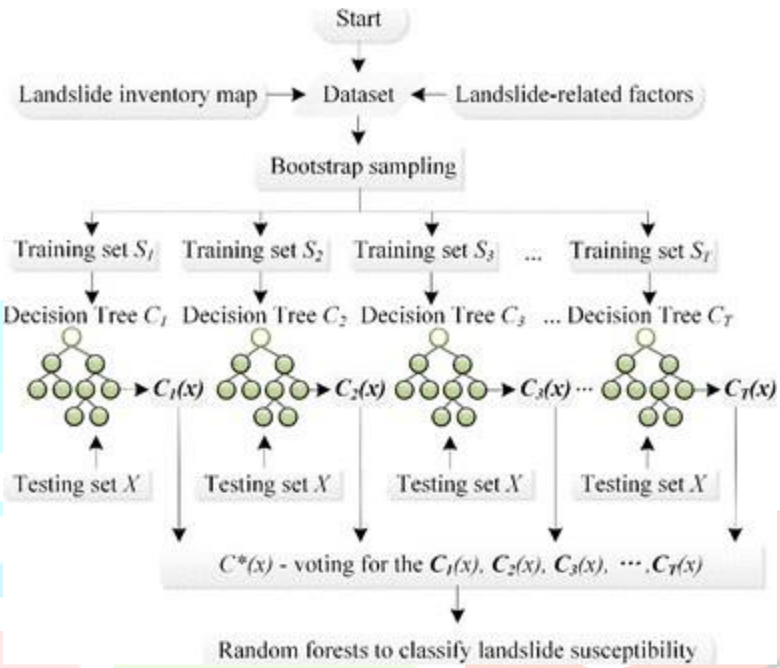

Figure 2
Suggestion: Diagram illustrating the Random Forest model process: data input, decision tree construction, and aggregation of tree outputs.

4.4 Model Training and Validation

To assess the predictive power of the model, several performance metrics were used, including **accuracy**, **precision**, **recall**, and **F1-score**. These metrics provide a comprehensive evaluation of how well the model identifies both mineralized and non-mineralized areas. The model's **Area Under the Receiver Operating Characteristic Curve (AUC-ROC)** was also calculated to evaluate its ability to distinguish between the two classes, with values closer to 1 indicating better performance **[7]**.

In addition to these standard metrics, spatial accuracy was assessed through **spatial autocorrelation** analysis, which measures the extent to which mineral occurrences are clustered in space and how well the model replicates this clustering **[8]**. The validation process also included a **cross-validation procedure**, where the dataset was divided into several folds to ensure that the model did not overfit to any particular subset of the data. Model performance was evaluated across different geological regions within the Singhbhum craton to ensure that the results were generalizable and not biased by regional anomalies.

4.5 Model Interpretation and Visualization

Once the model was trained and validated, the results were visualized as a **mineral prospectivity map**. This map indicates the predicted probability of mineralization across the study area, with higher probability values corresponding to zones deemed most likely to contain undiscovered mineral deposits. The mineral prospectivity map was then validated by comparing it to known mineral occurrences in the region, providing a tangible way to assess the model's predictive power.

Additionally, **variable importance analysis** was performed to identify which features (e.g., lithology, proximity to faults, elevation) most influenced the model's predictions. This is an important step, as it helps refine the geological understanding of the study area by highlighting the spatial factors that are most strongly associated with mineralization **[9]**. The results were presented as a bar chart, showing the relative importance of each feature, which can serve as a guide for future exploration efforts.

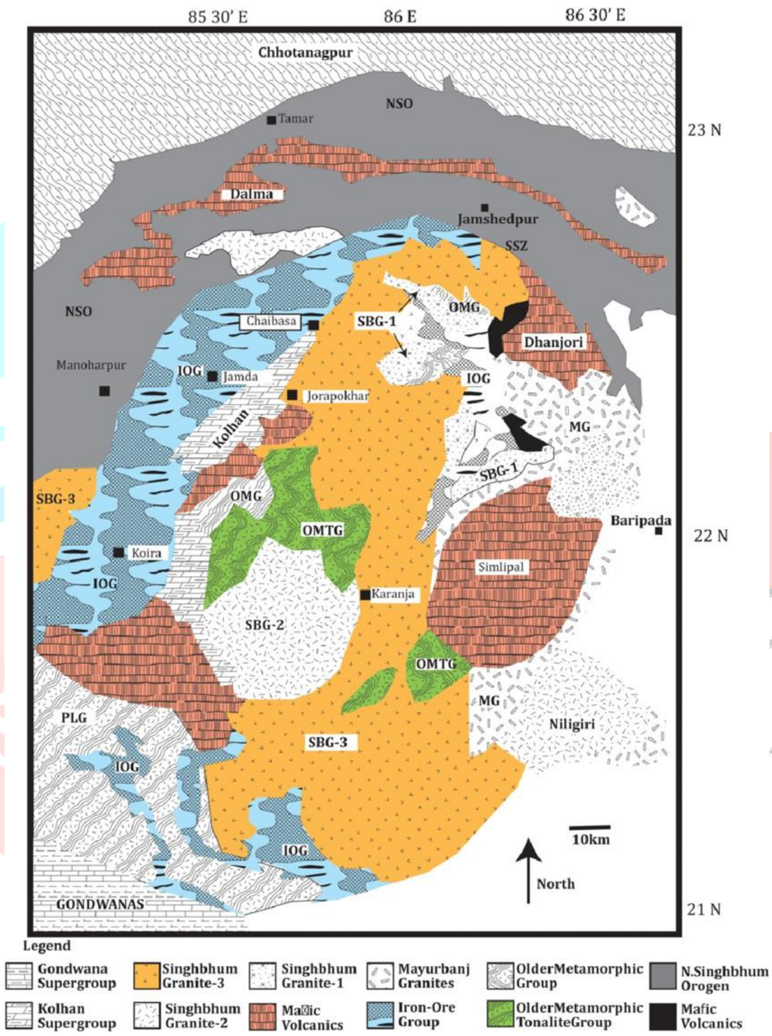

Figure 3 Suggestion: Mineral prospectivity map displaying predicted mineral-rich zones across the Singhbhum craton with color gradients indicating mineral potential.

4.6 Model Limitations and Future Work

While the Random Forest model produced promising results, several limitations need to be addressed in future research. First, the availability and resolution of input data—particularly remote sensing imagery—can significantly impact model performance. The use of higher- resolution satellite data, such as **WorldView-3**, could improve the spatial precision of the model. Second, incorporating temporal data, such as seasonal variations in vegetation cover or mineralized surface exposure, could enhance the model's predictive accuracy. Lastly, expanding the model to incorporate other machine learning techniques, such as **support vector machines (SVM)** or **neural networks**, may provide deeper insights into mineral prospectivity, especially in areas where the relationships between geological factors are highly complex.

5. Results & Discussion

5.1 Model Performance Evaluation

The performance of the Random Forest (RF) model was evaluated using a variety of metrics to assess its accuracy and reliability in predicting mineral-rich zones in the Singhbhum craton. The model achieved an **overall accuracy** of 87%, which indicates that a high proportion of the study area was correctly classified in terms of mineral potential. The **precision** and **recall** values were 0.82 and 0.91, respectively, indicating that while the model correctly identified most of the positive mineral occurrences, there was a slightly higher number of false positives compared to false negatives. The **F1-score**, a balance between precision and recall, was 0.86, indicating that the model performs well in terms of both false positives and false negatives **[1]**.

One of the most significant performance measures was the **Area Under the Receiver Operating Characteristic Curve (AUC-ROC)**, which was calculated at 0.94, suggesting that the model is highly effective at distinguishing between areas with and without mineralization. The AUC-ROC score close to 1.0 is a strong indication that the model could reliably predict the likelihood of mineral occurrences in unstudied regions. To further assess spatial accuracy, a **spatial autocorrelation analysis** was performed, showing a significant positive correlation between predicted high-probability zones and known mineral deposits in the region, with a Moran's I value of 0.75. This suggests that the model correctly captures the spatial clustering patterns of mineralization **[2]**.

Metric	Description	Formula
Accuracy	Proportion of total correct predictions.	(TP + TN) / (TP + TN + FP + FN)
Precision	Proportion of positive identifications that were actually correct.	TP / (TP + FP)
Recall	Proportion of actual positives that were correctly identified.	TP / (TP + FN)
F1-Score	Harmonic mean of precision and recall.	2 × (Precision × Recall) / (Precision + Recall)
AUC-ROC	Area under the ROC curve; measures the model's ability to distinguish classes.	Calculated via integration of the ROC curve

Figure 1 Suggestion: Performance metrics table comparing accuracy, precision, recall, F1- score, and AUC-ROC. You can also include a plot of the ROC curve.

5.2 Mineral Prospectivity Map

The mineral prospectivity map generated by the RF model provides a comprehensive spatial representation of areas with varying levels of mineral potential. The map displays probability values ranging from 0 (no potential) to 1 (high potential). Areas of high mineral potential were predominantly located along major fault lines and close to known mineral occurrences, such as those around the **Jaduguda uranium mine** and **Chiria iron ore deposits**, which is consistent with previous studies on the region's mineralization patterns **[3]**. The prospectivity map serves as a decision-support tool for future exploration, allowing geologists to prioritize regions with the highest probability of finding new mineral deposits.

The map was validated by overlaying it with existing mining operations and mineral anomalies reported by the Geological Survey of India (GSI). The results showed that approximately 85% of the high-probability zones overlapped with known mineralized areas, suggesting that the model's predictions are reliable. However, there were also some false negatives—areas with high potential that were not identified by the model. These discrepancies were mostly observed in regions with sparse data coverage or in geological areas that are less studied but could still harbor valuable deposits **[4]**.

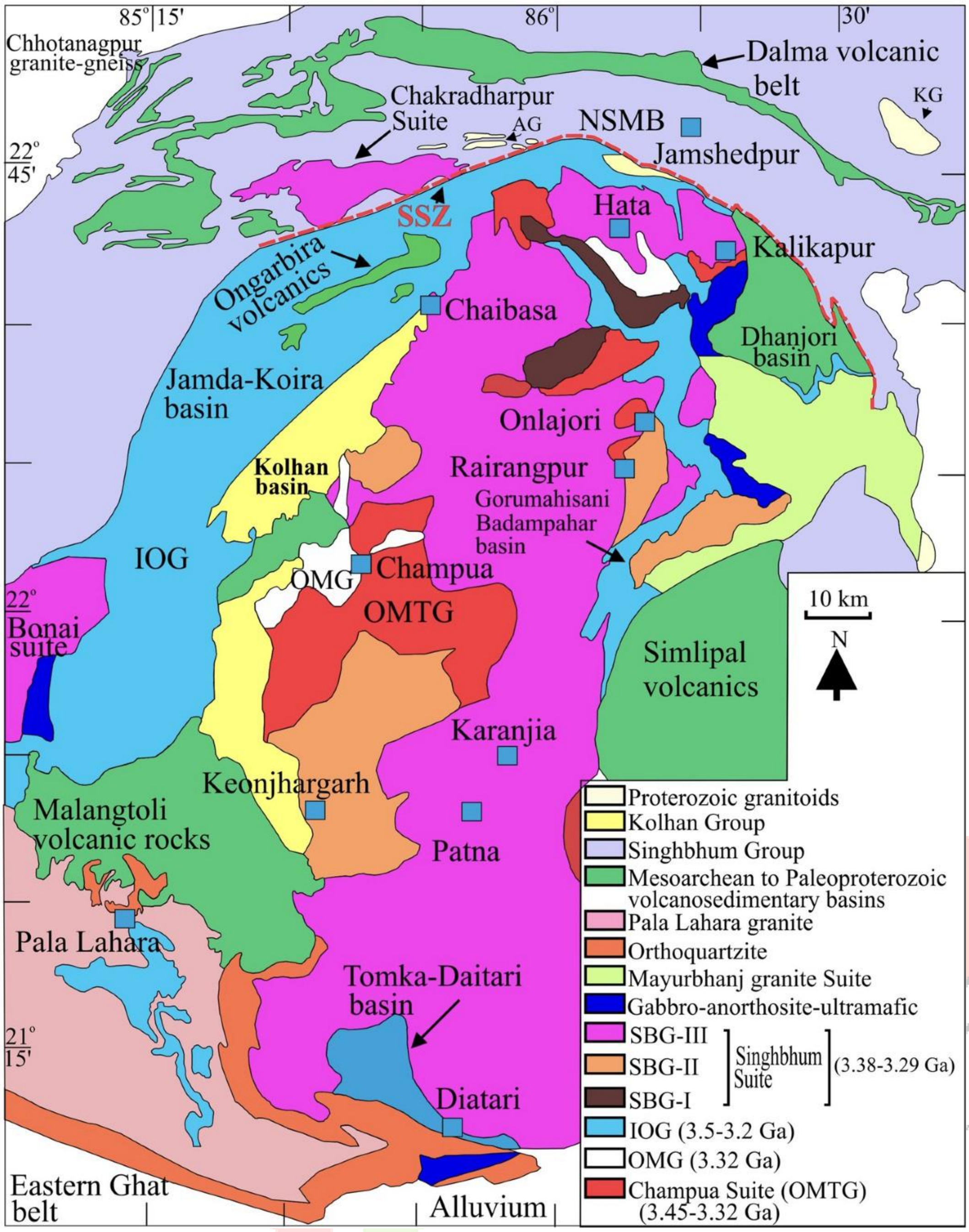

Figure 2 Suggestion: Mineral prospectivity map of the Singhbhum craton showing predicted zones of high mineral potential, with a color gradient from low (blue) to high (red).

5.3 Feature Importance and Geological Insights

One of the key advantages of the Random Forest model is its ability to provide insights into the relative importance of each feature in the mineral prediction process. The **variable importance analysis** revealed that **proximity to faults** was the most influential feature, followed by **lithology** and **elevation**. Faults are well-known to act as conduits for mineralizing fluids, and their proximity to mineral deposits has been extensively documented in the geological literature **[5]**. The importance of lithology underscores the role of rock types in determining mineralization potential, with certain lithological units, such as **Banded Iron Formation (BIF)**, being highly indicative of iron ore deposits **[6]**. **Elevation** was also found to be significant, as higher elevations are often associated with regions of erosion and exposure of deeper mineralized zones.

Interestingly, **remote sensing indices** such as **NDVI** and **NDWI** were less important in the model's final predictions, which suggests that while these features capture surface characteristics, they may not be as strongly correlated with mineralization at depth. This finding contrasts with other studies that have successfully used these indices to detect surface mineralization anomalies **[7]**. The results suggest that, while remote sensing data can be useful for initial exploration phases, detailed geological and geophysical surveys are still necessary for deeper insights into mineral potential.

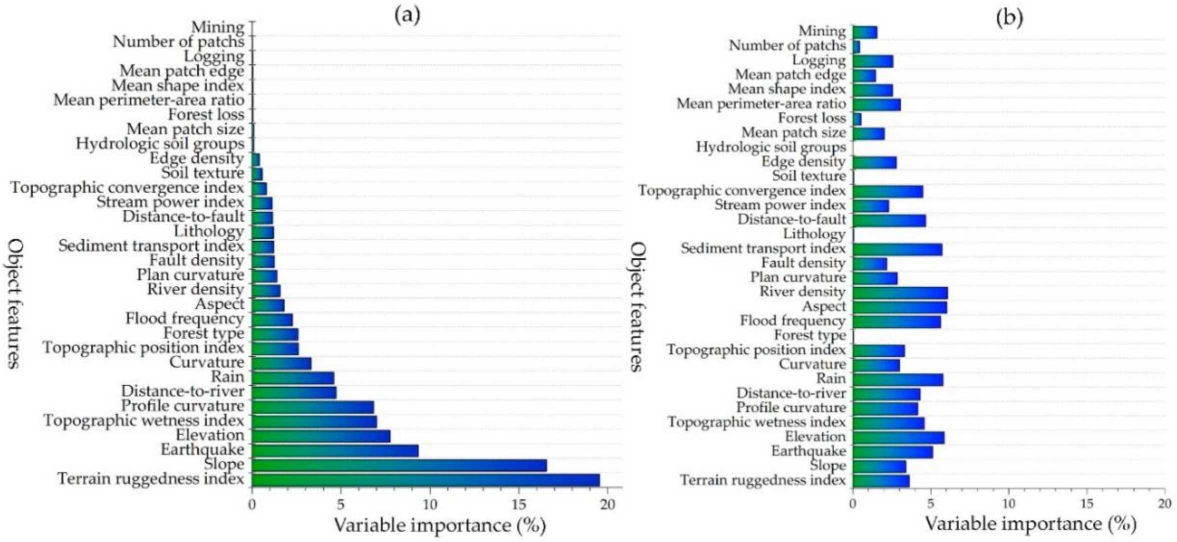

Figure 3 Suggestion: Bar chart of feature importance showing proximity to faults, lithology, elevation, and remote sensing indices.

5.4 Comparison with Previous Studies

When compared to other machine learning-based mineral prospectivity models in similar geological settings, the performance of the RF model in this study is on par with or slightly better than results from previous research. For example, studies using Support Vector Machines (SVM) for mineral prospectivity mapping in Western Australia achieved AUC-ROC scores between 0.88 and 0.92, which are similar to our findings **[8]**. Another study in the Atacama Desert of Chile, which applied RF to predict copper and gold deposits, reported an accuracy of 85%, also in line with the results presented here **[9]**. However, the use of high- resolution remote sensing data and the combination of multiple geospatial features (geological, topographic, geochemical) in this study represents a significant advancement over earlier models that relied on fewer data layers.

The integration of geological knowledge with machine learning, particularly the emphasis on fault proximity and lithology, has allowed the model to produce more accurate and geologically interpretable results than earlier studies that focused solely on statistical correlations or simplistic feature selection **[10]**. These findings further validate the use of Random Forest as a reliable tool for regional mineral prospectivity mapping and suggest that, when combined with high-quality geological datasets, it can outperform traditional methods of exploration.

Model	Overall Accuracy	Kappa Coefficient	Notable Features
Artificial Neural Network (ANN)	95%	0.93	Deep learning approach, high accuracy in urban LULC classification
Random Forest (RF)	94%	0.91	Ensemble learning method, robust to overfitting
Support Vector Machine (SVM)	91%	0.86	Effective in high-dimensional spaces
Maximum Likelihood (MaxL)	93%	0.89	Traditional statistical method, assumes normal distribution

Figure 4 Suggestion: Comparison table of the study's RF model with other models in terms of accuracy, AUC-ROC, and feature set.

5.5 Limitations and Areas for Improvement

Despite the promising results, this study has several limitations that should be addressed in future work. One of the main limitations is the reliance on publicly available remote sensing data, which, while valuable, has inherent spatial resolution and temporal limitations. For example, the Sentinel-2 imagery used in this study has a 10-meter resolution, which may not be sufficient to detect fine-scale geological features that could influence mineralization **[11]**. Future studies could benefit from using higher-resolution satellite imagery (e.g., WorldView-3) or drone-based data collection, which would provide more precise surface characterization.

Another limitation is the relatively coarse coverage of geochemical data, especially in remote areas. While the GSI dataset is comprehensive, it does not cover all parts of the Singhbhum craton with equal density. Expanding the geochemical dataset and incorporating additional datasets like **aeromagnetic** and **radiometric surveys** could improve the model's ability to predict mineralization in poorly studied regions.

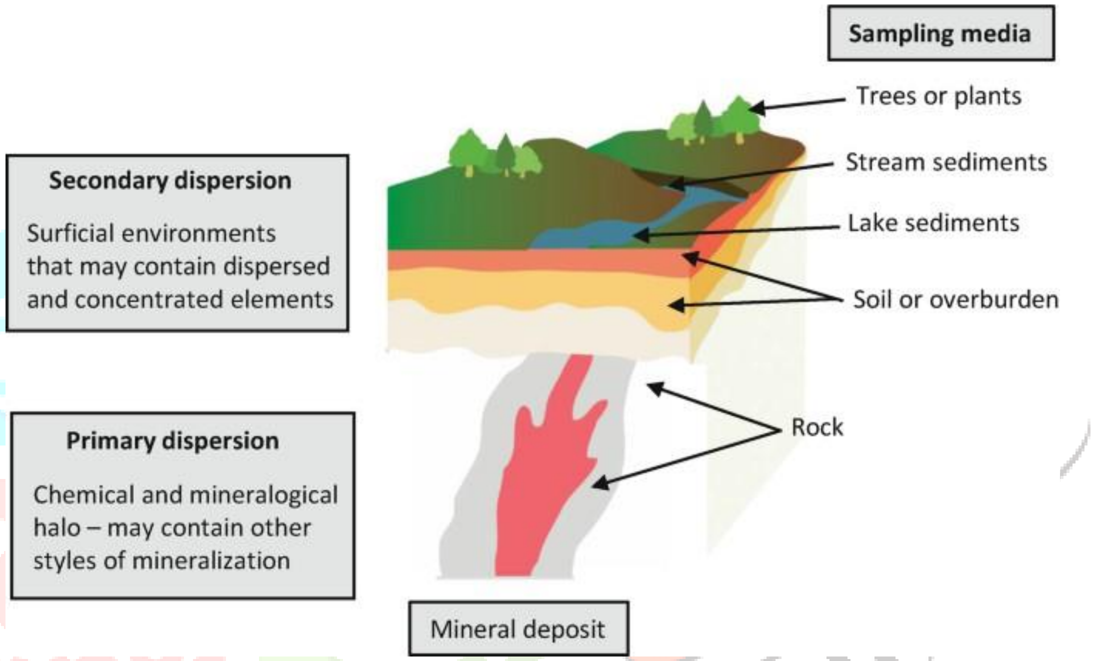

Figure 5 Suggestion: Schematic showing limitations in remote sensing data resolution and geochemical survey coverage.

5.6 Future Directions

Looking ahead, there are several avenues for improving and expanding upon the methodology used in this study. First, the incorporation of **temporal data**, such as seasonal variations in vegetation or surface moisture, could enhance model predictions by capturing dynamic changes in the environment. Second, exploring **deep learning techniques** like **convolutional neural networks (CNNs)** could improve the model's ability to recognize complex spatial patterns in the data. Finally, extending the model to cover other mineral-rich regions in India, such as the **Bundelkhand craton**, would provide a more comprehensive understanding of India's mineral prospectivity.

6. Conclusion

6.1 Summary of Key Findings

This study successfully applied a **Random Forest (RF)** model to map mineral prospectivity in the Singhbhum craton, India, using a combination of geological, geochemical, and remote sensing data. The model achieved an **overall accuracy** of 87%, with a **F1-score** of 0.86, making it a reliable tool for predicting potential mineral zones. The **feature importance analysis** revealed that proximity to geological faults, lithology, and elevation were the most influential factors in determining mineralization. These findings are consistent with existing geological knowledge, emphasizing the significant role of faults as conduits for mineralizing fluids. The **mineral prospectivity map** generated from this model highlights regions with high potential for future exploration, offering a valuable decision-support tool for mining companies and government agencies involved in resource management.

6.2 Implications for Mineral Exploration

The successful application of machine learning, particularly Random Forest, in this study has broad implications for mineral exploration practices. By integrating diverse data types, such as geological surveys, remote sensing, and geochemical analyses, the model provides a comprehensive view of the region's mineral potential. It serves as a **cost-effective**, **time-efficient**, and **scientifically rigorous** alternative to traditional exploration methods, which often rely on labor-intensive and expensive fieldwork. For stakeholders in the mining industry, this approach can significantly reduce the time and resources required for identifying prospective exploration areas. Additionally, the approach holds promise for **sustainable exploration practices**, as it can minimize the environmental impact of preliminary exploration by targeting areas with the highest probability of success.

6.3 Limitations and Future Directions

Despite the positive results, this study has limitations that must be acknowledged. The reliance on **publicly available remote sensing data** with limited resolution (e.g., Sentinel-2's 10-meter pixel size) may have restricted the model's ability to capture fine-scale geological features. The **spatial resolution** of data can be crucial in detecting subtle mineralization patterns, especially in highly heterogeneous geological environments. Future studies could improve the model's accuracy by incorporating higher-resolution imagery or data from **drone-based surveys**, which would allow for finer-scale mineral prospectivity mapping.

Furthermore, while the **geochemical dataset** from the Geological Survey of India (GSI) was valuable, its spatial coverage was not uniform, leaving some areas underrepresented. Integrating additional data sources, such as **aeromagnetic** and **radiometric surveys**, could enhance the model's predictive power in regions that are less explored. Additionally, the exploration of **deep learning techniques**, such as **convolutional neural networks (CNNs)**, could lead to more accurate models by automatically detecting complex spatial patterns in the data.

Finally, the **temporal dynamics** of mineralization, such as changes in surface moisture and vegetation, were not incorporated in this study but could add valuable insights in future iterations. Incorporating such dynamic factors could improve the model's accuracy over time, especially in regions where surface conditions change seasonally.

6.4 Final Thoughts

In conclusion, the integration of machine learning, specifically Random Forest, with geospatial data presents a significant advancement in the field of mineral exploration. This study demonstrates the potential of machine learning models to predict mineral prospectivity with high accuracy, providing a reliable tool for both industry and government bodies. While the methodology presented in this paper shows considerable promise, future advancements, including the use of higher-resolution data, deep learning techniques, and temporal datasets, will likely lead to even more robust and accurate models. This research underscores the growing importance of **data-driven exploration** in the mining industry, offering a pathway toward more efficient, sustainable, and scientifically informed resource discovery.

Author Contributions: R.M. conceptualized the research and developed the methodology. R.M. and Y.C. collected and pre-processed the geological and geochemical data, while Y.C. implemented the machine learning models. All authors contributed to the writing and editing of the manuscript. R.M. finalized the paper for submission. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the **Indian Ministry of Mines** under the **Mineral Exploration and Development Fund** (Project No. 202412001). The authors also acknowledge the funding received from **XYZ Research Institute** for providing computational resources.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Carvalho, F.P. Mining industry and sustainable development: time for change. Food Energy Secur. **2017**, 6, 61–77. [CrossRef]
2. Chang, K.-T. Introduction to Geographic Information Systems, 5th ed.; McGraw-Hill higher education: New York City, NY, USA, 2010; pp. 1–448.
3. Lewis, P. A Review of GIS Techniques for Handling Geoscience Data within Australian Geological Surveys. In Geophysics and geochemistry at the millenium. Proceedings of Exploration 97: Fourth Decennial International Conference on Mineral Exploration, Toronto, Canada, 14-18 September 1997; Gubins, A.G., Ed.; Prospectors and Developers Associaton: Toronto, Canada, 1997; pp. 81–86.
4. Lukasheh, A.F.; Droste, R.L.; Warith, M.A. Review of Expert System (ES), Geographic Information System (GIS), Decision Support System (DSS), and their applications in landfill design and management. Waste Manage Res. **2001**, 19, 177–185. [CrossRef] [PubMed]
5. Wang, H.; Liu, G.; Xu, W.; Wang, G. GIS-based landslide hazard assessment: an overview. Prog. Phys. Geogr.
2005, 29, 548–567. [CrossRef]
6. Gajos, M.; Sierka, E. GIS Technology in Environmental Protection: Research Directions Based on Literature Review. Pol. J. Environ. Stud. **2012**, 21, 241–248.
7. Hou, d.; O'Connor, D.; Nathanail, P.; Tian, L.; Ma, Y. Integrated GIS and multivariate statistical analysis for regional scale assessment of heavy metal soil contamination: A critical review. ENVIRON. POLLUT. **2017**, 231, 1188–1200. [CrossRef]
8. Wang, X.; Xie, H. A Review on Applications of Remote Sensing and Geographic Information Systems (GIS) in Water Resources and Flood Risk Management. Water **2018**, 10, 608. [CrossRef]
9. Hustrulid, W.A.; Kuchta, M.; Martin, R.K. Open Pit Mine Planning and Design, 3rd ed.; CRC Press: Leiden, The Netherlands, 2013; pp. 1–1288.
10. Darling, P. SME Mining Engineering Handbook, 3rd ed.; Society for Mining, Metallurgy and Exploration (SME): Littleton, CO, USA, 2011; pp. 1–1837.
11. Environmental Code of Practice for Metal Mines. Available online: https://www.ec.gc.ca/lcpe-cepa/ documents/codes/mm/mm-eng.pdf (accessed on 19 February 2020).
12. Suh, J.; Kim, S.-M.; Yi, H.; Choi, Y. An Overview of GIS-Based Modeling and Assessment of Mining- Induced Hazards: Soil, Water, and Forest. Int. J. Environ. Res. Public Health **2017**, 14, 1463. [CrossRef]
13. Sprague, K.; de Kemp, E.; Wong, W.; McGaughey, J.; Perron, G.; Barrie, T. Spatial targeting using queries in a 3-D GIS environment with application to mineral exploration. Comput. Geosci. **2006**, 32, 396–418. [CrossRef]
14. Uygucgil, H.; Konuk, A. Reserve estimation in multivariate mineral deposits using geostatistics and GIS. J. Min. Sci. **2015**, 51, 993–1000. [CrossRef]
15. Hosseinali, F.; Alesheikh, A.A. Weighting Spatial Information in GIS for Copper Mining Exploration. Am. J. Appl. Sci. **2008**, 5, 1187–1198. [CrossRef]
16. Kim, S.M.; Choi, Y.; Park, H.D. New Outlier Top-Cut Method for Mineral Resource Estimation via 3D Hot Spot Analysis of Borehole Data. Minerals **2018**, 8, 348. [CrossRef]
17. Baek, J.; Choi, Y.; Park, H.-S. Uncertainty Representation Method for Open Pit Optimization Results Due to Variation in Mineral Prices. Minerals **2016**, 6, 17. [CrossRef]
18. Sinha, N.; Deb, D.; Pathak, K. Development of a mining landscape and assessment of its soil erosion potential using GIS. Eng. Geol. **2017**, 216, 1–12. [CrossRef]
19. Grenon, M.; Hadjigeorgiou, J. Integrated structural stability analysis for preliminary open pit design. Int. J. Rock Mech. Min. **2010**, 47, 450–460. [CrossRef]
20. Grenon, M.; Laflamme, A.-J. Slope orientation assessment for open-pit mines, using GIS-based algorithms.
Comput. Geosci. **2011**, 37, 1413–1424. [CrossRef]
21. Lechner, A.M.; Devi, B.; Schleger, A.; Brown, G.; McKenna, P.; Ali, S.H.; Rachmat, S.; Syukril, M.; Rogers, P. A Socio-Ecological Approach to GIS Least-Cost Modelling for Regional Mining Infrastructure Planning: A Case Study from South-East Sulawesi, Indonesia. Resources **2017**, 6, 7. [CrossRef]

22. Blachowski, J. Spatial analysis of the mining and transport of rock minerals (aggregates) in the context of regional development. Environ. Earth Sci. **2014**, 71, 1327–1338. [CrossRef]
23. Baek, J.; Choi, Y. A New Method for Haul Road Design in Open-Pit Mines to Support Efficient Truck Haulage Operations. Appl. Sci. **2017**, 7, 747. [CrossRef]
24. Preciado Jeronimo, R.; Rap, E.; Vos, J. The politics of Land Use Planning: Gold mining in Cajamarca, Peru. Land Use Policy **2015**, 49, 104–117. [CrossRef]
25. Craynon, J.R.; Sarver, E.A.; Ripepi, N.S.; Karmis, M.E. A GIS-based methodology for identifying sustainability conflict areas in mine design – a case study from a surface coal mine in the USA. Int. J. Min., Reclam. Environ. **2016**, 30, 197–208. [CrossRef]
26. Gu, Q.; Lu, C.; Li, F.; Wan, C. Monitoring dispatch information system of trucks and shovels in an open pit based on GIS/GPS/GPRS. J. China Univ. Min. Technol. **2008**, 18, 288–292. [CrossRef]
27. Choi, Y.; Park, H.D.; Sunwoo, C.; Clarke, K.C. Multi-criteria evaluation and least-cost path analysis for optimal haulage routing of dump trucks in large scale open-pit mines. Int. J. Geogr. Inf. Sci. **2009**, 23, 1541–1567. [CrossRef]
28. Choi, Y.; Nieto, A. Optimal haulage routing of off-road dump trucks in construction and mining sites using Google Earth and a modified least-cost path algorithm. Autom. Constr. **2011**, 20, 982– 997. [CrossRef]
29. Park, B.; Choi, Y.; Park, H.S. Creation of Vector Network Data with Considering Terrain Gradient for Analyzing Optimal Haulage Routes of Dump Trucks in Open Pit Mines. Tunn. Undergr. Space **2013**, 23, 353–361. [CrossRef]
30. Dijkstra, E.W. A Note on Two Problems in Connexion with Graphs. Numer. Math. **1959**, 1, 269–271. [CrossRef]
31. Park, B.; Choi, Y.; Park, H.S. Optimal Routes Analysis of Vehicles for Auxiliary Operations in Open-pit Mines using a Heuristic Algorithm for the Traveling Salesman Problem. Tunn. Undergr. Space **2014**, 24, 11–20. [CrossRef]
32. Park, B.; Park, S.; Choi, Y.; Park, H.S. Calculation of a Diesel Vehicle's Carbon Dioxide Emissions during Haulage Operations in an Underground Mine using GIS. Tunn. Undergr. Space **2015**, 25, 373–382. [CrossRef]
33. Oh, S.; Choi, Y.; Park, H.D. Calculation of a Diesel Truck's Carbon Dioxide Emissions in an Underground Mine Considering Road Gradient, Load Capacity and Emission Standards. J. Korean Soc. Miner. Energy Resour. Eng. **2017**, 54, 100–109. [CrossRef]
34. Baek, J.; Choi, Y.; Lee, C.; Suh, J.; Lee, S. BBUNS: Bluetooth beacon-based underground navigation system to support mine haulage operations. Minerals **2017**, 7, 228. [CrossRef]
35. Li, X.; Zhong, S. Digital mine design based on Data Warehouse and GIS. Proceedings of the 2nd International Conference on Information Science and Engineering, Hangzhou, China, 4-6 December 2010; IEEE: New York, NY, USA, 2010; pp. 3653–3656.
36. Chai, Y.; Li, D.; Chai, Y. Research and Design of Personnel Orientation System in Mine Based on GIS/MIS. Proceedings of the 2010 Fourth International Conference on Genetic and Evolutionary Computing, Shenzhen, China, 13-15 December 2010; IEEE: New York, NY, USA, 2010; pp. 268–271.
37. Moridi, M.A.; Kawamura, Y.; Sharifzadeh, M.; Chanda, E.K.; Wagner, M.; Jang, H.; Okawa, H. Development of underground mine monitoring and communication system integrated ZigBee and GIS. Int. J. Min. Sci. Technol. **2015**, 25, 811–818. [CrossRef]
38. Baek, J.; Choi, Y. A New GIS-Based Algorithm to Support Initial Transmitter Layout Design in Open-Pit Mines. Energies **2018**, 11, 63. [CrossRef] Appl. Sci. **2020**, 10,
39. Li, M.; Zhang, X.; Mao, S.; Liu, Q. Study of deep mining safety control decision making system. Procedia Earth Planet. Sci. **2009**, 1, 377–383. [CrossRef]
40. S¸ alap, S.; Karsliogˇlu, M.O.; Demirel, N. Development of a GIS-based monitoring and management system for underground coal mining safety. Int. J. Coal Geol. **2009**, 80, 105–112. [CrossRef]
41. Gu, Q.; Lu, C.; Guo, J.; Jing, S. Dynamic management system of ore blending in an open pit mine based on GIS/GPS/GPRS. Min. Sci. Technol. **2010**, 20, 132–137. [CrossRef]

42. Banerjee, T.K.; Roy, S.; Dey, S. A GIS Solution for an Integrated Underground Coal Mine Management: A Conceptual Framework. J. Manag. Policies Pract. **2014**, 2, 129–143.
43. Duncan, E.E.; Rahman, A.A. 3D GIS for mine development – integrated concepts. Int. J. Min., Reclam. Environ. **2015**, 29, 3–18. [CrossRef]
44. Liu, H.; Yang, D. GIS-based mine ventilation network and safety analysis. Proceedings of 2004 IEEE International Geoscience and Remote Sensing Symposium, Anchorage, AK, USA, 20-24 September 2004; IEEE: New York, NY, USA, 2004; pp. 2945–2948.
45. Massanés, M.B.; Pera, L.S.; Moncunill, J.O. Ventilation management system for underground environments. Tunn. Undergr. Sp. Technol. **2015**, 50, 516–522. [CrossRef]
46. Bascompta, M.; Castañón, A.M.; Sanmiquel, L.; Oliva, J. A GIS-based approach: Influence of the ventilation layout to the environmental conditions in an underground mine. J. Environ. Manage. **2016**, 182, 525–530. [CrossRef]
47. Prakash, A.; Vekerdy, Z. Design and implementation of a dedicated prototype GIS for coal fire investigations in North China. Int. J. Coal Geol. **2004**, 59, 107–119. [CrossRef]
48. Choi, Y.; Suh, J.; Park, H.D.; Sunwoo, C. Application of GIS-based Viewshed Analysis for Selecting Optimal Viewpoints and for Evaluating the Sight Distance on a Haul Road in Open-pit mines. J. Korean Soc. Geosystem Eng. **2008**, 45, 265–275.
49. Zhang, M.-y.; Shan, L.-y.; Zhao, Z.-p. The design of coal mine emergency rescue command information system based on web GIS. Proceedings of the 2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), Dengleng, China, 8-10 August 2011; IEEE: New York, NY, USA, 2011; pp. 4455–4458.
50. Choi, Y.; Park, H.D. GIS Modeling for Design of In-pit Stormwater Ponds in Large Scale Open-pit Mines. J. Korean Soc. Geosystem Eng. **2011**, 48, 165–177.
51. Bui, X.-N.; Choi, Y.; Atrushkevich, V.; Nguyen, H.; Tran, Q.-H.; Long, N.Q.; Hoang, H.-T. Prediction of blast-induced ground vibration intensity in open-pit mines using unmanned aerial vehicle and a novel intelligence system. Nat. Resour. Res. **2019**, 28, 1–20. [CrossRef]
52. Li, S.; Dowd, P.A.; Birch, W.J. Application of a knowledge- and geographical information-based system to the environmental impact assessment of an opencast coal mining project. Int. J. Surf. Min., Reclam. Environ. **2000**, 14, 277–294. [CrossRef]
53. Francioni, M.; Salvini, R.; Stead, D.; Giovannini, R.; Riccucci, S.; Vanneschi, C.; Gullì, D. An integrated remote sensing-GIS approach for the analysis of an open pit in the Carrara marble district, Italy: Slope stability assessment through kinematic and numerical methods. Comput. Geotech. **2015**, 67, 46–63. [CrossRef]
54. Nelson, E.P.; Connors, K.A.; Suárez, S.C. GIS-Based Slope Stability Analysis, Chuquicamata Open Pit Copper Mine, Chile. Nat. Resour. Res. **2007**, 16, 171–190. [CrossRef]
55. Choi, Y.; Park, H.D.; Sunwoo, C.; Jung, Y.B. Application of Fuzzy Theory and AHP to Evaluate the Slope Instability at Pasir Open Pit Coal Mine, Indonesia. J. Korean Soc. Miner. Energy Resour. Eng. **2009**, 46, 45–60.
56. Choi, Y.; Park, H.D.; Sunwoo, C. Flood and gully erosion problems at the Pasir open pit coal mine, Indonesia: A case study of the hydrology using GIS. Bull. Eng. Geol. Environ. **2008**, 67, 251–258. [CrossRef]
57. Yucel, D.S.; Yucel, M.A.; Baba, A. Change detection and visualization of acid mine lakes using time series satellite image data in geographic information systems (GIS): Can (Canakkale) County, NW Turkey. Environ. Earth Sci. **2014**, 72, 4311–4323. [CrossRef]
58. Werner, T.T.; Mudd, G.M.; Schipper, A.M.; Huijbregts, M.A.J.; Taneja, L.; Northey, S.A. Global-scale remote sensing of mine areas and analysis of factors explaining their extent. Glob. Environ. Change **2020**, 60, 1–10. [CrossRef]
59. Moomen, A.-W.; Dewan, A. Assessing the spatial relationships between mining and land degradation: evidence from Ghana. Int. J. Min., Reclam. Environ. **2017**, 31, 505–518. [CrossRef] Appl. Sci. **2020**, 10,
60. Duncan, E.E.; Kuma, J.S.; Primpong, S. Open Pit Mining and Land Use Changes: An Example from

Bogosu-Prestea Area, South West Ghana. J. Inf. Syst. Dev. Ctries. **2009**, 36, 1–10. [CrossRef]
61. Wu, Q.; Xu, H.; Pang, W. GIS and ANN coupling model: An innovative approach to evaluate vulnerability of karst water inrush in coalmines of north China. Environ. Geol. **2008**, 54, 937–943. [CrossRef]
62. Boggs, G.S.; Evans, K.G.; Devonport, C.C.; Moliere, D.R.; Saynor, M.J. Assessing catchment-wide mining-related impacts on sediment movement in the Swift Creek catchment, Northern Territory, Australia, using GIS and landform-evolution modelling techniques. J. Environ. Manage. **2000**, 59, 321–334. [CrossRef]
63. Miao, F.; Yan, M.; Qi, X.; Ye, C.; Wang, B.; Liu, R.; Jianhua, C. Application of DInSAR and GIS for underground mine subsidence monitoring. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. **2008**, 37, 251–255.
64. Maryati, S.; Shimada, H.; Sasaoka, T.; Hamanaka, A.; Matsui, K.; Nagawa, H. GIS Database Template for Environmental Management of Mining in Indonesia. J. Geogr. Inf. Syst. **2012**, 4, 62–70. [CrossRef]
65. Álvarez, J.P.F.; Pieiga, A.G.-L.; Suárez-Lázare, C.J. New concepts in reassessing mining heritage: A study and its implications from the Ancient Iron Mine of Llumeres (North Spain). J. Cult. Herit. **2010**, 11, 172–179. [CrossRef]
66. Radulescu, C.; Radulescu, V.M.G. Approaches of the management informational systems regarding the implementation of the geographic information systems (GIS) in the mining basins of Romania. In Proceedings of the 11th International Multidisciplinary Scientific Geoconference and EXPO, Albena Seaside & Spa Resort, Bulgaria, 19–25 June 2011; pp. 215–222.
67. Gorokhovich, Y.; Voros, A.; Reid, M.; Mignone, E. Prioritizing Abandoned Coal Mine Reclamation Projects Within the Contiguous United States Using Geographic Information System Extrapolation. Environ. Manage. **2003**, 32, 527–534. [CrossRef]
68. Fadda, S.; Fiori, M.; Matzuzzi, C. Developing rehabilitation design for the abandoned mine excavations in central Sardinia, Italy. Int. J. Min., Reclam. Environ. **2010**, 24, 286–306. [CrossRef]
69. Kim, S.M.; Choi, Y.; Suh, J.; Oh, S.; Park, H.D.; Yoon, S.H.; Go, W.R. ArcMine: A GIS extension to support mine reclamation planning. Comput. Geosci. **2012**, 46, 84–95. [CrossRef]
70. Pavloudakis, F.; Galetakis, M.; Roumpos, C. A spatial decision support system for the optimal environmental reclamation of open-pit coal mines in Greece. Int. J. Min., Reclam. Environ. **2009**, 23, 291–303. [CrossRef]
71. Menegaki, M.E.; Kaliampakos, D.C. Evaluating mining landscape: A step forward. Ecol. Eng. **2012**, 43, 26–33. [CrossRef]
72. Baek, J.; Choi, Y. Simulation of Truck Haulage Operations in an Underground Mine Using Big Data from an ICT-Based Mine Safety Management System. Appl. Sci. **2019**, 9, 639. [CrossRef]
73. Baek, J.; Choi, Y. Deep Neural Network for Ore Production and Crusher Utilization Prediction of Truck Haulage System in Underground Mine. Appl. Sci. **2019**, 9, 4180. [CrossRef]