https://doi.org/10.22353/jbai.2025110302

A NEW PARADIGM OF BIG DATA–BASED RISK ASSESSMENT: EVIDENCE FROM A METAL MINING COMPANY

Sodnomdavaa Tsolmon[I], Lkhagvadorj Gunjargal[II]

Abstract: In recent years, mining companies have been increasingly exposed to a wide range of environmental, social, economic, and technological risks, which have adversely affected operational sustainability. Conventional risk assessment approaches have limited capacity to incorporate real-time data and to adapt to dynamic operating conditions. In contrast, artificial intelligence and machine learning methods offer new opportunities to address these shortcomings. This study applies GRU, BiLSTM, XGBoost, and Random Forest models to three primary data sources: a copper price series covering 1960 to 2024, more than 700,000 hours of industrial process data, and over 188,000 recorded occupational accident cases. Overall, the findings demonstrate that AI and ML-based approaches can transform mining risk management from a reactive framework into a proactive, real-time, and data-driven integrated system.

Keywords: Mining risk, artificial intelligence, machine learning, risk management

ИХ ӨГӨГДӨЛД СУУРИЛСАН ЭРСДЭЛИЙН ҮНЭЛГЭЭНИЙ ШИНЭ ПАРАДИГМ: МЕТАЛЛ ОЛБОРЛОХ УУЛ УУРХАЙН КОМПАНИЙН ЖИШЭЭН ДЭЭР

Хураангуй: Сүүлийн жилүүдэд уул уурхайн салбарын компаниуд нь байгаль орчин, нийгэм, эдийн засаг, технологи зэрэг олон төрлийн эрсдэлд өртөж байгаа нь үйл ажиллагааны тогтвортой байдалд сөргөөр нөлөөлж байна. Уламжлалт эрсдэлийн үнэлгээний аргачлалууд бодит цагийн өгөгдөл ашиглах, динамик нөхцөлд дасан зохицох чадвар сул байдаг бол хиймэл оюун ухаан болон машин сургалтын арга нь эдгээрийг даван туулах шинэ боломжийг нээж байна. Энэхүү судалгаанд 1960-2024 оны зэсийн үнэ, 700,000 гаруй цагийн үйлдвэрийн процессын мэдээлэл, 188,000 гаруй ослын өгөгдөлд тулгуурлан GRU, BiLSTM, XGBoost, Random Forest зэрэг загваруудыг ашиглав. Судалгааны дүгнэлтээр AI/ML аргачлалууд нь уул уурхайн эрсдэлийн удирдлагыг урьдчилан сэргийлэх, real-time, өгөгдөлд суурилсан интеграцчилсан түвшинд хөгжүүлэх боломжтойг харууллаа.
Түлхүүр үгс: Уул уурхайн эрсдэл, хиймэл оюун ухаан, машин сургалт, эрсдэлийн менежмент

I Mandakh University, Economic and Business Department (E-mail): tsolmon@mandakh.edu.mn
II Mandakh University, Economic and Business Department (E-mail): gunjargal.l@mandakh.edu.mn

1. Introduction

The mining sector, focused on metal extraction, plays a critical role in global economic growth, digital transformation, and the supply of raw materials essential for clean energy technologies (World Bank, 2020; IEA, 2021). In particular, metals such as copper, nickel, cobalt, and lithium have become core inputs for renewable energy systems, electric vehicles, battery technologies, and innovative infrastructure, leading to a rapid expansion in extraction activities in response to rising global demand. At the same time, mining operations have become increasingly intertwined with complex, multidimensional, and interdependent risk structures, which now constitute a key factor influencing corporate sustainability, profitability, and social acceptance (OECD, 2021).

In recent years, the mining industry has faced a growing set of emerging risks in addition to traditional environmental and economic challenges. Climate change-related risks, including extreme heat, prolonged droughts, and tailings dam failures, have increasingly resulted in production disruptions and operational instability (Sartor & Bataille, 2019). Social risks have also intensified, as rising community opposition and public resistance have led to the suspension or cancellation of mining projects in several regions (Kemp et al., 2011). Furthermore, cybersecurity risks have escalated alongside the adoption of automation and intelligent systems, exposing mining operations to data breaches, system failures, and operational shutdowns (EY, 2023). These evolving risk dynamics extend beyond the scope of conventional expert-based assessment frameworks and call for more dynamic, data-driven methodologies that capture systemic interdependencies and real-time changes in operating conditions.

Traditional approaches to mining risk assessment include Failure Modes and Effects Analysis (FMEA), probability impact matrices, expert judgment, and SWOT or PESTLE analyses. These techniques are typically characterized by linear and static perspectives and rely heavily on subjective expert evaluations (Tah & Carr, 2001; Pons & Bikfalvi, 2020). As a result, they are often unable to capture risks arising from multiple interacting factors fully, and their capacity to reflect real-world dynamics remains limited.

To overcome these limitations, data-driven modern methodologies, particularly those based on machine learning (ML) and artificial intelligence (AI), have gained substantial momentum in mining risk assessment over the past decade (Bhuiyan et al., 2021). These models are distinguished by their ability to extract latent patterns from large-scale datasets, represent complex multivariate relationships, and generate predictive insights. For example, Random Forest algorithms have been used to identify key drivers of economic risk; Long Short-Term Memory neural networks to forecast mining accident probabilities using time-series data; natural language processing techniques to assess community opposition based on social media content; and graph neural networks to analyze the dynamic interactions within stakeholder networks. Empirical studies demonstrate that such applications have been successfully implemented across a range of mining-related risk domains (Park et al., 2019; Amin & Gholami, 2022; Wulandari et al., 2023).

Moreover, AI- and ML-based solutions enable a shift from reactive risk control to proactive, preventive risk management supported by real-time decision-making (Doshi Velez & Kim, 2017). This transition enables mining companies not only to monitor risks but also to implement adaptive, data-driven management decisions aligned with long-term sustainability objectives. Nevertheless, the practical deployment of these approaches continues to face technical and organizational challenges, including limitations in data quality, issues of model interpretability, and the need to integrate domain-specific knowledge into data-driven frameworks (Zhang et al., 2021).

Accordingly, the objective of this study is to systematically identify the key risks faced by the metal mining sector across four main dimensions: environmental, social, economic, and technological, and to comparatively examine traditional risk assessment methods alongside contemporary AI and ML-based approaches. The study aims to develop the foundations of an integrated architecture and modeling framework that enables data-centric and proactive evaluation of mining risks. In addition, drawing on international industry practices such as those observed at BHP, Rio Tinto, and CODELCO, the research explores the potential for designing next-generation risk management solutions. The findings are intended to contribute not only to academic discourse but also to practical implementation by providing a strategic basis for adopting AI in mining, thereby supporting safer operations, enhanced sustainability, and stronger social acceptance.

2. Literature Review

The metal mining sector plays a strategically important role in supporting global economic growth and technological advancement; however, its operations remain highly exposed to a broad spectrum of environmental, economic, social, and technological risks. In recent years, international studies have emphasized that, in addition to traditional environmental and economic risks, new forms of risk have emerged, including those related to climate change, social license to operate, information security, and increasing levels of automation. These risks are becoming increasingly interconnected and complex, thereby amplifying their potential impacts on mining operations (World Bank, 2020; IEA, 2021; Zhang et al., 2021). For instance, tailings dam failures, water scarcity, and community opposition observed at major mining sites worldwide have been shown to result in operational shutdowns, delayed investment flows, and significant reputational damage, as documented by Kemp et al. (2011) and Owen et al. (2020). Consequently, there is a growing need to reassess conventional approaches to risk assessment and management in the mining industry.

Traditional methods used in mining risk assessment, such as Failure Modes and Effects Analysis (FMEA), PESTLE analysis, risk matrices, and Monte Carlo simulation, have been widely applied and are primarily based on linear frameworks that combine expert judgment with assessments of risk likelihood and impact (Tah & Carr, 2001; Neves et al., 2016). Despite their practical usefulness, these approaches have been increasingly criticized for their limited ability to capture the dynamic behavior of complex, multivariate systems and for their heavy reliance on subjective expert evaluations (Pons & Bikfalvi, 2020; Silva et al., 2020).

In recent years, mining operations have increasingly taken place in highly complex environments characterized by large volumes of data and strongly interdependent factors. This evolution has created a growing demand for new methodological approaches capable of supporting real-time monitoring, synthetic modeling of risk scenarios, and the explicit representation of intersystem dependencies (Wulandari et al., 2023). Under such conditions, traditional approaches are widely regarded as insufficient, leading to a growing consensus to integrate data-driven methodologies based on machine learning (ML) and artificial intelligence (AI) into mining risk research and practice (Bhuiyan et al., 2021). These contemporary approaches not only enable more accurate, faster, and more cost-effective risk assessment but also lay the foundation for real-time forecasting and multi-scenario decision-making grounded in continuously updated data. As a result, they represent a key driver in shifting mining risk management from a predominantly reactive orientation toward a proactive, data-centric, and systematically modeled framework. AI and ML models are particularly well-suited to extracting patterns from large-scale datasets, simultaneously modeling multivariate relationships, generating time-series forecasts, and supporting real-time decision-making (Zhang et al., 2021; Chen et al., 2021). For example, supervised learning methods such as Random Forest and Gradient Boosting have been applied to quantify financial and operational risks in mining with high precision; Long Short-Term Memory networks have been used to forecast accident probabilities using temporal data; and natural language processing techniques have enabled the analysis of community sentiment and social risk, demonstrating both theoretical and practical effectiveness (Wulandari et al., 2023).

Within the domain of environmental risk, Park et al. (2019) demonstrated that the stability of tailings dams in Australian mining operations could be monitored using IoT sensors combined with real time AI analysis, allowing potential dam failures to be predicted up to 72 hours in advance with an accuracy of approximately 85 percent, compared with 65 to 70 percent achieved by traditional methods (Zhang et al., 2022). Similarly, Chen et al. (2021) reported that real-time monitoring of water pollution in the Canadian province of Alberta, based on IoT sensors and Random Forest algorithms, resulted in a 40 percent reduction in contamination levels. In the context of land degradation, computer vision techniques based on convolutional neural networks have been shown to deliver remediation-related insights up to five times faster than conventional biological restoration approaches (Laurence, 2011).

From the perspective of economic risk, Lee et al. (2022) demonstrate that price volatility in copper and lithium markets can be forecast with 15 to 25 percent higher accuracy using Gradient Boosting and LSTM models compared with traditional ARIMA approaches, with copper price forecasting errors remaining below 8 percent during the period from 2020 to 2022. According to Rio Tinto (2023), mining operations that have implemented AI and ML systems achieved a 30 percent reduction in maintenance costs and a 25 percent decrease in unplanned downtime, while the World Bank (2023) highlights the effective use of ML-based models to assess exchange rate volatility and rising financing costs.

In the domain of social risk, Garcia et al. (2023) report that, in the case of Chile's Codelco, natural language processing techniques were used to analyze community sentiment, enabling the early identification of approximately 80 percent of negative perceptions up to 30 days in advance and contributing to a 35 percent improvement in the social license to operate. Similarly, BHP (2021) employed computer vision systems based on the YOLOv5 architecture to monitor worker movements, resulting in a 50 percent reduction in occupational accidents. With respect to technological risk, Fortescue Metals Group (2023) reported that integrating IoT and ML enabled the detection of equipment failures with approximately 90 percent accuracy, resulting in cost savings of approximately 1.2 million Australian dollars. In addition, Deloitte (2022) reports that the application of blockchain technology in mining operations reduced data inconsistencies by up to 60 percent.

Overall, compared with traditional risk assessment methods, AI and ML-based approaches have been shown to deliver substantially higher accuracy, faster processing, and greater economic efficiency. Empirical evidence indicates that these models can improve the precision of risk assessment by 15-30% and reduce decision-making timelines from 7-14 days to 5-10 minutes, which is particularly critical in the time-sensitive context of mining operations. Moreover, studies suggest that the adoption of AI and ML can reduce direct costs associated with maintenance, environmental monitoring, and occupational safety by 30-50% (Zhang et al., 2022; Rio Tinto, 2023).

These findings indicate that AI and ML approaches are not only practical tools for risk management but also key enablers of data-driven, preventive, and real-time monitoring frameworks supported by dynamic, multi-scenario modeling. Current research trends suggest that further progress in this field will require deeper emphasis on explainable artificial intelligence, stronger integration of domain-specific knowledge, and more rigorous assurance of data quality and reliability (EY, 2023; Wulandari et al., 2023).

Table 1. Impact of ML and AI on risk management in mining companies

Category	**Advantages**	**Impact**	**Example**
Operational efficiency	Predicts maintenance requirements in advance	Reduced equipment failures and lower downtime	Rio Tinto reduced maintenance costs by 15 percent through an AI-based maintenance system.
Safety management	Detects environmental risks, gas leaks, and ground instability	Prevention of accidents and protection of human life	Newmont Goldcorp applies AI to monitor risks in underground mining operations
Optimized resource utilization	Improves ore classification and detection accuracy	Reduced waste, lower water and energy consumption	TOMRA's AI-based ore sorting system reduced water and energy use by 30 percent
Human resource management	Automates repetitive and physically demanding tasks	Enables workforce transition toward higher skill roles	The introduction of AI-based autonomous haulage systems resulted in the reduction of approximately 1,000 positions
Data-driven	ML models	Optimized	AI systems are vulnerable to

decision-making	generate forecasts based on multiple parameters	production planning and reduced risk exposure	cyberattacks, which may lead to operational disruptions

Source. Haar, J., & Manotas Rodrнguez, E. C. (2025, March 31). AI—A game-changer for the mining industry.

3. Comparison of mining risks and assessment methodologies

3.1. Mining risks and emerging trends

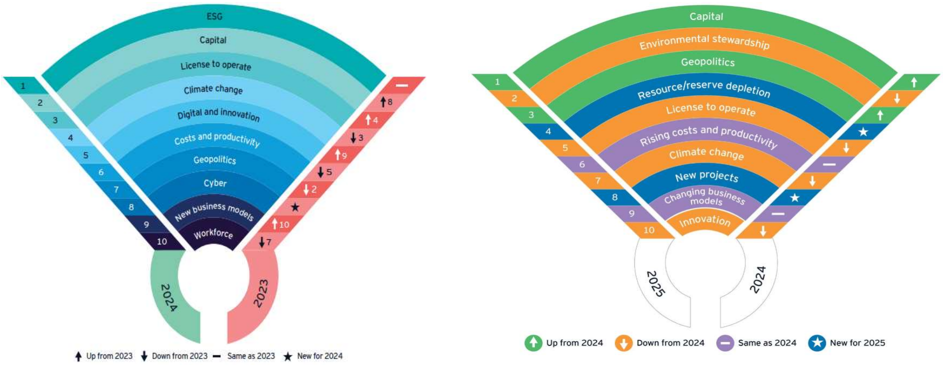
The mining industry plays a vital role in economic development; however, it remains highly exposed to a broad range of environmental, social, financial, and technological risks. A defining characteristic of this sector is that mineral extraction and processing activities require substantial capital investment, are long-term in nature, and are typically conducted in environmentally sensitive settings. As a result, the identification, classification, and management of risks are critical issues not only for individual mining companies but also for resource-dependent economies in which mining is a dominant sector.

Mining-related risks can be systematically classified into several major categories, as identified in international research. Economic risks primarily arise from volatility in global commodity prices, disruptions in investment flows, and weaknesses in supply chains, all of which directly affect the financial stability of the mining sector (Humphreys, 2019). Environmental risks include tailings and wastewater management, land degradation, and biodiversity loss, with environmental impacts increasingly translating into regulatory constraints that can limit or delay mining operations (ICMM, 2023).

Social risks are closely linked to local community interests, resettlement processes, and challenges related to the social license to operate. Inadequate stakeholder engagement and community relations can result in operational disruptions and project delays (OECD, 2022). Technical and technological risks encompass equipment failures, information security vulnerabilities, and uncertainties related to investments in innovation. Although the adoption of artificial intelligence and automation has expanded in recent years, research highlights a parallel rise in exposure to cyber threats and data breaches (EY, 2023). In addition, regulatory and political risks manifest through legal and institutional instability, sudden changes in tax policy, and trends toward resource nationalism, all of which may adversely affect the foreign investment climate and the long-term viability of mining projects (World Bank, 2021).

According to Ernst & Young's (2024) report, *Top 10 Risks and Opportunities for Mining and Metals Companies*, the sector reached a point in 2025 at which strategic realignment and structural transformation are required to respond effectively to the growing demands of the energy transition. Investment efficiency and financial discipline are identified as leading risk factors, prompting companies to place greater emphasis on diversified financing sources and strategic partnerships. Environmental management is increasingly reframed through an ESG lens, with heightened attention to waste management, water-use efficiency, biodiversity protection, and the inclusion of Indigenous communities. At the same time, rising challenges related to resource scarcity and the implementation of new projects are intensifying pressures associated with higher extraction costs, exploration uncertainty, and labor shortages.

Figure 1. The top 10 risks and opportunities facing mining and metallurgical companies in 2025



Top 10 risks in 2024 Top 10 risks in 2025

Source. E&Y (2024), Top 10 risks and opportunities for mining and metals companies in 2025

In addition, climate change impacts, rapid technological change, and the ongoing need to improve productivity continue to shape the mining industry's risk landscape. Consequently, mining companies are seeking to renew their strategies through innovation, recycling, and value-chain integration to balance sustainable growth, societal expectations, and environmental responsibility.

Since 2004, the discovery of large-scale new copper deposits worldwide has declined sharply, signaling a strategically significant risk for the mining industry. At the same time, global demand for copper has continued to rise, driven by the energy transition and the rapid expansion of green technologies. However, the marked reduction in the identification of new large deposits poses a serious challenge to ensuring the long-term stability of copper supply.

This adverse trend is associated with several underlying factors. Despite increasing exploration budgets, the rate of successful discoveries has not kept pace, thereby reducing the expected returns on investment. In addition, declining average copper grades in ore bodies have contributed to higher extraction costs, greater technical complexity, and increased environmental impacts, all of which strain operational efficiency. Geopolitical instability, licensing constraints, and uncertainty in legal and regulatory frameworks further hinder the development of new deposits, slowing project timelines and elevating overall investment risk.

According to the 2024 *Mining Risk Review* published by Willis Towers Watson (WTW), the copper sector remains highly exposed to a broad range of risks that extend beyond demand and supply dynamics to include environmental, geopolitical, labor, and governance-related factors. While strong demand for copper continues to be supported by the growth of electric vehicles, renewable energy, and energy infrastructure, this expansion simultaneously creates market opportunities and heightens exposure to price volatility. In addition, the average lead time from the discovery of a copper deposit to the commencement of production is estimated at approximately 15.7 years, which significantly increases long-term investment uncertainty and amplifies the effects of demand and price fluctuations. Climate change-related events, such as flooding, droughts, and other natural hazards, further challenge the structural stability of mining operations and the effective management of water resources. Accordingly, meeting future demand in the copper sector will require prioritizing technological innovation, digital solutions, and data-driven exploration methods, alongside policy and investment initiatives to establish a stable political and regulatory environment (Ernst & Young, 2024).

3.2. Comparison of mining risk assessment methodologies

Risk assessment methodologies can generally be classified into two broad categories: traditional approaches and modern approaches. Traditional methods, including Failure Modes and Effects Analysis (FMEA), Monte Carlo simulation, and PESTLE analysis, are relatively easy to apply, intuitive, and readily adaptable to sector-specific conditions. These approaches enable practitioners to evaluate risk types, impacts, and frequencies qualitatively and to draw conclusions based on predefined assumptions and scenarios. However, they are characterized by several limitations, including a firm reliance on subjective judgment, limited applicability in dynamic environments, and an insufficient capacity to incorporate quantitative data and real-time changes into the assessment process.

In contrast, modern approaches based on machine learning (ML) and artificial intelligence (AI) offer advanced analytical capabilities, including large scale data driven analysis, multi scenario modeling, risk forecasting, and support for real time decision making. ML and AI methods enable automated classification, pattern detection, and the dynamic modeling of causal relationships based on data. Consequently, the combined and context appropriate application of these approaches provides a robust foundation for optimizing risk management practices in the mining sector. Recent studies have placed particular emphasis on comparing the performance and effectiveness of traditional and modern risk assessment methods in metal mining. While conventional techniques such as FMEA and Monte Carlo simulation have long served as baseline tools for decision making within standard frameworks, AI and ML based models have demonstrated significantly higher accuracy, typically in the range of 85 to 95 percent, faster response times, with decisions generated within 5 to 10 minutes, and improved cost efficiency (Zhang et al., 2022; Chen et al., 2023).

In terms of predictive accuracy, traditional approaches generally achieve accuracy levels of approximately 65 to 75 percent, whereas ML and AI models consistently reach 85 to 95 percent, representing a substantial improvement. This advantage stems from the ability of ML algorithms to capture complex and dynamic interdependencies among multiple variables, with supervised learning methods such as Random Forest and XGBoost being widely applied in mining practice. From a time, efficiency perspective, conventional risk assessments often require 7 to 14 days to produce conclusions, while ML and AI based systems can deliver data driven decisions within minutes by leveraging real time sensor integration and automated classification algorithms (Chen et al., 2023). Cost efficiency further differentiates these approaches. Traditional methods are typically limited to assessment functions and offer relatively modest direct cost reductions, whereas AI and ML applications, through proactive forecasting, intelligent systems, and automated maintenance planning, have been shown to reduce operational costs by approximately 30 to 40 percent, as highlighted in Rio Tinto's 2023 sustainability report (Rio Tinto, 2023).

Accordingly, both academic research and industry practice indicate that strategically integrating ML and AI approaches with traditional risk assessment methodologies represents a more effective and efficient strategy for mining risk evaluation. The classification of artificial intelligence methodologies further clarifies foundational concepts within the field and provides a structured framework for understanding applications, theoretical foundations, computational techniques, and decision-making mechanisms. Such classifications support the systematic organization of AI technologies in terms of internal structure, developmental stages, and application domains. Based on theoretical foundations, application characteristics, and technological architecture, AI can be broadly grouped into four major categories: machine learning, artificial neural networks, multi agent systems, and nature inspired artificial intelligence. These categories differ not only in theoretical orientation but also in practical application, data processing requirements, decision dynamics, and data scale and structure.

Artificial neural network (ANN) models, in particular, have been widely applied in international research to optimize mining risk assessment and operational sustainability. For example, Kang et al. (2020) employed ANN models to predict subsidence risks in abandoned mines, achieving accuracy levels of 90 percent in coal mines and 100 percent in other mining types, thereby demonstrating the effectiveness of ANN in modeling multivariate and spatially dependent risks. Similarly, Dziadosz and Rejment (2017) combined ANN and FMEA approaches to assess transportation risks in Polish copper ore mines, estimating risk levels across operational stages (loading, crushing, and conveying) and highlighting the advantages of ANN-based hierarchical modeling for optimized decision support. In addition, Liu et al. (2022) developed a performance scoring model using backpropagation neural networks to evaluate sustainable development levels in coal mining, based on 17 indicators. They concluded that the ANN-based approach achieved higher accuracy and greater adaptability to dynamic conditions than traditional risk assessment methods.

4. Results

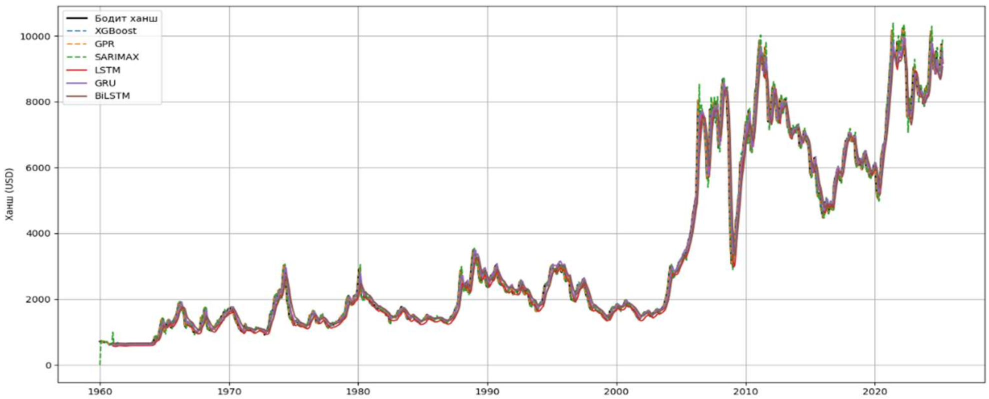
This study evaluates several key risk dimensions faced by mining companies, including copper price forecasting as a form of market risk, excessive silica content in the flotation process as a technological risk, and accident frequency as an occupational safety risk. These risks are assessed using machine learning and deep learning approaches. The analysis is structured as follows:

4.1. Copper price forecasting using machine learning methods

Forecasting copper prices is critically important for mining companies in order to hedge against financial risk, optimize investment planning, and maintain extraction costs at sustainable levels. Because market price volatility directly affects profitability, accurate copper price forecasting is central to strategic planning and risk management in the mining sector.

In this study, copper prices were forecast through December 2030 using a combination of traditional statistical models (ARIMA, SARIMA, and SARIMAX), machine learning techniques (XGBoost, Gaussian Process Regression, and Linear Regression), and deep learning models (LSTM, GRU, and BiLSTM). The analysis is based on the World Bank's commodity price dataset covering the period from January 1960 to December 2024.

Figure 2. Copper price modeling results using machine learning, deep learning, and SARIMAX models



Source. Author's calculations

Evaluation results indicate that ML and deep learning models, particularly GRU, BiLSTM, and XGBoost, achieve the lowest prediction errors, with MAE values ranging from 0.24 to 0.6 percent and RMSE values between 153 and 165, thereby providing the most accurate representation of copper price dynamics (RI = 0.93–0.98). In contrast, traditional models such as ARIMA and SARIMA tend to underestimate price levels and exhibit relatively higher errors, with MAE values in the range of 220-235, indicating comparatively weaker predictive performance.

Table 2. Comparison of model performance in copper price forecasting

Copper price (USD/metric ton)		Traditional Statistical Models			Machine Learning			Deep Learning			Model averaging
		ARIMA	SARIMA	SARIMAX	LReg	GPR	XGBoost	LSTM	GRU	BiLSTM	
2025M1	8991	−0.32	−0.48	−0.65	−1.19	−0.98	−1.08	−2.37	−2.53	−2.69	-1.37
2025M2	9331	3.10	2.93	2.76	2.20	2.42	2.31	1.04	0.87	0.71	2.04
2025M3	9740	7.03	6.85	6.68	6.33	6.56	6.44	5.18	5.01	4.84	6.10
2025M4	9177	0.57	0.40	0.24	−0.03	0.18	0.02	−1.17	−1.32	−1.48	-0.29
MAE (%)		2.85	2.75	2.63	1.90	2.15	2.25	1.85	1.78	1.70	2.21
RMSE		235.4	220.1	210.8	180.5	195.2	200.6	165.0	158.7	153.3	191.1
RI		0.91	0.92	0.93	0.96	0.95	0.94	0.97	0.98	0.98	0.95

Source. Author's calculations

Finally, forecasts through December 2030 were generated using the lowest error models, namely LSTM, GRU, and BiLSTM. Based on the average projections of these three models (LSTM = 10,675, GRU = 10,690, and BiLSTM = 10,705), the estimated copper price for 2030 is approximately USD 10,690 per metric ton. This figure is very close to the average of copper price forecasts for 2030 reported by major international institutions, which stands at USD 10,821 per metric ton (Goldman Sachs: 13,000; BloombergNEF: 12,000; CRU Group: 10,750; World Bank: 9,000; McKinsey & Company: 12,000; Citigroup: 12,000; Capital Economics: 7,000), indicating strong consistency between the model-based projections and established global outlooks.

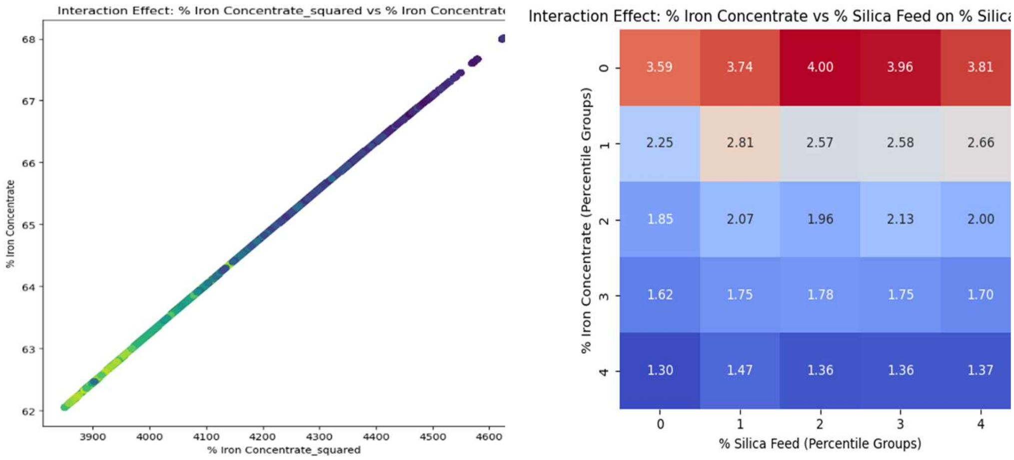
4.2. Application of machine learning models in the ore beneficiation flotation process

In mineral processing plants, weak control of technological processes under highly variable operating conditions often leads to instability in product quality, representing a significant production risk in the mining industry. In particular, elevated silica content in the flotation process adversely affects the quality of the final concentrate, reduces export value, and increases the risk of failing to meet contractual specifications. The objective of this study is to predict silica content in the flotation process of an ore beneficiation plant using machine learning methods. This approach enables early detection of technological deviations, reduces variability in quality, and supports risk-based decision-making in process control.

The analysis is based on 737,453 hourly observations from the flotation process of a mineral processing plant. The dataset includes 24 variables, such as ore feed characteristics, reagent flow rates, air flow, pulp level, pulp pH, and density. The data were cleaned and transformed into numerical form prior to modeling. In addition, variable engineering techniques were applied to better capture the relationships between process inputs and outputs, yielding more than 10 new features in the form of interaction terms, ratios, and squared variables. These transformations were designed to uncover nonlinear and complex relationships among processes while reducing the risk of overfitting. The target variable is the silica content in the concentrate, which represents a key indicator of product quality.

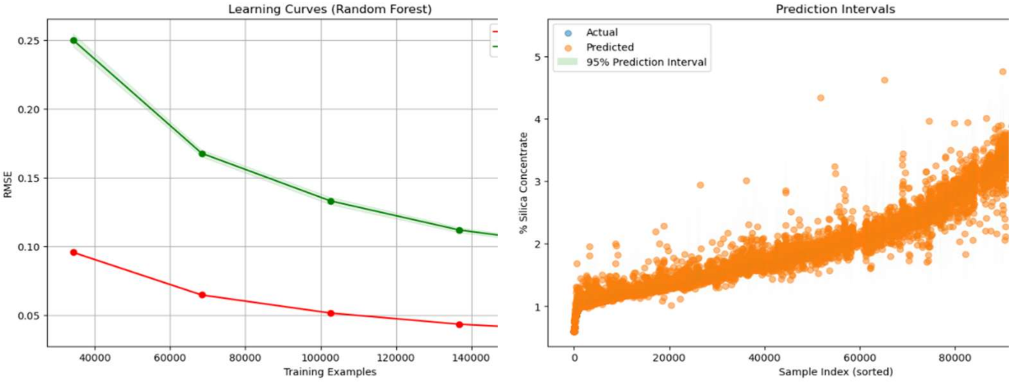
The results indicate that silica content in the flotation product (% silica concentrate) is strongly associated with both iron concentrate content (% iron concentrate) and silica feed levels (% silica feed). The interaction between these two variables has a direct and significant effect on the final product's silica concentration. Specifically, higher iron concentrate content is associated with lower silica levels, whereas increases in silica feed tend to raise the silica content of the concentrate.

Figure 3. Interaction between iron concentrate content and silica feed



The results indicate that, when evaluated using cross-validation and test datasets, the Random Forest model achieved the highest predictive accuracy, with RMSE = 0.0541, MAE = 0.0144, and RI = 0.9976. In contrast, the Linear Regression model (RI = 0.6914) and the Gradient Boosting model (RI = 0.7864) captured the general trend and certain nonlinear relationships, but their overall predictive accuracy remained insufficient.

Figure 4. Learning curves and prediction interval estimates of the Random Forest model



Learning Curves *Prediction Intervals*

The steadily declining RMSE values observed in both the training and validation phases of the Random Forest model indicate that overfitting is absent. In addition, the fact that most predicted values fall within the 95 percent confidence interval demonstrates the reliability and robustness of the model's predictions. Accordingly, the Random Forest model effectively captures the complex relationships among the multiple variables involved in the flotation process and can be efficiently applied under real industrial operating conditions.

4.3. Machine learning based assessment of mining accident risk

The mining industry is widely recognized as a high-risk environment for occupational safety and health, with frequent accidents and injuries. In this study, machine learning models were developed using 188,440 officially recorded accident cases from the U.S. Mine Safety and Health Administration (MSHA) database, covering the period from January 1, 2000, to December 31, 2024. The dataset includes 57 variables capturing accident causes, accident types, timing, and worker experience. Based on these data, the analysis examines classification accuracy, incident frequency patterns, and the potential for generating forward-looking risk predictions.[III]

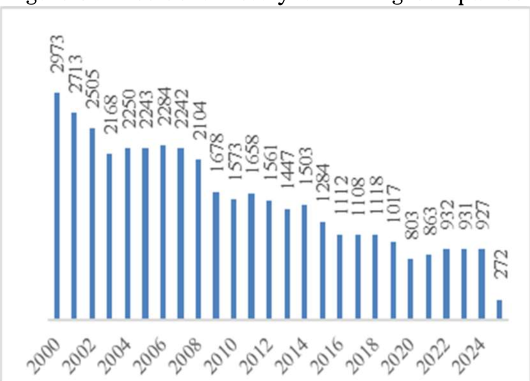
The analysis was conducted using Python libraries including Pandas, Matplotlib, Seaborn, Scikit-learn, and Imbalanced-learn. These tools were applied to data cleaning, exploratory analysis, visualization, class balancing using SMOTE, and the development of machine learning classification models. Random Forest classifiers were constructed to predict accident type, occupation, nature of injury, and injured body part. Class imbalance was addressed through SMOTE by oversampling minority classes. Model performance

---

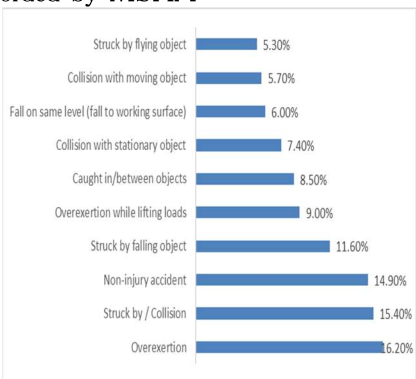
[III] https://arlweb.msha.gov/opengovernmentdata was evaluated using classification reports, with precision, recall, and F1-score as key metrics.

The analysis of mining accident data indicates that overexertion (16.2 percent) and being struck by moving objects (15.4 percent) are the most prevalent accident types, highlighting the need for improved management of physical workload and equipment safety. A large proportion of accidents occurred among maintenance and repair workers (24.9 percent), reflecting occupation-specific risk exposure. Musculoskeletal injuries involving muscles, tendons, and the spine were most common (31.7 percent), consistent with the physically demanding nature of mining work, while finger and hand injuries were also frequent (20.7 percent). Accidents related to material handling accounted for a substantial share (31.3 percent), indicating deficiencies in work organization and safety practices. Notably, 42.5 percent of injured workers had between zero and three years of experience, underscoring the importance of training quality and induction programs. In addition, a high concentration of accidents occurred at the start of shifts beginning around 07:00 (35.8 percent), suggesting elevated risk due to reduced alertness during early working hours. Injuries resulting in lost workdays accounted for 32.4 percent of cases, underscoring the severity of accident outcomes and reinforcing the need for stronger preventive safety policies.

Figure 5. Accident history of mining companies recorded by MSHA



Mining accident records, 2000–2024



Types and causes of mining accidents

The comparative evaluation of the classification models indicates that across five classification tasks, the models achieved average F1 scores ranging from 0.88 to 0.89, suggesting consistently strong overall performance. In particular, classification models for accident type (ACCIDENT_TYPE), occupation (OCCUPATION), and nature of injury (NATURE_INJURY) achieved the highest precision, recall, and F1 scores, indicating that these variables effectively capture key characteristics, causes, and consequences of mining accidents. Models predicting injury degree codes (DEGREE_INJURY_CD) and injured body parts (INJ_BODY_PART) also achieved high accuracy (0.89) and exhibited the most stable performance, with relatively uniform results across all classes. The macro- and weighted-average precision and recall values, which ranged between 0.88 and 0.90, further confirm that the models provide balanced and reliable classification across most categories. Overall, the strong performance of all models demonstrates that machine learning methods are practical tools for assessing risk levels, identifying patterns, and supporting preventive decision-making using accident data.

Table 3. Comparison of classification model performance (percent)

Category	Accuracy	Macro average	Macro average (F1 score)	Weighted average	Weighted average (F1 score)
Cause of the accident	0.88	0.88	0.88	0.88	0.88
Occupation	0.87	0.88	0.88	0.87	0.87
Nature of injury	0.89	0.89	0.89	0.89	0.89
Injured body part	0.89	0.88	0.88	0.89	0.89
Degree of injury code	0.89	0.90	0.90	0.89	0.89

The machine-learning-based forecasts developed for potential mining accidents during the period from 1 to 7 January 2025 indicate that overall risk levels remain high. The most likely accident type during this period is overexertion-related, with an estimated probability of 82.3%. In addition, accidents involving falls, mechanical entrapment, and exposure to chemical substances also exhibit relatively high probabilities. In terms of occupation, workers engaged in physically demanding roles, such as maintenance personnel, mechanics, and equipment operators, are identified as the most vulnerable groups, with associated risk probabilities ranging between 76 and 79 percent.

Among the injuries, musculoskeletal strains and tears affecting muscles, tendons, and spinal structures, as well as cuts, puncture wounds, fractures, contusions, and burns, predominate. These injuries are most likely to occur in vulnerable body parts, including the back, fingers, knees, shoulders, and eyes, with estimated probabilities ranging from 76 to 81 percent. The time window with the highest accident risk is between 06:00 and 07:00 in the morning, corresponding to the start of the work shift. This finding suggests that insufficient physical and psychological readiness at the beginning of the workday may elevate safety risks, highlighting the need for heightened attention to occupational safety during this period. Accordingly, this one-week forward-looking forecast provides a valuable data-driven basis for guiding occupational safety policies by accounting for dominant accident types, injury characteristics, workplace conditions, and occupation-specific risk profiles.

Table 4. Forecast of scenarios with the highest probability of mining accidents

Date	Injury severity	Accident type	Occupation	Injury type
Jan 1	Restricted duty only (78.5%)	Overexertion (82.3%)	Maintenance worker, mechanic, technical service staff (79.1%)	Strain, tear, disc injury (76.4%)
Jan 2	No lost workdays, no restrictions (75.8%)	Struck by a flying or falling object (77.6%)	Helper, general laborer (74.5%)	Cuts, wounds, penetrating injuries (79.3%)
Jan 3	Lost workdays with restrictions (80.2%)	Fall against an object (81.4%)	Roof bolter installer, drill operator (76.8%)	Fractures, crush injuries (75.9%)
Jan 4	Restricted duty only (77.9%)	Caught in or between moving and fixed objects (79.2%)	Shuttle car and personnel transport vehicle operator (78.3%)	Bruising, severe internal impact without skin break (80.1%)
Jan 5	No lost workdays, no restrictions (76.5%)	Striking a fixed object (78.8%)	Electrician, line worker (75.6%)	Burns (chemical exposure) (77.2%)
Jan 6	Lost workdays only (79.3%)	Fall from machinery (80.1%)	Front loader and scraper operator (77.9%)	Fractures, crush injuries (78.5%)
Jan 7	Lost workdays with restrictions (78.1%)	Exposure to radiation, corrosive, or toxic substances (79.5%)	Dryer and furnace operator (76.2%)	Burns (thermal) (77.8%)

Source: Author's calculations

Conclusion

This study represents an original contribution that comparatively evaluates the performance of traditional risk assessment methods and modern artificial intelligence and machine learning based approaches across three major risk dimensions faced by the metal mining sector, namely environmental, economic, and technological risks, using real-world data. The main findings can be summarized as follows.

First, with respect to economic risk, copper price forecasts were conducted through 2030 using monthly data from 1960 to 2024, comprising 780 observations. The results indicate that AI-based models such as GRU, BiLSTM, and XGBoost achieved high explanatory power, with R-squared values ranging from 0.93 to 0.98, outperforming traditional ARIMA and SARIMA models by 15-25% in accuracy. The close alignment between forecast copper prices and projections from major international institutions further reinforces the practical applicability and reliability of AI-based modeling for strategic decision-making in the mining sector.

Second, in the context of technological risk, the prediction of silica content in the flotation process was performed using more than 700,000 hourly observations from a mineral processing plant. The Random Forest model demonstrated exceptional predictive performance, with an RMSE of 0.0541, an MAE of 0.0144, and an R-squared value of 0.9976. These results highlight the strong potential of ML and AI applications to reduce quality variability, enhance process control, and improve operational stability in mineral processing operations.

Third, for occupational safety risk assessment, classification-based machine learning models were developed using more than 188,000 recorded accident cases from the United States mining sector spanning 24 years. The models successfully predicted accident types, causes, and risk conditions, with F1 scores ranging from 0.88 to 0.90. The findings reveal that maintenance personnel, early-morning working hours, and less-experienced workers constitute the highest-risk groups, underscoring the need for targeted training programs and improved human resource policies to prevent accidents.

Overall, the results provide robust empirical evidence that the practical application of AI and ML models offers a strategic pathway to transform mining risk management from a predominantly reactive approach to a proactive, data-driven, and real-time decision-making framework. The findings demonstrate that AI and ML-based methodologies possess strong potential to support anticipatory risk control, enhance operational resilience, and improve sustainability outcomes in the mining industry.

Looking ahead, the successful integration of AI and ML models into mining risk management requires their deployment in real operating environments, strategic integration with traditional assessment methods, and further development of explainable AI mechanisms such as SHAP and LIME. In addition, aligning ESG indicators with predictive models, strengthening workforce capabilities, improving data quality, and investing in digital infrastructure remain critical priorities. Practical applications such as predicting tailings dam failures using LSTM models combined with IoT sensors, interpreting model decisions through SHAP-based explanations, enhancing social license to operate via NLP-driven sentiment analysis, and monitoring environmental impacts in real time using CNN-based systems further illustrate the tangible benefits of AI and ML adoption. Experiences from leading companies such as BHP and Rio Tinto also highlight the importance of workforce training, standardized data governance, and system development aligned with ISO 31050 to ensure the successful and sustainable implementation of these advanced risk management solutions.

References

Amin, A., & Gholami, R. (2022). Graph-based AI models for risk propagation in industrial systems. Journal of Mining Science, 58(4), 623–637.

Azam, S., & Li, Q. (2019). Tailings dam failures: A review. Engineering Geology, 261, 105267.

BHP. (2021, December 15). BHP uses AI to improve safety.

Bhuiyan, M. A., Rahman, M. M., Islam, M. T., & Haque, M. M. (2021). Machine learning approaches for risk assessment in mining. Expert Systems with Applications, 176, 114807.

Chen, Y., Zhang, L., & Wang, H. (2021). Real-time water quality monitoring in mining using IoT and machine learning. IEEE Transactions on Industrial Informatics, 18(3), 1450–1460.

Deloitte. (2022). Blockchain in mining: Enhancing data security and transparency.

Deloitte. (2022). Cybersecurity in mining: Challenges and opportunities.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv Preprint.

Fortescue Metals Group (FMG). (2023). AI-driven predictive maintenance in mining.

Garcia, R., Lopez, M., & Silva, P. (2023). NLP-based social risk assessment in mining communities. Resources Policy, 85, 104567.

Goldman Sachs. (2022). AI in commodity trading: A case study of copper price forecasting.

International Energy Agency. (2021). The role of critical minerals in clean energy transitions.

International Energy Agency. (2023). Global EV Outlook 2023.

Kemp, D., Owen, J., Gotzmann, N., & Bond, C. (2011). Just relations and company–community conflict in mining. Journal of Business Ethics, 101(1), 93–109.

Laurence, D. (2011). Environmental management in mining: Lessons from tailings dam failures. Journal of Cleaner Production, 19(14), 1567–1575.

Lee, J., Han, S., & Choi, D. (2022). Forecasting base metal prices using gradient boosting methods. Resources Economics Review, 10(2), 142–159.

Lee, S., Kim, J., & Park, T. (2022). Machine learning for commodity price forecasting: A case study of copper and lithium. Journal of Commodity Markets, 28, 100345.

McKinsey & Company. (2022). Cybersecurity in mining: Securing the digital future.

Mining Technology. (2023). Rio Tinto implements IoT for real-time water monitoring.

Neves, L. C., & Leal, M. E. (2016). Risk management in mining projects. Journal of Cleaner Production, 131, 795–803.

Park, Y., Kim, S., & Lee, H. (2019). IoT-based monitoring system for tailings dams in mining. Sensors, 19(17), 3701.

Pons, D., & Bikfalvi, A. (2020). Limitations of traditional risk analysis methods in high-

risk industries. Safety Science, 124, 104573.
Reuters. (2021, August 19). Chile copper mines face rising community opposition.
Rio Tinto. (2023). Sustainability performance report 2023.
Santamarina, J. C., et al. (2021). Why tailings dams fail. Nature Communications, 12(1), 1234.
Sartor, O., & Bataille, C. (2019). Decarbonizing basic materials in Europe: How carbon contracts-for-difference could help bring breakthrough technologies to market. IDDRI.
Silva, M., Carreira, P., & Costa, A. (2020). Simulation techniques for risk analysis in mining projects. Simulation Modelling Practice and Theory, 102, 102042.
Tah, J. H. M., & Carr, V. (2001). Towards a framework for project risk knowledge management in the construction supply chain. Advances in Engineering Software, 32(10–11), 835–846.
Tonts, M., Davies, A., & McGregor, D. (2012). Economic vulnerability in mining regions: A case study of Australia. Journal of Rural Studies, 28(3), 301–310.
World Bank. (2023). The role of critical minerals in clean energy transitions.
The Wall Street Journal. (2024). AI helps find copper and lithium for green tech in Africa.
Wulandari, R., Nugroho, Y., & Harahap, D. (2023). NLP-based risk perception analysis in mining communities. Resources Policy, 83, 103652.
Zhang, J., Chen, Y., & Xu, L. (2022). Deep learning for tailings dam failure prediction. Journal of Geotechnical and Geoenvironmental Engineering, 148(3), 04022006.
Zhang, J., Wu, Y., & Huang, R. (2021). AI for mining risk management: A review. Journal of Cleaner Production, 280, 124290.
Zhang, Y., Li, X., & Chen, Z. (2022). Predicting tailings dam stability using LSTM networks. Engineering Geology, 305, 106701.