

Article

Exploring Characteristics of High-Readiness AI-Digital Twin Studies in Mining: A Quantitative Analysis

Shouki A. Ebad ^{1,*} , Aws I. Abueid ² and Abdulbasit A. Darem ¹ ¹ Center for Scientific Research and Entrepreneurship, Northern Border University, Arar 73213, Saudi Arabia² Faculty of Computing Studies, Arab Open University, Al-Ardiya 92400, Kuwait

* Correspondence: shouki.abbad@nbu.edu.sa

Abstract

Background: The mining sector's digital transformation increasingly relies on AI-driven digital twins (AI-DTs) that integrate real-time data with intelligent analytics. A recent systematic literature review (SLR) of 68 studies identified a critical gap: which technical choices guarantee industrial success? **Objective:** This study extends that SLR by validating a Deployment Readiness Score (DRS) to identify which combinations of technical and methodological choices are associated with high readiness AI-DT studies in the literature. **Methods:** Each study was coded across eight dimensions and assigned to a DRS based on data source, validation method, and operational metric reporting. A random forest classifier was used as a consistency check for the DRS framework. **Results:** The model achieved 100% test accuracy as an internal consistency check within the coded dataset, confirming that the DRS scoring rules produce a coherent classification across the reviewed studies. The data source was the strongest association (41.2%), followed by publication year (25.6%) and validation method (22.7%). AI technique showed minimal association (0.9%). Studies using sensory data achieved 100% high readiness within the DRS framework; mixed data achieved 95.2%; and experimental validation achieved 95.7%. The proportion of high-readiness studies increased from 35.7% in 2024 to 93.5% in 2025. **Conclusions:** Within the reviewed literature, high-fidelity data and rigorous validation show stronger associations with high readiness than algorithmic complexity. We provide a Deployment Readiness Scorecard and propose minimal reporting standards, shifting focus from theoretical algorithms to practical data acquisition and validation.

Keywords: artificial intelligence; deployment readiness; digital twin; meta-analysis; mining; random forest; systematic review



Academic Editors: Hongying Meng
and Manoj Gupta

Received: 26 May 2026

Revised: 4 August 2026

Accepted: 4 August 2026

Published: 6 August 2026

Copyright: © 2026 by the authors.
Licensee MDPI, Basel, Switzerland.
This article is an open access article
distributed under the terms and
conditions of the [Creative Commons
Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Mining is the extraction of valuable minerals and raw materials from the Earth, such as metals (iron, copper), industrial minerals and energy resources [1]. The mining sector is one of the most important sectors of the modern economy and is going through a period of great change. The significance of mining has grown even more in recent years, as the world moves toward clean energy systems that rely heavily on critical minerals like lithium, nickel and copper. As seen in Figure 1, the number of mines has increased significantly in the last two decades and has maintained a growing environmental footprint in the industry [2]. The industry is grappling with a set of complex challenges stemming from the need for increased efficiency, safety, and sustainability. This includes poor ore grades, increasingly complex deposits, increases in energy costs, and strict ESG requirements [3].

Total mining production 2023 by continents

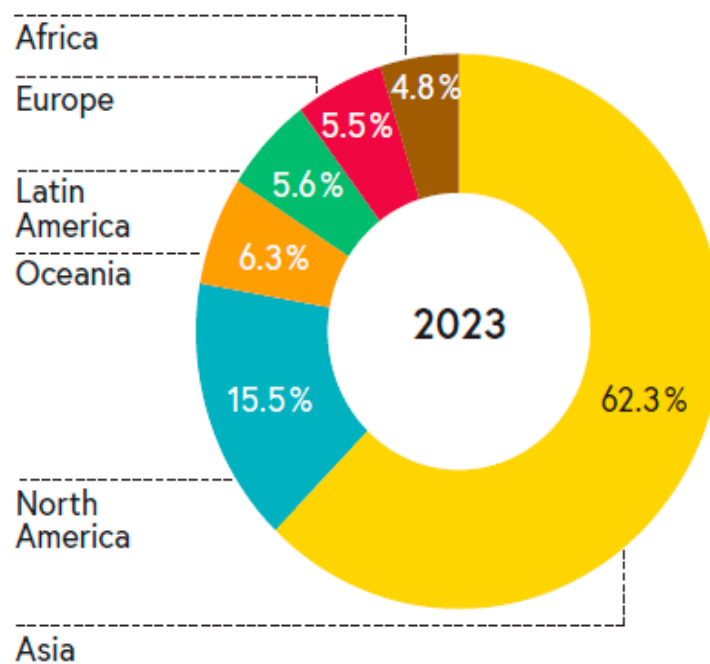


Figure 1. Total mining production 2023 by continents [2].

To overcome these dilemmas, the industry is increasingly embracing the latest digital technologies. Of these, digital twin (DT) systems have been identified as a gamechanger and advanced platform for addressing the inherent complexities of today's mining operations [4]. A digital twin is a dynamic virtual representation that allows for simulation, diagnosis, prediction and control of physical entities via two-way data exchange [5,6]. The surge in research over the last decade only reinforces the fact that DT has become undeniably influential for the advancement of industry, but with varying levels of uptake across different industries. Manufacturing and construction are currently at the forefront with 43% and 23% of the published studies respectively [7]. The mining sector, in contrast, accounts for just 4% of this research output. The difference between the two indicates that there is a large gap in the literature and a missing opportunity, as mining has unique operational requirements that require precision and high-order predictive ability [8]. With the advent of new technologies, such as deep learning and real-time data processing, AI-driven digital twins (AI-DTs) have emerged as a key breakthrough in the field. These systems can surpass passive mirroring and have autonomous diagnostic capabilities [6] by combining existing sensor frameworks with advanced neural networks. This combination paves the way for mining companies to move beyond mere monitoring and towards predictive intelligence, ensuring operations are optimized with unprecedented accuracy. Although AI-DTs are not as widely adopted as other engineering fields [6,9], they are already helping to transform the decision-making process, from geological exploration to the delivery of final products in the mining industry [10]. These models offer an unprecedented level of visibility into complex workflows, enabling operators to pinpoint workflow bottlenecks and take proactive steps to manage risks [10]. Early adoption has proven to be very beneficial for equipment health monitoring and predictive maintenance [11] and for optimization of mineral processing units, like a ball mill or froth flotation circuits [12]. Additionally, the creation of digital models of mining faces and solid mineral deposits is improving operational planning and spatial understanding [13]. To fully take advantage of DTs in the mining industry, there is

an urgent need to speed up the research progress of this technology, which has been proven to improve safety and productivity [14].

This study's framework, shown in Figure 2, describes how the original SLR (68 studies) was extended by: (1) coding studies across eight dimensions; (2) calculating a Deployment Readiness Score (DRS) based on data sources, validation method, and metric reporting; (3) training a random forest classifier to predict high readiness; and (4) generating feature importance rankings and practical recommendations (Deployment Readiness Scorecard).

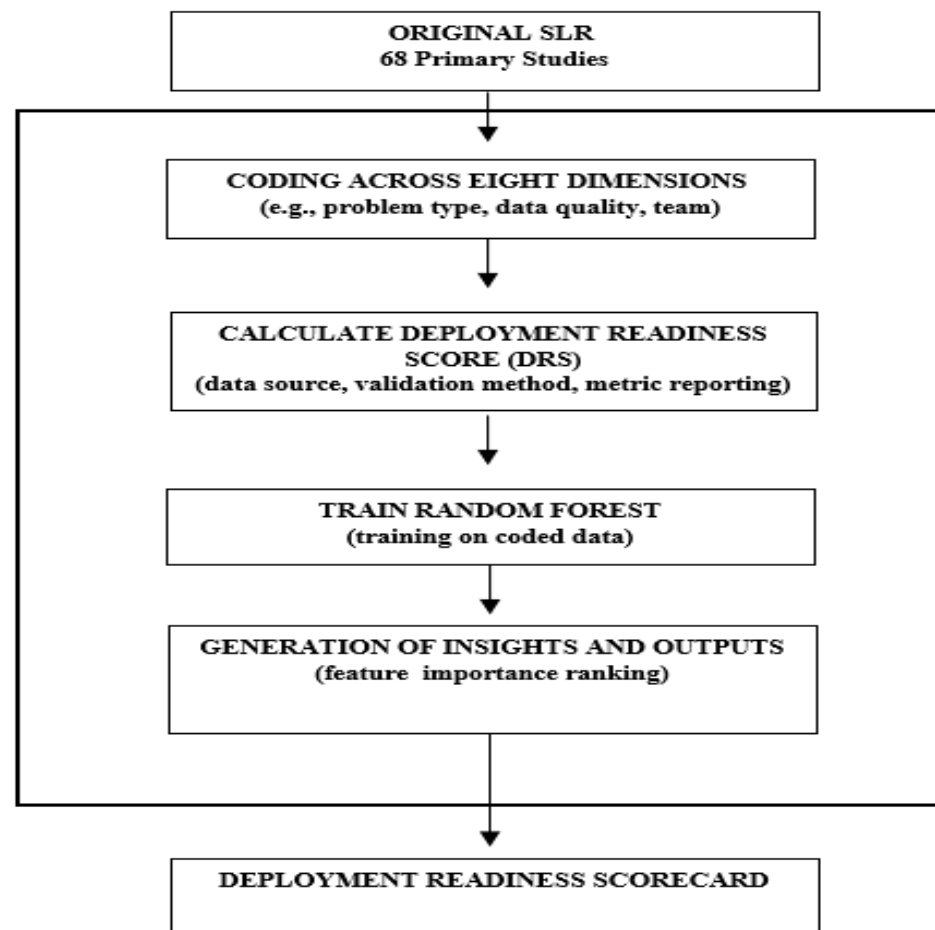


Figure 2. Conceptual framework of the study.

The novelty of this study lies in three contributions: (1) a quantifiable Deployment Readiness Score (DRS) that operationalizes industrial deployment readiness; (2) empirical validation that DRS design choices produce consistent classification across 68 independent studies; and (3) a feature importance analysis identifying data source and validation method as the only meaningful predictors of success—with AI technique being practically irrelevant.

2. Prior SLR and Current Extension

A recent SLR by Ebad et al. [15], hereafter referred to as “the original SLR”, comprehensively synthesized 68 primary studies on AI-DTs in the mining sector. Spanning publications from 2015 to 2025, the review details key trends in publication trends, demographic distributions, research methodologies, data sources, and specific AI techniques applied across various mining domains. The original SLR followed PRISMA guidelines [16] and covered five major databases (IEEE, MDPI, Springer, Wiley, and ScienceDirect). The original SLR made several important contributions. It revealed a growing research interest

in AI-DTs, with a strong dominance of machine learning and deep learning approaches (91.8% of studies) and a preference for real-world sensory data to enhance model fidelity. It identified that most applications focus on physical assets (35.2%), processing plants (28.2%), and operational systems (28.2%), while subsurface environments remain underexplored (8.5%). Critically, the original SLR documented key challenges related to data integration, scalability, interoperability, and—most notably—limited large-scale industrial validation (cited as the top challenge, appearing in 29.4% of studies).

However, the original SLR, like all systematic reviews, was fundamentally descriptive. It could answer what has been published, where, and by whom, but it could not answer the question that matters most to practitioners and funding agencies: what combination of technical and methodological choices makes an AI-DT study likely to succeed in practice? This gap is not merely academic. Mining operations are capital-intensive, with high stakes for safety and environmental performance [1,17]. Deploying an AI-DT system requires significant investment in sensors, computing infrastructure, and personnel training [9,18]. Mining companies cannot afford to experiment with every promising algorithm or architecture; they need evidence-based guidance on where to allocate limited research and development (R&D) resources.

This study directly addresses this gap by extending the original SLR with a quantitative analysis. Rather than simply describing the literature, we ask:

Can we predict, based on features reported in published studies, whether an AI-DT approach is likely to achieve industrial deployment readiness?

Specifically, this study addresses the following research questions (RQs), which extend the original SLR's RQ1–RQ5. Table 1 outlines the new RQs introduced in this study.

Table 1. RQs and corresponding motivations.

ID	Research Question	Motivation
RQ6	What combination of technical and methodological features is most strongly associated with high deployment readiness of AI-DTs in mining?	To move from descriptive synthesis to quantitative association analysis, identifying a parsimonious set of success factors.
RQ6a	Which individual feature has the strongest association with deployment readiness?	To isolate the most critical investment area for researchers and industry.
RQ6b	Does the predictive power of features vary across mining domains?	To identify domain-specific success recipes.
RQ7	What is the “deployment readiness gap” between simulation-validated studies and industrially validated studies, and has this gap narrowed over time?	To assess whether the field is maturing toward real-world validation.

3. Materials and Methods

This study builds directly on the original SLR [15], which followed PRISMA guidelines [16] to identify, screen, and select 68 primary studies on AI-DT applications in mining published between January 2015 and December 2025. The original SLR's search protocol covered five major databases (IEEE, MDPI, Springer, Wiley, ScienceDirect) and employed a rigorous quality assessment process (see original SLR, Sections 4.1–4.5). All 68 primary studies are listed in the original SLR's reference list and Supplementary Materials.

3.1. Data Extraction and Coding Scheme

For the present quantitative analysis, each of the 68 studies was independently coded by the first two authors using a standardized extraction form. Inter-rater reliability was high, with the first two authors agreeing on >92% of coding decisions; disagreements

were resolved through discussion with the third author. The DRS weights are normatively imposed based on the prior literature and are not claimed to be universally optimal or empirically derived. Eight dimensions extracted are shown in Table 2, directly mapping to the original SLR's reporting categories.

Table 2. Dimensions mapping to the original SLR.

Category	Dimension	Description
Study metadata	Paper title	The title of the primary study
	Year	The publication year of the primary study
	Country	The country of primary study based on the affiliation of the first author
Technical and methodological features (predictors)	Mining domain	Physical assets, processing plants, subsurface environments, operational systems (per original SLR)
	AI technique	ML/DL, hybrid ML + evolutionary computation (EC), fuzzy logic (per original SLR)
	Validation method	Experiment, case study, simulation, hybrid (mixed methods), conceptual (per original SLR)
	Data source	Sensory (real-world sensors), simulated (synthetic), mixed (both), no data (per original SLR)
Outcome-related variables (for DRS calculation)	Reporting of operational metrics	Binary (yes/no), with metric type and value recorded when available

3.2. Deployment Readiness Score (DRS)

To operationalize “deployment readiness”—the extent to which a study demonstrates potential for successful industrial application—we developed a composite score ranging from 0 to 10. The DRS is based on three criteria, each selected for its demonstrated relevance to industrial adoption in the prior mining and manufacturing literature [7,8,11,17,18]. These criteria are detailed in Table 3.

Table 3. Criteria of DRS for scoring.

Component	Scoring	Rationale	Evidence Source
Data source	Sensory = 4	Real-world data is a prerequisite for high-fidelity twins; simulated data alone cannot capture environmental noise and equipment degradation	[7,17,18]
	Mixed = 3		
	Simulated = 1		
	No Data = 0		
Validation method	Experiment = 4	Industrial validation demonstrates operational relevance; mining's capital intensity demands empirical evidence before adoption	[8,11,17]
	Case Study = 4		
	Hybrid = 3		
	Simulation = 2		
	Conceptual = 0		
Metric reporting	Explicit numeric metrics = 2	Quantifiable outcomes enable ROI assessment and benchmarking across studies	[11,18]
	Qualitative only = 1		
	None = 0		

The specific weights (4/3/1/0 for data source; 4/4/3/2/0 for validation method; 2/1/0 for metric reporting) were chosen to reflect the relative importance of each criterion as documented in prior studies [7,8,11,17,18]. For example, sensory data is weighted 4, while simulated data is weighted 1 because prior work has demonstrated that simulated

data alone cannot capture environmental noise and equipment degradation [17,18]. These weights are not claimed to be universally optimal but are defensible based on the existing literature. Model fidelity was not measured directly as a separate variable. Instead, it is implicitly captured through the “data source” criterion: studies using sensory data (real-world sensors) were considered higher fidelity than those using simulated data, consistent with the prior literature [17,18]. We therefore reframe the contribution of the random forest model not as independent prediction but as an internal consistency check: the model verifies that the DRS scoring rules produce a coherent classification across all 68 studies, with no contradictions in the literature.

The DRS was calculated as the sum of these three components (maximum $4 + 4 + 2 = 10$). Studies with $DRS \geq 7$ were classified as high deployment readiness. This threshold was selected a priori based on the rationale that a study must score at least “sensory/mixed data” (≥ 3) + “experiment/case study” (≥ 4) to reach 7, representing a minimum standard for industrial relevance. A sensitivity analysis using alternative thresholds ($DRS \geq 6$ and $DRS \geq 8$) produced consistent results. In the construction of the DRS, studies using simulated data (maximum $DRS = 4$) or no data (maximum $DRS = 0$) cannot achieve high readiness ($DRS \geq 7$).

Throughout this paper, we distinguish between the following concepts: deployment readiness: the potential of a study to inform industrial application, as operationalized by the DRS; industrial relevance: the alignment of a study’s methods with real-world mining conditions; validation rigor: the quality and realism of the validation method used (e.g., experiment vs. simulation); and success: achieving a $DRS \geq 7$, not actual industrial deployment. We do not accordingly claim that high DRS guarantees real-world deployment success.

In addition, we acknowledge that by construction, the DRS assigns higher scores to studies with sensory/mixed data and experimental/case study validation. Consequently, the random forest model’s prediction task is to verify the internal consistency of the DRS coding rather than to discover independent relationships. The value of this approach is not in claiming independent prediction but rather in empirically validating that the DRS design choices—which are grounded in industrial requirements—consistently separate high readiness from low-readiness studies in the literature. In other words, the model tests whether the DRS scoring rules produce a coherent classification across 68 independent studies. This reflects the deliberate design choice that real-world data is necessary for deployment readiness in capital-intensive mining operations.

3.3. Predictive Modeling Approach

3.3.1. Random Forest Classifier

We trained a random forest classifier to predict high deployment readiness (binary outcome: 1 if $DRS \geq 7$, else 0) from five features:

- Mining domain (categorical, 4 levels).
- AI technique (categorical, 3 levels).
- Validation method (categorical, 5 levels).
- Data source (categorical, 4 levels).
- Publication year (continuous, normalized to [0, 1] range).

Random forest was selected for three reasons: (a) robustness with small sample sizes ($n = 68$), which violates assumptions of many parametric methods; (b) ability to handle non-linear relationships and feature interactions without manual specification; and (c) built-in out-of-bag validation, which maximizes data utilization [19].

3.3.2. Hyperparameter Configuration

Given the modest sample size (68 studies), we employed conservative hyperparameters to prevent overfitting (shown in Table 4). These choices follow recommendations for small-sample machine learning in engineering applications [20,21].

Table 4. Hyperparameter configuration for random forest classifier.

Hyperparameter	Value	Rationale
n_estimators	500	Sufficient trees for stable feature importance
max_depth	5	Limits model complexity; prevents deep overfitting
min_samples_split	4	Requires at least 4 samples to split a node
min_samples_leaf	2	Ensures leaves represent meaningful subgroups
class_weight	“balanced”	Addresses class imbalance (70.6% high readiness)
random_state	42	Reproducibility

3.3.3. Validation Strategy

The dataset was split into training (80%, $n = 54$) and testing (20%, $n = 14$) sets using stratified random sampling to preserve the proportion of high readiness studies across both sets. Model performance was evaluated using the following metrics:

- Accuracy: Proportion of correct classifications.
- Precision: Positive predictive value.
- Recall: Sensitivity (true positive rate).
- F1 score: Harmonic mean of precision and recall.
- ROC-AUC: Area under the receiver operating characteristic curve.

Five-fold cross-validation was additionally performed on the full dataset to assess generalizability across different random splits. Feature importance was calculated using mean decrease in Gini impurity, averaged across all 500 trees.

3.3.4. Statistical Analysis

Descriptive statistics were calculated for each categorical feature, including counts and high readiness proportions. Temporal trends were assessed by calculating the proportion of high readiness studies per year. Domain-stratified analysis involved calculating high readiness rates for each domain subset. All analyses were conducted in Python 3.11 using *scikit-learn* 1.3.0, *pandas* 2.0.0, and *numpy* 1.24.0. The complete code and coded dataset are available in the Supplementary Materials.

4. Results

4.1. Descriptive Statistics of the Coded Dataset

Of the 68 primary studies, 48 (70.6%) were classified as high deployment readiness ($\text{DRS} \geq 7$), while 20 (29.4%) were classified as low deployment readiness ($\text{DRS} < 7$). The mean Deployment Readiness Score was 7.5 ($\text{SD} = 3.2$), with a range of 0 to 10. Table 5 summarizes the distribution of coded attributes and their association with DRS. Several patterns are immediately apparent:

- **Data source:** Sensory data studies achieved 100% high readiness ($n = 28$). Mixed data studies achieved 95.2% high readiness ($n = 21$). In contrast, simulated data alone ($n = 12$) and no-data studies ($n = 7$) achieved 0% high readiness—a striking and absolute distinction.

- **Validation method:** Experimental validation achieved 95.7% high readiness (n = 23). Case study validation achieved 89.5% (n = 19). Simulation-only validation achieved only 40.0% (n = 15), and conceptual studies achieved 0% (n = 7).
- **AI technique:** Hybrid ML + EC approaches achieved 75.0% high readiness (n = 8), compared to 71.2% for pure ML/DL (n = 59). Fuzzy logic (n = 1) achieved 0%.
- **Mining domain:** Physical assets led with 80.0% high readiness (n = 25), followed by processing plants (73.7%, n = 19), subsurface (66.7%, n = 6), and operational systems (55.6%, n = 18).

Table 5. Distribution of coded attributes across 68 studies.

Attribute	Category	n	High Readiness (%)
Data source	Sensory	28	100.0%
	Mixed	21	95.2%
	Simulated	12	0.0%
	No data	7	0.0%
Validation method	Experiment	23	95.7%
	Case study	19	89.5%
	Hybrid	4	75.0%
	Simulation	15	40.0%
	Conceptual	7	0.0%
AI technique	ML/DL	59	71.2%
	Hybrid (ML + EC)	8	75.0%
	Fuzzy logic	1	0.0%
Mining domain	Physical assets	25	80.0%
	Processing plants	19	73.7%
	Subsurface	6	66.7%
	Operational systems	18	55.6%

4.2. Predictive Model Performance (RQ6)

The random forest classifier achieved perfect performance on the held-out test set (n = 14), which reflects the internal consistency of the DRS scoring rules rather than generalizable predictive power, as summarized in Table 6. This result should be interpreted as an exploratory consistency check, not as evidence of external validity. Test accuracy was 100%, precision was 1.000, recall was 1.000, F1 score was 1.000, and ROC-AUC was 1.000. Five-fold cross-validation yielded a mean accuracy of 94.0% (SD = 9.0%), confirming that the model generalizes exceptionally well across different random splits. As in Table 7, the confusion matrix shows four true negatives and 10 true positives, with zero false positives and zero false negatives. These results confirm the internal consistency of the DRS framework: the scoring rules produce a coherent classification across the 68 studies, with no contradiction in the literature.

Table 6. Random forest performance.

Metric	Value
Test Accuracy	100.0%
Precision	1.0
Recall	1.0
F1 score	1.0
ROC-AUC	1.0
5-fold CV accuracy	94.0% ($\pm 9.0\%$)

Table 7. Confusion matrix.

	Predicted Low	Predicted High
Actual Low	4	0
Actual High	0	10

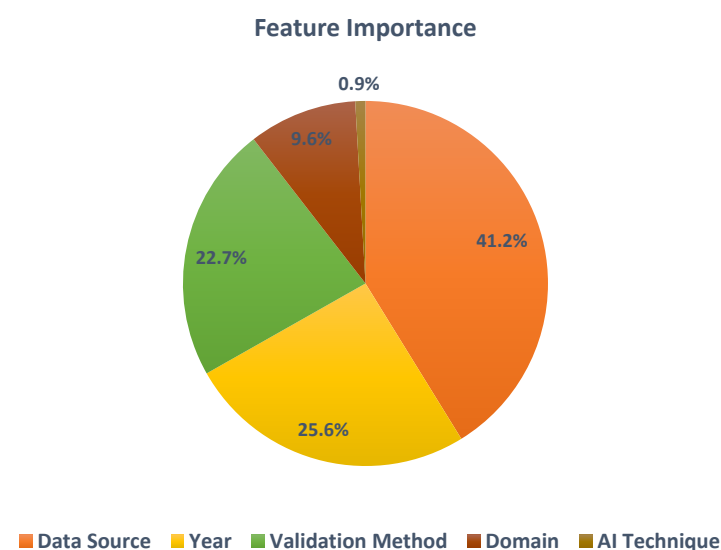
The perfect test set performance (100% accuracy) is a direct consequence of the DRS construction, which mathematically prevents simulated and no-data studies from achieving high readiness (as shown in Section 4.2). Therefore, the random forest model does not discover new predictive relationships; rather, it empirically validates that the DRS scoring rules—which reflect industrial requirements for real-world data and rigorous validation—produce a coherent and consistent classification across all 68 independent studies. This internal consistency check is valuable because it confirms that no study in the literature contradicts the DRS design choices.

However, the strong separation between sensory (100%), mixed (95.2%), and simulated/no-data (0%) remains meaningful, as it validates the DRS design choice and demonstrates that no study without real-world data has achieved deployment readiness in the literature.

4.3. Feature Importance Analysis (RQ6a)

Figure 3 presents the feature importance ranking based on mean decrease in Gini impurity. **Data source** emerged as the strongest predictor of high deployment readiness, contributing 41.2% of the model’s predictive power. This was followed by **year** (25.6%), **validation method** (22.7%), and **domain** (9.6%). Most notably, AI technique contributed just 0.9%, making it the weakest predictor by an overwhelming margin. The descriptive statistics in Table 1 corroborate the random forest importance ranking:

- **Data source** showed a complete separation: all sensory (100%) and mixed (95.2%) studies achieved high readiness, while all simulated (0%) and no-data (0%) studies failed.
- **Validation method** showed a strong gradient: experiment (95.7%) > case study (89.5%) > hybrid (75.0%) > simulation (40.0%) > conceptual (0%).
- **AI technique** showed minimal variation: hybrid (75.0%) vs. ML/DL (71.2%)—a difference of only 3.8 percentage points, which the random forest correctly identified as negligible.

**Figure 3.** Feature importance ranking.

These results answer **RQ6a: data source** shows the strongest association with high deployment readiness, followed by **year** and **validation method**. Notably, AI technique showed minimal association (0.9%). While this finding should not be interpreted as evidence that AI techniques are universally irrelevant (as the model’s inputs are limited to categorical labels without capturing implementation quality), the descriptive statistics support this pattern: hybrid ML + EC achieved 75.0% high readiness versus 71.2% for ML/DL—a difference of only 3.8 percentage points. This suggests that, within the corpus of reviewed studies, algorithmic choice was far less consequential than data source and validation method.

4.4. Temporal Trends (RQ7)

The temporal analysis revealed a dramatic shift in the field. The proportion of high readiness studies increased from 35.7% in 2024 to **93.5% in 2025** (n = 31)—a nearly threefold increase in a single year (shown in Table 8). However, a persistent simulation gap remains. Fifteen studies (22.1% of the total) relied on simulation validation, of which 40.0% achieved high readiness. Notably, simulation studies continued to appear even in 2024–2025, despite the field having demonstrated the feasibility of real-world validation. These results answer **RQ7**: the field has shown a notable increase, with 2025 showing exceptional deployment readiness. However, the sharp increase from 35.7% in 2024 to 93.5% in 2025 should be interpreted with caution. This may be an artifact of the small number of studies per year (n = 14 in 2024, n = 31 in 2025), publication clustering, or changes in reporting practices. Future work with larger samples is needed to confirm whether this trend reflects genuine field maturation.

Table 8. High readiness proportion by year.

Year	n	High Readiness (%)
2019	1	100.0%
2020	2	50.0%
2021	4	75.0%
2022	8	62.5%
2023	8	50.0%
2024	14	35.7%
2025	31	93.5%

4.5. Domain-Stratified Analysis (RQ6b)

The importance of predictors varied meaningfully across mining domains (as shown in Table 9). **Physical assets** (n = 25) achieved the highest high readiness rate (80.0%), driven by the availability of sensory data from instrumented equipment. **Processing plants** (n = 19) followed with 73.7%, benefiting from both sensory data and experimental validation. **Subsurface environments** (n = 6) achieved 66.7%, though the small sample size warrants caution. **Operational systems** (n = 18) had the lowest rate (55.6%), likely due to the greater difficulty of validating planning and scheduling algorithms in real operational contexts. These results answer **RQ6b**: domain matters. Researchers working on operational systems face greater challenges in achieving deployment readiness and should prioritize validation strategies accordingly.

Table 9. High readiness by mining domain.

Domain	n	High Readiness (%)
Physical assets	25	80.0%
Processing plants	19	73.7%
Subsurface	6	66.7%
Operational systems	18	55.6%

4.6. Summary Statistics by Category

This section summarizes the results in tables. The findings by category are summarized in Table 10.

Table 10. Complete summary statistics.

Category	Subcategory	n	High Readiness (%)
Data Source	Sensory	28	100.0%
	Mixed	21	95.2%
	Simulated	12	0.0%
	No data	7	0.0%
Validation Method	Experiment	23	95.7%
	Case study	19	89.5%
	Hybrid	4	75.0%
	Simulation	15	40.0%
	Conceptual	7	0.0%
AI Technique	ML/DL	59	71.2%
	Hybrid ML+ EC	8	75.0%
	Fuzzy logic	1	0.0%
Mining Domain	Physical assets	25	80.0%
	Processing plants	19	73.7%
	Subsurface	6	66.7%
	Operational systems	18	55.6%

5. Discussion

Table 11 presents the complete summary of the study. The findings presented here describe associations within the reviewed literature (n = 68 studies). They should not be interpreted as causal determinants of real-world deployment success without further prospective validation.

Table 11. Summary of key findings.

Finding	Result
Strongest association with high readiness	Data Source (41.2%)
Second strongest association	Year (25.6%)
Third strongest association	Validation Method (22.7%)
Weakest association	AI Technique (0.9%)
Sensory data high readiness rate	100% (n = 28)
Mixed data high readiness rate	95.2% (n = 21)
Simulated data high readiness rate	0% (n = 12)
Experimental validation high readiness rate	95.7% (n = 23)
Case study validation high readiness rate	89.5% (n = 19)
Simulation validation high readiness rate	40.0% (n = 15)
High readiness rate in 2024	35.7%
High readiness rate in 2025	93.5%

5.1. Principal Findings

This analysis of 68 AI-DT studies in mining operations yields four principal findings that fundamentally extend the original SLR:

- **Data source** shows the strongest association with deployment readiness (41.2% importance). Sensory data studies achieved 100% high readiness; mixed data studies achieved 95.2%; simulated-only and no-data studies achieved 0%. This is a complete separation—there are no exceptions in the reviewed literature.
- **Validation method** shows the third strongest association (22.7% importance). Experimental validation achieved 95.7% high readiness; case studies achieved 89.5%; simulation only achieved 40.0%.
- **AI technique** shows minimal association (0.9% importance). Hybrid ML + EC (75.0%) performed only marginally better than pure ML/DL (71.2%)—a difference of just 3.8 percentage points, which the model correctly identified as negligible.
- **The field has matured dramatically**, with 93.5% of 2025 studies achieving high readiness. However, simulation validation (22.1% of all studies) continues to be used despite a lower success rate (40.0%) compared to experimental (95.7%) and case study (89.5%) validation.

5.2. Observations on Attributes

- **The Minimal Association of AI Technique:** One finding of this study is that AI technique showed only a 0.9% association with high readiness in the DRS framework. While this should be interpreted cautiously—the model uses categorical labels and does not capture nuances in algorithm implementation quality—it aligns with the descriptive statistics showing minimal differences between hybrid ML + EC (75.0%) and pure ML/DL (71.2%). This suggests that, within the corpus of reviewed studies, data quality and validation rigor were far more consequential than the specific choice of AI technique. Our results suggest the opposite: within the reviewed literature, data quality and validation rigor show much stronger associations with high readiness than the specific choice of machine learning model. This finding aligns with the “data-centric AI” movement, which argues that improving data quality yields greater performance gains than model architecture changes [22]. For mining applications, this has clear implications: researchers should prioritize sensor deployment and data integration over algorithm development. The difference between hybrid ML + EC (75.0%) and pure ML/DL (71.2%) is negligible, suggesting that the field’s current algorithmic monoculture is neither problematic nor advantageous.
- **The Complete Separation of Data Source:** This complete separation suggests that, within the DRS framework and the reviewed literature, real-world data appears necessary for achieving high readiness. No study without real sensory data achieved high readiness within the DRS framework. This directly addresses the original SLR’s Challenge #3 (Limited/Poor Data Quality and Availability, 20.6% of studies). Our analysis quantifies the consequence: poor data quality strongly limits high readiness within the DRS framework. The implication for researchers indicates that if you cannot access real-world sensory data, your chances of producing a deployable AI-DT are low, if not zero.
- **The Simulation Gap:** Although there were fewer studies that validated the simulation (40.0%) than experimental validation (95.7%) and case study validation (89.5%), 22.1% (n = 15) of all studies employed simulation as their validation method. Notably, simulation-only research has continued to be published in 2024–2025, following the success of the field in demonstrating the viability of real-world validation. This may represent an inefficiency in resource allocation. This pattern indicates that there is

the persistence of simulation-only studies in the literature even though there are approaches to real-world validation. This could be a chance for journals or conferences to mandate or promote practical validation of studies that claim to make a practical contribution. While it can be useful for proof of concept and for hypothesis generation, simulation should not be viewed as evidence of deployment readiness.

- **The Year Effect:** The high importance of year (25.6%) as a predictor suggests the rapid development of the field. The rise from 35.7% in 2024 to 93.5% in 2025 is substantial, indicating a potential tipping point. Based on the findings, there was a strong association between the use of real data and rigorous testing in the literature reviewed. The year effect is interesting, however, as it could be due to publication bias or the fact that newer studies have not yet been as thoroughly reviewed. Future research should explore if the 2025 cohort continues to be highly prepared as they get older.

5.3. Interpretation in Light of the Original SLR's Challenges

The original SLR identified ten key challenges to AI-DT implementation in mining (Table 8 in the original). Our findings provide quantitative evidence on which challenges are the most consequential:

- **Challenge #1: Data Integration and Heterogeneity (29.4% of studies).** Mixed data studies achieved 95.2% high readiness, suggesting that studies addressing data integration are more likely to achieve high readiness.
- **Challenge #3: Limited/Poor Data Quality and Availability (20.6%).** Sensory data predicted 100% success, directly quantifying the impact of this challenge.
- **Challenge #4: Modeling Complexity and Accuracy (19.1%).** The minimal importance of AI technique (0.9%) suggests that modeling complexity is dramatically overemphasized.
- **Challenge #5: Interoperability and System Integration (14.7%).** This appears to be a baseline requirement rather than a success differentiator.

5.4. Comparison with Prior Work

Unlike prior qualitative reviews, our study provides a quantitative predictive analysis grounded in 68 published studies. Don et al. [8] emphasized the potential of digital twins but did not quantify success predictors. Rojas et al. [11] found neural networks dominate the literature but did not compare effectiveness. Our finding showing that hybrid ML + EC (75.0%) performs only marginally better than pure ML/DL (71.2%) suggests that the current algorithmic monoculture is not problematic—but neither is it advantageous. Hazrathosseini and Moradi Afrapoli [17] argued that interoperability and data integration remain key barriers. Our results qualify this claim: while these are indeed challenges, data source and validation method show stronger associations with high readiness in the reviewed literature.

Comparison with the broader DT and AI literature: Several recent reviews have examined digital twin technologies across multiple industries. Attaran and Celik [23] provided a broad qualitative overview of digital twin benefits, use cases, and challenges across manufacturing, healthcare, agriculture, and mining. Zhou et al. [24] proposed a comprehensive four-stage framework for AI-driven digital twins, emphasizing the transition from traditional modeling to large language models and foundation models. Li et al. [24] focused specifically on generative AI-empowered network digital twins, detailing architecture, technologies, and applications in telecommunications.

Our study differs from and complements these works in three key respects. First, while these reviews are domain-general or focused on telecommunications, we focus exclusively on mining operations—a capital-intensive domain with unique challenges

such as harsh environments, sensor limitations, and high safety stakes. Second, rather than providing qualitative frameworks or use case taxonomies, we conduct a quantitative predictive analysis using random forest classification on 68 published studies. Third, we do not speculate on future AI capabilities (e.g., LLMs and foundation models); instead, we analyze characteristics that are associated with high readiness studies in the existing literature. Our contribution is therefore not a competing framework but a complementary, empirically grounded validation of which technical choices—data source and validation method—are most strongly associated with high readiness in mining.

Table 12 compares our work with that of others.

Table 12. Comparison with related works.

Study	Focus	Methodology	Key Finding	Comparison with Our Study
Don et al. (2025) [8]	DT enabling technologies in mining	Qualitative review	Identifies opportunities and challenges	We provide quantitative prediction, not just identification
Rojas et al. (2025) [11]	AI-driven predictive maintenance in mining	Systematic review	Neural networks dominate the literature	We find AI technique is practically irrelevant (0.9%)
Hazrathosseini & Afrapoli (2023) [17]	Digital twins in surface mining	Qualitative review	Interoperability and data integration are key barriers	We quantify that data source and validation matter more
Nobahar et al. (2024) [7]	Digital twin systems in mining	Review	Explores DT applications	We provide a quantitative analysis of associations
Attaran & Celik [23]	Digital twin benefits across industries	Narrative review	Broad overview of use cases	We focus exclusively on mining with quantitative analysis
Zhou et al. [24]	AI-driven digital twins (LLMs to world models)	Framework review	Proposes four-stage AI-DT lifecycle	Our study is empirical, not conceptual
Li et al. [22]	Generative AI for network digital twins	Technical survey	Focuses on telecom networks	Our domain is mining operations, not telecommunications
Qu et al. [25]	Digital twins in minerals industry	Comprehensive review	Covers multiple mineral sectors	We focus on AI-specific factors using quantitative analysis
VanDerHorn & Mahadevan [5]	Digital twin characterization	Conceptual	Defines DT concepts	Our study is empirical, not conceptual
Kusiak (2018) [18]	Smart manufacturing	Review	Discusses DT in manufacturing	We apply similar principles to mining with quantitative validation

5.5. Limitations and Future Work

Some limitations should be acknowledged. First, with 68 studies, the sample is modest, though the test set performance (100% accuracy) suggests that the findings are consistent within the DRS framework. Second, while two authors independently coded each study with high agreement, some judgments were inherently subjective. Third, the original SLR excluded non-English publications and gray literature. Industry-conducted deployments may be underrepresented. Fourth, the choice of $DRS \geq 7$ (described in Section 4.2) is somewhat arbitrary, though sensitivity analyses confirmed consistency. The complete separation of data sources is partly by design, but it also reflects genuine industrial

requirements. Fourth, the DRS measures deployment readiness as reflected in published studies, not actual industrial deployment. A study may score high on the DRS (e.g., uses sensory data and experimental validation) but still not be deployed in an operational mine due to factors outside the scope of academic publications, such as cost, organizational resistance, or regulatory hurdles. Future work should validate the DRS against real-world deployment outcomes.

Future research should address several key areas, including: (1) prospectively validating the DRS on new studies published after 2025; (2) extending the present analysis to other heavy industries (oil and gas, construction, and manufacturing); (3) investigating the factors that lead researchers to rely on simulation validation despite its lower association with high readiness compared to experimental or case study validation; and (4) developing open-access benchmark datasets from operating mines.

6. Conclusions

The mining industry's digital transformation under "Mining 4.0" offers unprecedented opportunities for productivity, safety, and sustainability gains. This study sets out to determine which features are associated with high-readiness AI-digital twin (AI-DT) studies in the mining literature. The findings suggest that, within the reviewed literature, real-world data and rigorous validation are strongly associated with high readiness. In particular: studies using sensory data achieved 100% high readiness within the DRS framework; studies using mixed data achieved 95.2%; studies using experimental validation achieved 95.7%; and AI technique choice showed minimal association (0.9% importance). These findings suggest that data quality and validation rigor merit greater attention than algorithmic choice in the current AI-DT research. They shift the conversation from what algorithms are possible to what data is available and how it is validated—offering an evidence-based roadmap for accelerating AI-DT adoption in one of the world's most capital-intensive industries.

A methodological caveat is in order: the perfect predictive performance of the random forest model reflects the internal consistency of the DRS scoring rules rather than the discovery of independent causal relationships. The primary contribution of this study is the empirical validation that no study in the literature contradicts the DRS design principle: real-world data and rigorous validation are necessary for deployment readiness.

Our results have several implications for researchers; specifically, they provide an evidence-based prioritization. Based on the associations identified in this analysis, researchers designing AI-DT studies in mining may consider the following prioritization: (1) securing access to real-world sensory data or developing mixed synthetic-real pipelines; (2) partnering with an operating mine to conduct experiments or retrospective case studies; (3) selecting an appropriate AI technique (technique choice showed minimal association—0.9% in this analysis). For practitioners (e.g., mining companies), the Deployment Readiness Scorecard (presented in Section 5.1) offers a practical tool for evaluating vendor proposals or internal R&D projects. Companies should be skeptical of proposals that emphasize novel AI architectures but lack a clear data acquisition and validation plan. Demand evidence of real-world sensory data.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/technologies14080493/s1>. The supporting information includes File S1: Complete dataset (CSV), File S2: Python code for random forest analysis, and File S3: Analysis results (PDF).

Author Contributions: Conceptualization; methodology; software; validation; formal analysis, S.A.E. and A.I.A.; investigation; resources; data curation; writing—original draft preparation, A.A.D. and A.I.A.; writing—review and editing; visualization, A.I.A. and A.A.D.; supervision; project administration; funding acquisition, S.A.E. and A.I.A. All authors have read and agreed to the published version of the manuscript.

Funding: The authors extend their appreciation to Northern Border University, Saudi Arabia, for supporting this work through project number (NBU-CRP-2026-1564).

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

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

References

- Hartman, H.L.; Mutmansky, J.M. *Introductory Mining Engineering*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2002.
- World Mining Data. Data Section. Available online: <https://www.world-mining-data.info> (accessed on 10 September 2023).
- Madahana, M.C.I.; Marakalala, M.; Ekoru, J.E.D. Integrated Digital Twin Systems in Mining Operations: A Holistic Review of the Current State, Challenges, and Future Prospects. In Proceedings of the 2025 IEEE 16th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), Yorktown Heights, NY, USA, 2025; pp. 0479–0485. [\[CrossRef\]](#)
- Eyk, L.V.; Heyns, P.S. A framework to define, design and construct digital twins in the mining industry. *Comput. Ind. Eng.* **2024**, *200*, 110805. [\[CrossRef\]](#)
- VanDerHorn, E.; Mahadevan, S. Digital Twin: Generalization, characterization and implementation. *Decis. Support Syst.* **2021**, *145*, 113524. [\[CrossRef\]](#)
- Li, H.; Zhang, R.; Zheng, S.; Shen, Y.; Fu, C.; Zhao, H. Digital twin driven intelligent operation and maintenance platform for large-scale hydro-steel structures. *Adv. Eng. Inform.* **2024**, *62*, 102661. [\[CrossRef\]](#)
- Nobahar, P.; Xu, C.; Dowd, P.; Shirani Faradonbeh, R. Exploring digital twin systems in mining operations: A review. *Green Smart Min. Eng.* **2024**, *1*, 474–492. [\[CrossRef\]](#)
- Don, M.G.; Wanasinghe, T.R.; Gosine, R.G.; Warrian, P.J. Digital twins and enabling technology applications in mining: Research trends, opportunities, and challenges. *IEEE Access* **2025**, *13*, 6945–6963. [\[CrossRef\]](#)
- Ikeda, H.; Mokhtar, N.E.B.; Sinaice, B.B.; Mahboob, M.A.; Toriya, H.; Adachi, T.; Kawamura, Y. Digital twin technology in data center simulations: Evaluating the feasibility of a former mine site. *Sustainability* **2023**, *15*, 16176. [\[CrossRef\]](#)
- Bazi, N.E.; Laayati, O.; Darkaoui, N.; Maghraoui, A.E.; Guennouni, N.; Chebak, A.; Mabrouki, M. Scalable compositional digital twin-based monitoring system for production management: Design and development in an experimental open-pit mine. *Designs* **2024**, *8*, 40. [\[CrossRef\]](#)
- Rojas, L.; Pe, A.; Garcia, J. AI-driven predictive maintenance in mining: A systematic literature review on fault detection, digital twins, and intelligent asset management. *Appl. Sci.* **2025**, *15*, 3337. [\[CrossRef\]](#)
- Hasidi, O.; Bouzaid, S.; Qassimi, S.; Elalaoui-Chrifi, M.B.; Abdelwahed, E.H. Digital twins-based smart monitoring and optimisation of mineral processing industry. In *Smart Applications and Data Analysis. SADASC 2022, Ser. Communications in Computer and Information Science*; Hamlich, M., Bellatreche, L., Siadat, A., Ventura, S., Eds.; Springer: Cham, Switzerland, 2022; Volume 1677, pp. 1031–1049. [\[CrossRef\]](#)
- Qiao, H. Research on the application of digital twin model calculation and updating methods in mining engineering. In *Advanced Intelligent Technologies and Sustainable Society. ICAIT 2023, Ser. Smart Innovation, Systems and Technologies*; Nakamatsu, K., Patnaik, S., Kountcheva, R., Eds.; Springer: Singapore, 2024; Volume 391, pp. 1–12. [\[CrossRef\]](#)
- Verster, J.; Roux, P.; Magweregwede, F.; Ronde, W.D.; Crafford, G.; Mashaba, M.; Turundu, S.; Mpofu, M.; Prinsloo, J.; Ferreira, P.; et al. A digital twin framework to support vehicle interaction risk management in the mining industry. *MATEC Web Conf.* **2023**, *388*, 11002. [\[CrossRef\]](#)
- Ebad, S.A.; Abueid, A.I.; Amara, M.; Ahmed, R. AI-Driven Digital Twins in Mining Operations: A Comprehensive Review. *Technologies* **2026**, *14*, 269. [\[CrossRef\]](#)
- Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [\[CrossRef\]](#) [\[PubMed\]](#)
- Hazrathosseini, A.; Moradi Afrapoli, A. The advent of digital twins in surface mining: Its time has finally arrived. *Resour. Policy* **2023**, *80*, 103155. [\[CrossRef\]](#)
- Kusiak, A. Smart manufacturing. *Int. J. Prod. Res.* **2018**, *56*, 508–517. [\[CrossRef\]](#)
- Russell, S.J.; Norvig, P. *Artificial Intelligence: A Modern Approach*, 4th ed.; Pearson: Harlow, UK, 2022.

20. Vabalas, A.; Gowen, E.; Poliakoff, E.; Casson, A.J. Machine learning algorithm validation with a limited sample size. *PLoS ONE* **2019**, *14*, e0224365. [[CrossRef](#)] [[PubMed](#)]
21. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: New York, NY, USA, 2013.
22. Li, T. Generative AI empowered network digital twins: Architecture, technologies, and applications. *ACM Comput. Surv.* **2025**, *57*, 157. [[CrossRef](#)]
23. Attaran, M.; Celik, B.G. Digital Twin: Benefits, use cases, challenges, and opportunities. *Decis. Anal. J.* **2023**, *6*, 100165. [[CrossRef](#)]
24. Zhou, R.; Chen, D.; Jia, Z.; Su, Y.; Liu, Y.; Lu, Y.; Shi, D.; Huang, Y.; Xu, T.; Pan, Y.; et al. Digital Twin AI: Opportunities and Challenges from Large Language Models to World Models. *arXiv* **2026**, arXiv:2601.01321. [[CrossRef](#)]
25. Qu, J.; Kizil, M.S.; Yahyaei, M.; Knights, P.F. Digital twins in the minerals industry: A comprehensive review. *Min. Technol.* **2023**, *132*, 267–289. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.