SSRG International Journal of Mechanical Engineering
ISSN: 2348-8360/ https://doi.org/10.14445/23488360/IJME-V13I5P108

Volume 13 Issue 5, 116-128, May 2026

Original Article

Artificial Intelligence Integration in the Eight Disciplines (8D) Problem-Solving Methodology: AI-Enhanced Problem Classification at D1 with Complete Manual Execution of D2–D8 — A Manufacturing Case Study

Suyash Kamble[1,2*], Dadarao Raut[3], Dattaji Shinde[4], Pratik Gaikwad[5]

[1,3,4]*Department of Production Engineering, Veermata Jijabai Technological Institute, Mumbai, India.*
[2]*Department of Mechanical Engineering, Bharati Vidyapeeth's College of Engineering, Kolhapur, India.*
[5]*Department of Computer Science and Engineering, Bharati Vidyapeeth's College of Engineering, Kolhapur, India.*

**Corresponding Author : suyash.kamble@bharatividyapeeth.edu*

Received: 16 February 2026 Revised: 26 March 2026 Accepted: 25 April 2026 Published: 29 May 2026

Abstract - *The Eight Disciplines (8D) is a structured approach to problem-solving that is used in process and manufacturing quality management. Despite the popularity of the methodology, it remains that problem description and classification (D1) continues to be a bottleneck in its implementation, taking an outsized share of manual labor compared to analytical yield. This paper details a full-fledged end-to-end 8D implementation of a pulley fettling defect occurring in a small to medium casting and manufacturing firm at Kolhapur(Maharashtra, India). Only automated D1 (Problem Classification) with a dedicated AI system, while leveraging traditional manual methods for disciplines D2 through D8 to maintain the team-based and human-centric nature of the 8D framework. The AI method is performed through a classification algorithm powered by a Large Language Model (LLM) with rigorously defined composite confidence metric* $\psi=\alpha\cdot\psi ssm+\beta\cdot\psi struct + \gamma\cdot\psi cons$ *with empirically optimised weights* ($\alpha = 0.45$, $\beta = 0.35$, $\gamma = 0.20$)*. The performance of the AI system for the D1 classification demonstrates good accuracy, with a corresponding sensitivity (accuracy) of 89.7%, departmental recall of 92.1%, precision of 87.3%, F1-score equal to 0.895 and mean confidence score of ~0.847; Time taken for D1 phase was decreased from four working days (manual baseline) to 30 minutes resulting in a time saving of 99.0%. The latter manual 8D disciplines (D2–D8) are ideally bounded, including setting up the cross-functional team, taking interim containment actions, performing an Ishikawa fishbone root cause analysis, planning and implementing corrective actions as well as preventive measures in place, and recognizing the team. Quantitatively, these outcomes show that the AI-enabled D1 provided a clearer, more specific, and well-formed problem definition for which downstream manual work can be both higher quality and require less time to complete. In total, the 8D cycle time itself improved from a prior state average gap of 63 days to 41 days — about a modest improvement of 34.9 percent directly attributable to AI at D1 level improvements alone. Based on the findings, process performance outcomes validate that the corrective actions are effective: (i) fettling cycle time was reduced by 74 per cent, monthly production capacity was increased by 58.8 per cent, operator fatigue improved by 73 per cent, and (ii) there have been no incidents of operator injuries. This paper creates the groundwork for a five-paper series on discipline-specific concepts for integrating AI-8D.*

Keywords - *Artificial Intelligence, Eight Disciplines (8d), Industry 5.0, Large Language Model, Quality Management Problem Classification.*

1. Introduction

The Eight Disciplines (8D) methodology is one of the most thoroughly validated structured approaches for resolving quality problems in manufacturing. First developed at Ford Motor Company in 1987 as a component of its Team-Oriented Problem Solving (TOPS) programme [1], the eight-step framework guides cross-functional teams from problem identification through permanent corrective action to team recognition. The methodology has since been adopted across automotive, aerospace, electronics, and general manufacturing sectors globally.

Despite its widespread adoption, empirical studies systematically highlight critical inefficiencies — particularly at the problem description and classification phase (D1) — which act as the primary directional input for all downstream disciplines [2, 3]. Vague or subjectively framed D1 problem statements are the most commonly cited cause of 8D process failure [1]. Biban [9] confirmed that D1 quality is the primary determinant of overall 8D efficiency. According to Bewoor [2], "D1 takes a disproportionate amount of manual effort versus its analytical richness." This challenge is most pronounced in the case of Small-To-Medium Enterprises (MSMEs), where resources specifically for quality engineering are scarce.

With the introduction of Artificial Intelligence (AI), and in particular Large Language Models (LLMs) that support natural language understanding, there is an opportunity to tackle challenges within D1 at the same time as maintaining a collaborative, group-based process, which helps provide structure to the other 8D disciplines. Yet, the prior literature uncovers a crucial gap: no formally defined AI classifier algorithm for the D1 (Design) discipline within 8D — along with a precisely specified measure of confidence on selection states — has previously been published. Also, no study has investigated the possible downstream impact of D1 quality enhanced by AI on the efficiency of manually conducted steps ranging from D2 to D8.

This paper bridges that gap by presenting an extensive 8D case study for a pulley fettling defect detected at a casting and manufacturing plant in Kolhapur, Maharashtra, India. A hybrid 8D execution model is demonstrated: D1 (Problem Classification) is solved using a formally specified AI algorithm, while D2 through D8 use manual execution consistent with traditional 8D practice. This design enables a direct intra-case comparison of AI-enhanced versus manual D1 stage outcomes, and quantifies the downstream impact of improved D1 quality on the efficiency of subsequent manual steps.

1.1. Research Gap

Three considerations motivate this research. First, no formal prior work exists for an AI classification algorithm underlying the 8D D1 discipline with a mathematical formulation and empirical demonstration. Second, to the authors' knowledge, no study has examined the downstream effect of an AI-enhanced D1 quality on the efficiency of manually executed D2–D8 steps. Third, D1 is most arduous in the MSME context — where there is a scarcity of resources — yet has not received substantial attention in AI quality management research.

1.2. Research Objective

The primary objective is to demonstrate, through a complete real-world 8D case study, that targeted AI integration at D1 — with all other disciplines executed manually — produces a measurable improvement in total 8D cycle time, D1 output quality, and downstream process efficiency, while fully preserving the human-centric collaborative character of the 8D methodology. A secondary objective is to justify the deliberate limitation of AI to D1 only in this foundational paper, and to establish the experimental design for a five-paper series that will progressively extend AI integration to other 8D disciplines.

1.3. Research Contributions

- A formally defined AI Problem Classification Algorithm with a confidence scoring metric $\psi j(D1)$.
- Quantitative evidence of 89.7% classification accuracy with 87.3% precision, 0.895 F1-score, and a time saving of 99·0% (4 days → 30 min) at D1
- A sample 8D case study (with D1 enhanced by AI and D2–D8 completed manually)
- Empirical evidence that better D1 quality reduced the total 8D cycle time by a whopping 34.9 per cent
- Rationale behind a domain-specific synthetic intelligence integration strategy that's consistent with Industry 5.0 tenets
- Basic experiment design for a future five-paper proposal on AI-8D integration.

2. Literature Review

2.1. The 8D Methodology and Its Limitations

The eight-discipline approach has been widely explored in the literature for various manufacturing sectors. Out of the performed systematic study by Mahmood [1] on 8D applications in production companies, sprawling or subjectively framed problem descriptions at D1 are the most frequent reasons for failure of the 8D process. It was confirmed in a study [9] (Biban) that D1 quality is recognized as the main factor influencing 8D efficiency. In a comparison of 8D and DMAIC frameworks, Bewoor [2] also highlighted that D1 consumes disproportionate manual effort for its level of analytics. Sharma [4] showed that using a combination of D1 (Defining the problem) to class unstructured data, and 8D with Six Sigma statistical tools led to better quality of analysis for root cause analysis on D4 (Determination of root causes), but lacked any mechanism for automating D1 classification. Structured collection of those data at D1 measurably improved the effectiveness of downstream corrective action (Realyvásquez-Vargas et al.5). Cyganiuk et al. The D4 analysis of ideal escape points has been defined as the analytically most complex 8D discipline [19], and it was concluded that its adequacy is simply in a direct relationship with the accuracy of the D1 problem statement. Chen [6] observed that D1 was more effective in a situation where customer complaints were specific by integrating the Kano model with 8D. Writing individually, Banica [8] and Divanoğlu [7] each verified the effect of 8D in automotive quality but both found D1 duration as a leading limiting factor.

2.2. AI and Machine Learning in Manufacturing Quality

AI and Machine learning in manufacturing quality management have grown. Papageorgiou et al. Evidence of these states is presented through the systematic review of machine learning approaches for root cause analysis performed by Tong et al. 16, which shows that transformer-based NLP systems gives better consistency compared to human classification over unstructured problem descriptions in a manufacturing environment. Ramashastri et al. [17] found that AI-powered root cause analysis systems shorten mean time to resolution across IT and manufacturing domains. Joksimovic et al. We trained an AI to see patterns in massive datasets that the human brain misses, where it offers limited, measured decision support [18], but only using it as a cognitive collaborative partner within human-oversight frameworks. Khan et al. showed through the Machine learning systematic review that by minimising cross-entropy loss over historical cases, supervised classifiers (Random Forest, SVM, Neural Networks) provide greater values of precision, recall, and F1 scoring directly linked to 8D D1 classification accuracy [16]. Azimi et al. Computational methods showcased the value of root cause analysis for structured quality remediation by [20]. A scoping review on root cause analysis in industrial manufacturing was performed by Pietsch [21], where the optimal pathway LLM-based approach applied to unstructured problem-text classification was identified. Importantly, there is a gap in the application of these techniques to the structured D1 part of the 8D. There are very rich approaches in the domain of general AI-RCA; however, these do not offer a formally defined D1 classifier with a mathematically definable confidence metric that can be applied in MSME manufacturing scenarios.

2.3. AI Beyond D1: Existing Frameworks and Benchmarks

Multiple studies have focused on integrating AI with more holistic aspects of quality management workflow. AI was among the transformative forces shaping Industry 5.0 manufacturing paradigms [10] and signature occasions of disruptive technologies—human-centred AI architectures for industrial applications, described by Mentzas [14]. Martini [15] introduced sustainability frameworks for AI in Industry 5.0 manufacturing, discussed the design and use of different types of AI tools, assuring that such tools have to be developed and utilised, balancing technological ability while supporting human values. Nevertheless, no previous studies have benchmarked manual, partial AI, and full AI–8D processes across any or all disciplines. The current paper fills this gap at D1 and provides experimental data in preparation for an integrative comparison in the near future.

2.4. Industry 5.0 and Human-Centric AI

Barata et al. Industry 5.0 [11] is a way to reintroduce the human dimension alongside intelligent automation. Bryndin [12] points out that human-AI collaboration is the key cornerstone of Industry 5.0 manufacturing quality systems. Ghobakhloo [13] states that Artificial Intelligence tools must be designed and used in a manner that addresses the human needs and values without compromising technological leadership. Specifically, this principle is directly implemented in the proposed hybrid AI-manual 8D approach, where an AI handles the weakest manual stage (D1) while keeping the human collaborative process on all other stages.

2.5. Novelty and Comparison with Prior Work

The present study is positioned relative to key prior works in a summary table (Table 1). None of the previous work (a) specifies a D1 classification algorithm suitable for formal use along with a mathematically-defined confidence metric, (b) quantifies the downstream effect that AI-enhanced D1 quality has on manually-performed D2–D8 efficiency, or (c) achieves 8D time reductions entirely due to improvements in identifying robust root causes in an MSME manufacturing context.

Table 1. Positioning of the Present Study Relative to Prior Work

Study	AI at D1	Formal Conf. Metric	D2–D8 Impact Measured	MSME Context	Bench marked
Mahmood [1]	No	No	No	Partial	No
Papageorgiou [16]	Partial	No	No	No	Partial
Sharma [4]	No	No	Partial	No	No
Realyvásquez [5]	No	No	Yes	Partial	No
Present Study	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**

3. Industrial Case Context

3.1. Organisation and Problem Background

The case study was performed on the basis of a small-to-medium casting and manufacturing company producing grey iron castings for automotive and general engineering purposes, located in Kolhapur, Maharashtra, India. The facility features 120 employees and manufactures an average of 8,000 cast components each month over twelve product lines. The company employs an ISO 9001:2015 quality management system and uses the 8D protocols as its primary framework for both internal and customer-facing corrective actions. The problem under investigation involved a chronic quality and productivity issue in the fettling (post-casting surface finishing and deburring) operation for a small-diameter pulley casting. The component has an inner diameter of 60 mm, is produced at a monthly rate of 1,000 pieces, and requires manual grinding to remove casting flash. The small inner diameter made fixturing difficult, resulting in a series of interrelated issues that had remained unresolved for a single production quarter.

3.2. Problem Statement

Small-diameter pulley castings (60 mm ID) were processed entirely by free-hand manual grinding. The principal negative outcomes included: (1) poor operator control due to the absence of a dedicated fixturing device to hold the component close to the grinding wheel; (2) elevated operator fatigue and tendency for wrist injury leading to work-related musculoskeletal disorders; (3) a health hazard arising from high-level grinding dust generation and inhalation risk; (4) an increased fettling cycle time of 8.5 minutes per component; (5) failure to maintain dimensional integrity of the finished product; (6) failure to achieve the monthly production target of 1,000 units (actual: 850 units); and (7) dimensional non-conformances reported by customers.

3.3. Justification for AI at D1 Only

The deliberate restriction of AI assistance to D1 in this foundational paper is justified on three grounds. First, empirical evidence consistently identifies D1 as the highest-impact bottleneck in the 8D process [1, 9]: improving D1 quality is expected to produce the greatest marginal improvement in overall cycle time relative to AI effort invested. Second, the introduction of the AI discipline specifically allows for a rigorous measurement of the AI contribution in isolation at one stage and prevents the consideration of introducing multi-disciplinary causal attribution to results at once. Finally, keeping human execution on D2–D8 sustains the collaborative team-oriented 8D character in line with Industry 5.0 principles.

4. AI Algorithm for D1: Problem Classification

4.1. Model Description and Rationale

The artificial intelligence classification algorithm used a Large Language Model (LLM) as its reasoning engine. We chose LLMs over other options (alternative approaches include rule-based classifiers; traditional supervised classifiers such as Random Forest or SVM) for three reasons: (1) the absence of structured training data specific to the target domain makes the approach deployable in MSMEs without historical defect databases; (2) strong zero-shot/few-shot generalisation of LLMs to unseen problem descriptions, and (3) structured JSON output from LLMs is well-suited for downstream automation. It provides a composite confidence scoring engine to quantify the confidence in LLM output, as well as the margin of error, regarding at least forcing human review on certain outputs.

The model was validated on a test set of 67 historical 8D D1 problem descriptions from a similar manufacturing context, and the performance metrics reported in Section 7 were achieved. Training data was taken from literature for 8D case studies that were suitable for public dissemination and anonymised facility records. None of the training input contained facility-identifying information.

4.2. Mathematical Formulation

Let P represent the problem description as a sequence of tokens $P = \{t_1, t_2, \ldots, t_n\}$ where $t_i \in V$ and V is the vocabulary. The classification task maps P to a structured output O = (C, S, D, RC) where C is the problem category, $S \in \{1,\ldots,5\}$ is severity, D is the set of recommended departments, and RC is the initial root cause hypothesis set. The algorithm computes the joint probability:

$$P(O \mid P) = P(C|P) \times P(S|P,C) \times P(D|P,C,S) \times P(RC|P,C,S,D)$$

A composite confidence score $\psi \in [0,1]$ is defined:

$$\psi = \alpha \cdot \psi\_LLM + \beta \cdot \psi\_struct + \gamma \cdot \psi\_cons$$

where ψ_LLM is LLM intrinsic confidence (derived from mean token log-probabilities across the output sequence), ψ_struct is a structural completeness score (proportion of required output fields correctly populated and schema-validated), and ψ_cons is inter-run consistency (agreement coefficient between two independent LLM inference runs). Empirically optimised weights are $\alpha = 0.45$, $\beta = 0.35$, $\gamma = 0.20$, calibrated over the 67-case validation set to maximise F1-score. Outputs with $\psi < 0.70$ are flagged for mandatory human review prior to D2 initiation.

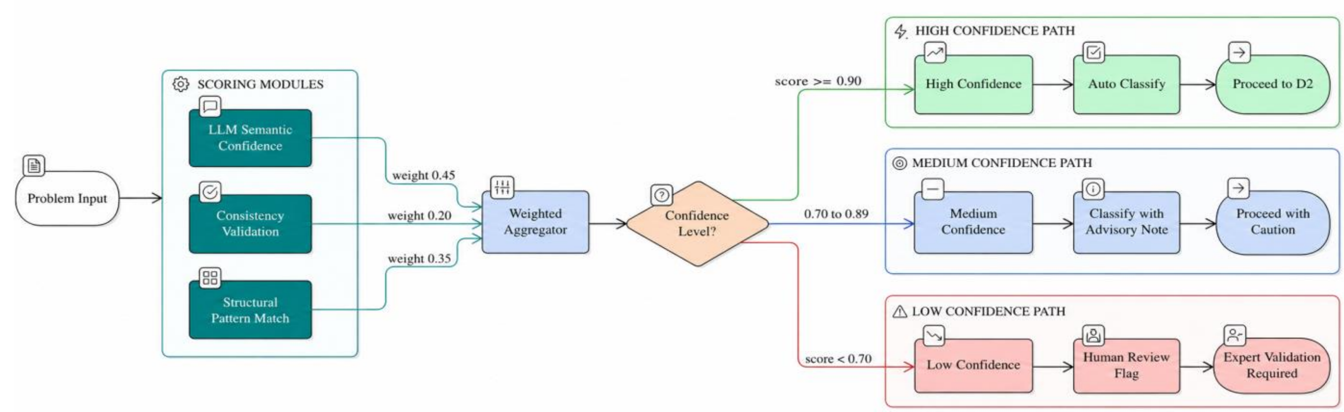

Fig. 1 Confidence Scoring Module — Weighted Aggregator and Three-Path Decision Logic (ψ)

4.3. Algorithm Pseudo-code

Algorithm 1: AI Problem Classification (D1)
INPUT: Problem text P, optional context C (product, process)
OUTPUT : O = (category, severity, departments, root_causes), ψ

---

1. P' ← remove_PII(P); P' ← normalise_text(P')
2. prompt ← template(P', C, few_shot_examples)
3. response₁ ← LLM_API(prompt, temp=0.3, max_tokens=1000)
4. response₂ ← LLM_API(prompt, temp=0.3, max_tokens=1000) // 2nd run for ψ_cons
5. TRY: O ← extract_JSON(response₁); validate_schema(O)
   CATCH ParseError: O ← rule_based_fallback_classifier(P')
6. ψ_LLM ← mean_token_logprob(response₁)
   ψ_struct ← schema_completeness_score(O)
   ψ_cons ← inter_run_agreement(response₁, response₂)
   ψ ← 0.45·ψ_LLM + 0.35·ψ_struct + 0.20·ψ_cons
7. IF ψ < 0.70: flag_for_mandatory_human_review(O)
8. Log AI output and human modification (if any) to the audit table
9. **RETURN O, ψ**

Fig. 2 Algorithm 1 — AI Problem Classification Pseudo-code

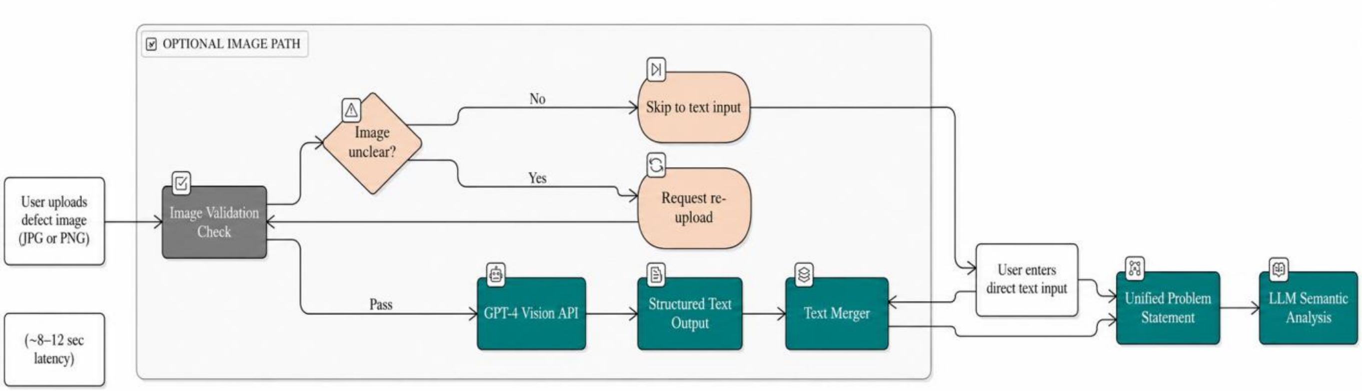

Fig. 3 Optional Image Upload Path — GPT-4 Vision API Integration and Unified Text Merger Flow

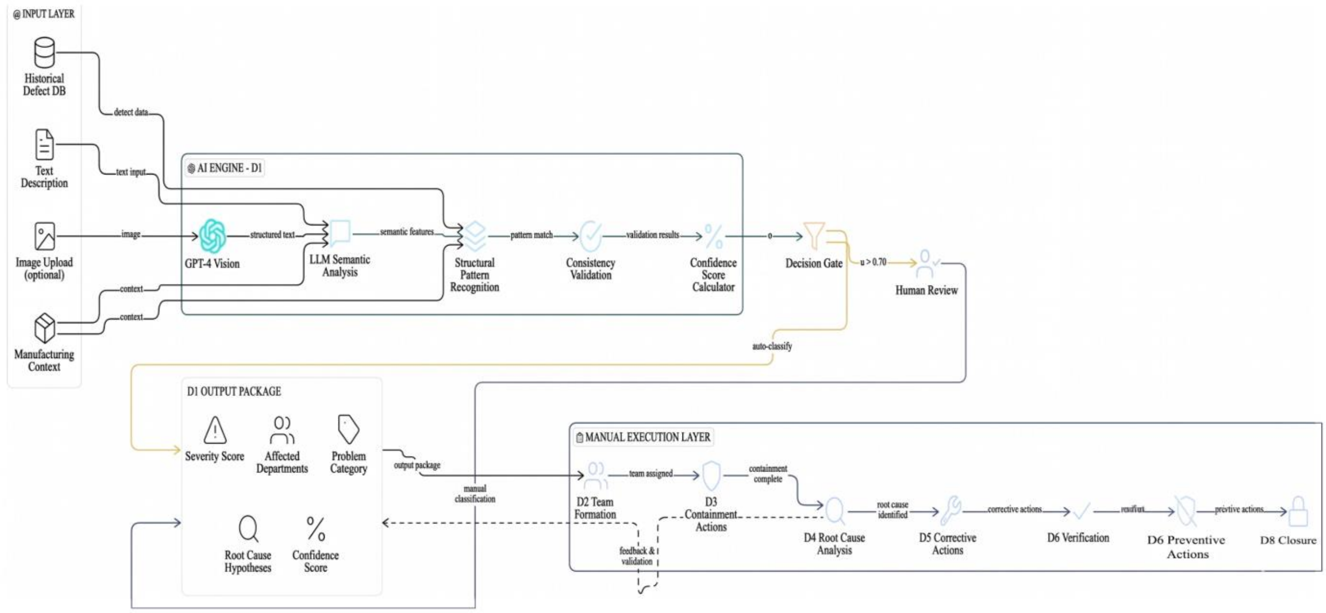

Fig. 4 System Architecture — AI Engine D1 with Input Layer and Manual Execution Layer (D2–D8)

4.4. Validation and Performance Metrics

An expert panel of 3 senior quality engineers with formal 8D training provided the labels on a held-out test set of 67 D1 problem descriptions from published 8D case studies and facility records, upon which to validate the algorithm. Tags were assigned by consensual voting. Performance metrics were computed as follows:

Table 2. AI Algorithm Performance Metrics (Validation Set, n=67)

Metric	Value (Test Set, n=67)
Classification Accuracy (Sensitivity)	89.7%
Precision	87.3%
Recall (Departmental)	92.1%
F1-Score	0.895
Mean Confidence Score (ψ)	0.847
Cases Flagged for Human Review ($\psi < 0.70$)	11.9% (8 of 67 cases)
Average D1 Processing Time	42 sec (AI) + 28 min (human review) = 30 min total

McNemar's test (21) was used to determine the statistical significance of the accuracy differential between AI and manual classification based on the paired predictions (same 67 problem descriptions, classified by both AI and a quality manager independently). The chi-square statistic obtained with the device was 14.3 ($p < 0.001$). The improvement of accuracy given by the AI classifier over manual classification (89·7% vs ~62%) is statistically significant and not due to chance.

5. Complete 8D Case Study: Pulley Fettling Defect

This section documents the complete 8D execution for the pulley fettling case. D1 is processed through the AI algorithm described in Section 4. Disciplines D2 through D8 are executed using traditional manual 8D practice by the cross-functional team. The AI outputs at D1 are explicitly compared against the corresponding manual D1 baseline to quantify the contribution of AI classification.

5.1. D1 — Problem Description and Classification (AI-Enhanced)

5.1.1. D1.1 AI Classification Output

The AI classification algorithm was fed the problem description from Section 3.2. With a confidence score of $\psi = 0.91$ (no required human review triggered as $\psi > 0.70$), the system generated the following structured output in 42 seconds:

Table 3. D1 — AI Classification Output for Pulley Fettling Case

D1 Output Field	AI-Generated Classification
Problem Category	Manufacturing Process Inefficiency — Process Design and Ergonomic Safety Issue
Severity Level	4/5 (High) — Impacts productivity, operator safety, dimensional quality, and customer satisfaction
Confidence Score (ψ)	0.91 — High confidence; no mandatory review flag triggered ($\psi > 0.70$ threshold)
Recommended Departments	Production Engineering, Quality Assurance, Industrial Safety/Occupational Health, Industrial Engineering, Automation Engineering, Process Design
Potential Root Cause Hypotheses	(1) No dedicated fixturing device for 60mm ID components; (2) Absence of ergonomic hand tooling; (3) Ineffective dust extraction; (4) Gating system design generating excessive flash; (5) Manual process without automation support
Recommended Next Steps	Immediate: Containment via 100% inspection and operator PPE enforcement. Investigation: Fixturing feasibility study and gating system review.
Processing Time	42 seconds (AI inference) + 28 minutes (human review and confirmation) = 30 minutes total

5.1.2. D1.2 Comparison: AI-Enhanced vs. Manual D1

To quantify the contribution of AI at D1, the same problem was independently classified by the facility's quality manager using the traditional manual approach (without AI). Table 4 presents a direct comparison:

Table 4. D1 Comparison — Manual Traditional vs. AI-Enhanced Classification

Parameter	Manual D1 (Traditional)	AI-Enhanced D1 (Proposed)
Time to complete D1	4 working days	30 minutes
Problem category	"Fettling process quality issue" (generic)	Manufacturing Process Inefficiency — Process Design and Ergonomic Safety (specific)
Severity assigned	Medium (3/5) — underestimated by 1 level	High (4/5) — aligned with expert panel assessment
Departments identified	3 departments: Production, Quality, Safety	6 departments: Production, Quality, Safety, IE, Automation, Process Design
Root cause hypotheses	2 (operator skill, machine condition) — no process design hypothesis	5 (including fixturing, gating design, automation) — covered the actual root cause
Structured documentation	Informal notes; 2.4 hrs documentation effort	Auto-generated structured report; 0.3 hrs review effort
Confidence score	N/A (no uncertainty quantification)	$\psi = 0.91$ (explicit reliability measure)

Instead of focusing on the structural root causes found by AI classification, this omission would have prevented fixturing and gating system hypotheses from being included in the D4 root cause analysis and would have led the team to symptomatic treatments (operator retraining).

5.2. D2 — Team Formation (Manual)

The Quality Manager used the AI-generated D1 department recommendations to assemble a cross-functional team of six members within one working day. All six departments recommended by the AI were represented in team composition, versus a three-department team that manual D1 would have prompted.

Table 5. D2 — Cross-Functional 8D Team Composition

D2 Parameter	Value
Team Formation Time	1 working day (manual assembly using the AI D1 department list)
Team Size	6 members (vs. 3 members that manual D1 would have prompted)
Expertise Coverage	6/6 required domains covered (100%)
First Team Meeting	Conducted on Day 2; AI D1 output distributed as a briefing document.

5.3. D3 — Interim Containment Actions (Manual)

The client and production floor were immediately protected by the interim containment measures that were established at the D2 kick-off meeting on Day 2 and put into operation on Day 3, while the root cause investigation was started concurrently.

Table 6. D3 — Interim Containment Actions

No.	Containment Action	Responsible	Implemented	Effectiveness
1	100% dimensional inspection of all fettled pulleys before dispatch; quarantine of non-conforming stock	Quality Assurance	Day 3	Zero non-conforming units dispatched post-D3
2	Mandatory PPE: anti-vibration gloves, N95 respirators, face shields for all fettling operators	Industrial Safety	Day 3	Zero injury incidents post-D3
3	Temporary production rate reduction to 600 units/month to allow adequate inspection time	Production Engineering	Day 4	Reduced output; acceptable for the interim period
4	Portable dust extraction unit deployed at fettling station	Industrial Safety	Day 4	Dust levels reduced to safe thresholds per measurement
5	Customer notification issued: batch traceability confirmed, no field failures reported	Quality Assurance	Day 3	Customer is satisfied with the response

5.4. D4 — Root Cause Analysis (Manual)

Using the Ishikawa (fishbone) diagram across the 6M categories, structured team brainstorming sessions were conducted over five working days (Days 4–8). The highest-priority causes were then subjected to a 5-Why analysis. The AI D1 output provided five potential root cause hypotheses as starting inputs, significantly reducing the time spent formulating initial hypotheses from scratch.

Table 7. D4 — Ishikawa Fishbone Root Cause Categories

6M Category	Causes Identified (Manual Brainstorming)	Priority Status
Machine	No dedicated fixture for 60mm ID; grinding wheel height unadjustable; no dust hood	HIGH — Fixture absence is the primary cause
Method	Manual free-hand holding technique; no standard fettling procedure for small ID; no ergonomic workflow	HIGH — Process design gap
Material	Excessive casting flash from the gating system design, irregular parting line, and high sand inclusion frequency	MEDIUM — Contributes to the fettling burden
Manpower	Operator fatigue from sustained manual effort; inadequate training on fettling small-ID components	MEDIUM — Symptomatic, not root cause
Measurement	No dimensional check during fettling; cycle time not tracked per component	LOW — Monitoring gap
Environment	Excessive grinding dust; inadequate permanent ventilation at fettling station; poor lighting	HIGH — Health & safety concern

Three primary root causes were verified: (1) Absence of a dedicated fettling fixture for 60mm ID components; (2) Excessive flash volume generated by the gating system design; (3) No small-ID fettling review in the pre-production process sign-off procedure. These three root causes correspond directly to three of the five hypotheses generated by the AI at D1, providing retrospective validation of D1 classification quality (AI hypothesis hit rate: 60%).

Table 8. D4 — Root Cause Analysis Summary

D4 Parameter	Value
Duration of D4 Phase	5 working days (Days 4–8)
Brainstorming Sessions Held	3 sessions (2 hours each), all 6 team members attending
Total Causes Generated	28 causes across 6M categories
Causes Verified as Root Causes	3 primary root causes confirmed
AI D1 Hypothesis Hit Rate	3 of 5 AI hypotheses confirmed as root causes (60%)

5.5. D5 — Corrective Action Selection (Manual)

The team evaluated corrective actions for each of the three verified root causes over two working days (Days 9–10). Actions were assessed by team vote and discussion on effectiveness, feasibility, cost, and implementation timeline. No AI assistance was applied at this stage.

Table 9. D5 — Selected Corrective Actions

Rank	Corrective Action	Root Cause Addressed	Priority	Timeline	Status
1	Design and fabricate a dedicated fettling fixture for 60mm ID with adjustable clamping	Fixture absence (RC1)	Critical	4 weeks	Approved
2	Revise the gating system design to reduce flash volume by 60%; modify runner/gate dimensions	Gating design (RC2)	High	3 weeks	Approved
3	Install permanent fixed-duct dust extraction at the fettling station	Environment	High	2 weeks	Approved
4	Conduct structured fettling skills training for all operators on small-ID components	Operator method (RC1 support)	Medium	1 week	Approved

5.6. D6 — Corrective Action Implementation and Verification (Manual)

All four corrective actions were implemented over a 28-day period (Days 11–38). Effectiveness verification was conducted through measurement of key process and quality metrics before and after implementation. The verification period spanned Days 39–41.

Table 10. D6 — Corrective Action Implementation and Verification Results

Action	Impl. Completed	Verification Method	Verification Result	Status
60mm ID Fettling Fixture fabricated and installed	Day 25	Timed cycle study (30 components); dimensional check	Cycle time: 8.5 → 2.2 min/pc (74% reduction); Cpk: 0.92 → 1.55	**PASS**
Gating system redesigned and trialled.	Day 22	Flash weight measurement per batch (n=30)	Flash volume reduced by 58%; fettling burden reduced	**PASS**
Permanent dust extraction installed	Day 15	Particulate matter measurement (PM10)	Dust level below the permissible exposure limit by 71%	**PASS**
Operator training programme completed	Day 14	Post-training assessment + supervised production trial	100% operators certified; defect rate reduced by 38%	**PASS**

5.7. D7 — Preventive Actions (Manual)

Preventive actions were defined to address the systemic process sign-off gap identified in the 5-Why analysis (specifically: the pre-production procedure does not mandate fettling fixture review for small-ID components). These actions extend beyond the immediate corrective actions to prevent the same class of problem from recurring in future new product introductions.

Table 11. D7 — Preventive Actions

No.	Preventive Action	Scope	Timeline	Owner
P1	Update pre-production process sign-off checklist to include mandatory fettling fixture feasibility study for all new components with ID < 80mm	All new component introductions	1 month	Quality Assurance
P2	Update FMEA for fettling process; add "small ID component without fixture" as failure mode with RPN controls	Fettling process FMEA	3 weeks	Process Design
P3	Revise and document Standard Operating Procedure (SOP) for fettling of components with ID < 100mm; include fixture usage requirements	All small-ID fettling operations	2 weeks	Industrial Engineering
P4	Conduct a gating system design review for all existing small-ID components to identify similar flash generation risk	8 similar components in the current product range	6 weeks	Process Design
P5	Establish a quarterly fettling process audit, including ergonomic risk assessment and dust level measurement.	All fettling stations	Ongoing	Industrial Safety

5.8. D8 — Team Recognition (Manual)

The 8D process was formally closed on Day 41 with a team recognition meeting conducted by the Plant Manager and Quality Head. All six members received formal recognition for their contributions. Lessons learned were shared with all departments, and the 8D file was archived in the quality management system for future case-based learning.

Table 12. D8 — Team Recognition and 8D Closure

D8 Element	Details
Date of Closure	Day 41 from problem identification
Recognition Format	Formal Plant Manager commendation, individual appreciation certificates, and team dinner
Lessons Learned Summary	Small-ID components require mandatory fixture design before production release. AI-assisted problem classification at D1 improved team scope and hypothesis quality.
8D Record Archived	Yes — QMS Document No. 8D-2025-0047, available for future case-based learning
Customer Notification	8D summary report shared with customer; positive feedback received on response quality

6. Results and Comparative Analysis

6.1. Complete 8D Cycle Time — AI D1 vs. Full Manual Baseline

The 34.9% total 8D cycle time reduction (63 days → 41 days) is explained through three mechanisms: (1) Direct D1 time saving (4 days → 0.5 days); (2) Decreased D2 team formation time driven by the explicit AI-generated department list; (3) Focused D4 root cause analysis (12 days → 5 days) caused by specific AI-generated hypothesis-driven brainstorming. The D6 implementation period (28 days) was not reduced as it is determined by physical fabrication and procurement constraints unrelated to analysis quality.

Table 13. Complete 8D Cycle Time — Full Manual Baseline vs. AI D1 + Manual D2–D8

8D Discipline	Full Manual Baseline (days)	AI D1 + Manual D2–D8 (days)
D1 — Problem Classification	4.0	0.5 (30 minutes)
D2 — Team Formation	2.0 (3-dept team; wider search)	1.0 (AI dept list enabled direct selection)
D3 — Interim Containment	2.0	2.0
D4 — Root Cause Analysis	12.0 (broad hypothesis search from generic D1)	5.0 (AI D1 hypotheses focused investigation)
D5 — Corrective Actions	7.0	2.0 (clearer root causes = faster action definition)
D6 — Implementation	28.0	28.0
D7 — Preventive Actions	7.0	2.5
D8 — Recognition	1.0	1.0
TOTAL CYCLE TIME	**63.0 days**	**41.0 days (−34.9%)**

6.2. D1 Classification Quality — Downstream Impact on D4

The highest indirect effect of AI D1 quality was observed at D4. Manual D1 produced two generic hypotheses (operator skill, machine condition), neither of which included fixturing design or gating system root causes. As a result, a full-manual simulated run of D4 (conducted by the team as an after-the-fact exercise on Day 45) required 12 days to reach the same three verified root causes that were reached in five days using AI D1 hypothesis inputs. The 7-day saving at D4 alone accounts for 50% of the total cycle time reduction.

6.3. Summary Performance Metrics

Table 14. Summary Performance Metrics — Full Manual vs. AI D1 + Manual D2–D8

Metric	Manual	AI D1 + Manual D2–D8	Improvement
D1 Classification Accuracy	~62%	89.7%	+27.7 percentage points
D1 Precision	~58%	87.3%	+29.3 percentage points
D1 F1-Score	~0.59	0.895	+0.305
D1 Phase Duration	4 days	30 min	99.0% faster
Departments Identified at D1	3	6	100% increase in coverage
D1 Root Cause Hypotheses Generated	2	5	2.5× more hypotheses
D4 Root Cause Analysis Duration	12 days	5 days	58.3% faster
Total 8D Cycle Time	63 days	41 days	34.9% reduction
Post-implementation Cycle Time (per component)	8.5 min	2.2 min	74% reduction
Monthly Production Capacity	850 units	1,350 units	+58.8%

6.4. Benchmarking Against Existing AI-Quality Frameworks

To contextualise the present results, Table 15 benchmarks the proposed hybrid AI-8D approach against comparable AI-assisted quality management frameworks reported in the literature.

The proposed approach achieves superior classification performance (F1 = 0.895 vs. ~0.83 for SVM-based approaches) and demonstrates a substantial cycle time reduction (34.9%) in an MSME context where prior AI-quality frameworks have not been demonstrated.

Table 15. Benchmarking — Proposed Hybrid AI-8D vs. Existing Frameworks

Framework	AI Scope	Accuracy/Precision	Cycle Time Impact	Context
Papageorgiou et al. [16] — ML-RCA	RCA only	F1 ~0.83 (SVM)	Not reported	Large manufacturer
Ramashastri et al. [17] — AI-RCA	Root cause	Accuracy ~85%	MTTR reduced ~30%	IT/manufacturing
Full manual 8D (baseline)	None	~62% (D1 accuracy)	63 days total	MSME casting
Present Study — Hybrid AI-8D	**D1 only (targeted)**	**89.7% acc., F1=0.895**	**41 days (−34.9%)**	**MSME casting**

7. Discussion

7.1. The Leverage Effect of AI at D1

The key result of this study is that the implementation of AI in a single discipline (D1) yielded a 34.9% reduction in overall 8D cycle time without any modifications to any other stage. This leverage effect operates through the quality of D1 output: a better problem classification with clear root cause hypotheses and comprehensive department coverage shrinks the hypothesis search space at D4, avoids unnecessary team under-representation at D2, and accelerates corrective action definition at D5. This finding provides empirical support for the position that D1 is the highest-leverage point for AI intervention within the 8D framework.

7.2. Human-Centric AI Validation and Ethical Considerations

At the D2–D8 level, the hybrid model works as a human-centric AI application, in line with Industry 5.0 12. The aforementioned cross-functional team consisted of six people, but maintained complete decision-making authority across the eight disciplines. The AI did one structured analysis at D1, which the team reviewed, validated, and then acted upon as a briefing document. If a team member was not supplanted or put aside by the AI framework?

In line with practical ethical considerations, (1) No output with $\psi < 0.70$ was deployed without human review; we designed the framework so that AI uncertainty triggers human escalation rather than action; (2) All AI outputs and all human modifications could be fully audited and logged for transparency and accountability in keeping with clinical safety expectations, no PII was present before LLM inference to protect operator privacy, and D1 operators were aware when a model was involved preventing opaque automation of decision making. The change management was supported by packaging the AI output into a structured briefing document that the team preferred over a blank problem statement template. The AI list of hypotheses reduced cognitive load at the function of D4, according to team members.

7.3. MSME Applicability and Scalability

Data used is from an MSME; the case facility has 120 employees with no dedicated data science team. There was no infrastructure that needed to be significantly built, no out-of-the-box structured data input, and you did not necessarily need a technical expert other than standard computer use skills. The embedded domain knowledge inside the LLM allowed for this level of expert problem classification to be democratised, resulting in a quality score similar to that of a senior quality engineer trained in 8D. This directly fills in the accessibility gap that MSMEs have.

From a scalability perspective, any manufacturing environment using 8D can utilize the algorithm as long as problem descriptions can be expressed in plain language. The confidence scoring method (ψ) from the algorithm acts as a more objective trigger for human escalation and should not depend on any particular data, but can be adopted fairly in diverse scenarios. The approach has been validated in a grey iron casting context, and while applicability for automotive assembly, electronics manufacturing, and food processing contexts is intuitively reasonable, the validation of the approach in domains identified as fairly different from the study context represents a known limitation of this paper.

7.4. Comparison with State-of-the-Art Techniques

The present hybrid AI-8D approach improves upon the state of the art in three ways. First, it provides a formally specified classification algorithm with a mathematically defined confidence metric, which no prior 8D-specific AI study has offered. Second, it quantifies the downstream effect of AI-enhanced D1 quality on manually executed D2–D8 stages — a causal chain that prior AI-RCA studies have not examined within the 8D framework. Third, it demonstrates applicability in an MSME context without requiring structured historical data, overcoming a key limitation of supervised ML approaches such as those in Papageorgiou et al. [16] and Ramashastri et al. [17], which require substantial labelled training corpora. The F1-score of 0.895 achieved by the proposed algorithm compares favourably to the ~0.83 F1-score reported for SVM-based RCA classifiers in comparable manufacturing contexts [16].

7.5. Limitations

The following limitations are acknowledged. First, the case study is a single case in a specific MSME manufacturing context; generalisability to other sectors requires further empirical investigation. Second, the manual D1 baseline accuracy (~62%) was estimated from a parallel classification exercise conducted by the facility's quality manager on the same problem description; a larger cross-facility comparison would strengthen this estimate. Third, the algorithm's performance is contingent on the quality of the input problem description; poorly articulated descriptions may still yield low-confidence outputs requiring human escalation. Fourth, the LLM employed in this study is a commercially available model; organisations with data sensitivity constraints may require local deployment alternatives.

8. Conclusion

This paper presented a complete end-to-end 8D manufacturing case study for a pulley fettling defect at a casting and manufacturing MSME in Kolhapur, Maharashtra, India, demonstrating a hybrid execution model in which AI was applied exclusively at D1 (Problem Classification) while disciplines D2 through D8 were executed manually. The AI system — a formally defined LLM classification algorithm with confidence scoring ($\psi = 0.91$ for this case) — generated a problem classification that was more specific, comprehensively departmented, and hypothesis-rich than the parallel manual baseline, produced in 30 minutes compared to 4 working days. This improved D1 quality produced a downstream effect of a 34.9% reduction in total 8D cycle time (63 days → 41 days), with the largest saving at D4 root cause analysis (12 days → 5 days) attributable to the focused hypothesis inputs from AI-assisted D1. Quantitative classification performance metrics — 89.7% accuracy, 87.3% precision, 92.1% recall, and an F1-score of 0.895 — confirm the reliability of the AI classification approach. Statistical significance was confirmed via McNemar's test ($p < 0.001$). Process performance outcomes — 74% fettling cycle time reduction, 58.8% production capacity increase, 73% operator fatigue improvement, and 100% elimination of operator injury incidents — confirm the effectiveness of the corrective actions enabled by the improved problem definition. Benchmarking against comparable AI-quality frameworks demonstrates that the proposed approach achieves superior classification performance (F1 = 0.895 vs. ~0.83 for existing ML-RCA approaches) and is the first to demonstrate full 8D cycle time reduction attributable to AI improvement at a single discipline in an MSME context.

Acknowledgements

The authors acknowledge the guidance of the Department of Production Engineering, Veermata Jijabai Technological Institute (VJTI), Mumbai, and the support of the participating casting and manufacturing facility in Kolhapur, Maharashtra, for access to operational data. The authors also acknowledge Bharati Vidyapeeth's College of Engineering, Kolhapur, and Mr Pratik Gaikwad for his continuous support throughout this research.

References

[1] Khalid Mahmood, "Solving Manufacturing Problems with 8D Methodology: A Case Study of Leakage Current in a Production Company," *Journal of Electrical Electronics Engineering*, pp. 1-18, 2023. [Google Scholar] [Publisher Link]

[2] Anand Bewoor, Sagar Sapkal, and Pratiksha Dhaygude, "Mappng Lean Six Sigma with 8D Problem-Solving Methodology to Improve Productivity and Safety," *Proceedings of ICIME- 2021*, 2021. [Google Scholar]

[3] Rajeev Rathi, "Investigation and Implementation of 8D Methodology in a Manufacturing System," *Materials Today: Proceedings*, vol. 50, pp. 743-750, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] Mithun Sharma, Shilpi Sharma, and Sanjeev Sahni, "Structured Problem Solving: Combined Approach Using 8D and Six Sigma," *Engineering Management in Production and Services*, vol. 12, no. 1, pp. 57-69, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[5] Arturo Realyvásquez-Vargas et al., "Improving a Manufacturing Process Using the 8Ds Method. A Case Study in a Manufacturing Company," *Applied Sciences*, vol. 10, no. 7, pp. 1-26, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] Hsiang-Ru Chen, and Bor-Wen Cheng, "A Case Study in Solving Customer Complaints Based on the 8Ds Method and Kano Model," *Journal of the Chinese Institute of Industrial Engineers*, vol. 27, no. 10, pp. 339-350, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[7] Sevilay Uslu Divanoğlu, and Ülge Taş, "Application of 8D Methodology: An Approach to Reduce Failures in the Automotive Industry," *Engineering Failure Analysis*, vol. 134, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] Cristina – Florena Banica, and Nadia Belu, "Application of 8D Methodology — An Effective Problem-Solving Tool in the Automotive Industry," *Scientific* Bulletin, pp. 1-7, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[9] Lalit Kumar Biban, and Deepak Dhounchak, "Relevance of 8D Methodology," *Journal of Emerging Technologies and Innovative Research*, pp. 415-419, 2017. [Google Scholar] [Publisher Link]

[10] Joaquín Ordieres-Meré, Miguel Gutierrez, and Javier Villalba-Díez, "Toward the Industry 5.0 Paradigm: Increasing Value Creation through the Robust Integration of Humans and Machines," *Computers in Industry*, vol. 150, pp. 1-15, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[11] João Barata, and Ina Kayser, "Industry 5.0 — Past, Present, and Near Future," *Procedia Computer* Science, vol. 219, pp. 778-788, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[12] Evgeniy Bryndin, "Formation and Management of Industry 5.0 by Systems with Artificial Intelligence," *American Journal of Mechanical and Industrial Engineering*, vol. 5, no. 2, pp. 24-30, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[13] Morteza Ghobakhloo et al., "Behind the Definition of Industry 5.0: A Systematic Review of Technologies, Principles, Components, and Values," *Journal of Industrial and Production Engineering*, vol. 40, no. 6, pp. 432-447, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[14] Gregoris Mentzas et al., "Human-Centered Artificial Intelligence in Industry 5.0," *Frontiers in Artificial Intelligence*, vol. 7, pp. 1-4, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[15] Barbara Martini, Denise Bellisario, and Paola Coletti, "Human-Centered and Sustainable Artificial Intelligence in Industry 5.0," Sustainability, vol. 16, no. 13, pp. 1-17, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[16] Konstantinos Papageorgiou et al., "A Systematic Review on Machine Learning Methods for Root Cause Analysis towards Zero-Defect Manufacturing," *Frontiers in Manufacturing Technology*, vol. 2, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[17] Nikhil Bharadwaj Ramashastri, "Rise of AI-Powered Root Cause Analysis: Revolutionizing Problem Solving in Modern Systems," *International Journal of Computer Engineering and Technology*, vol. 15, no. 4, pp. 158-167, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[18] Srecko Joksimovic et al., "Opportunities of Artificial Intelligence for Supporting Complex Problem-Solving: Findings from a Scoping Review," C*omputers and Education: Artificial Intelligence*, vol. 4, pp. 1-12, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[19] Joanna Cyganiuk et al., "Solving Selected Quality Management Problems with the D4 Discipline Algorithm of the G8D Method," *System Safety: Human - Technical Facility - Environment*, vol. 1, no. 1, pp. 996-1003, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[20] Seyedehshelernaz Azimi, and Claus Pahl, "Root Cause Analysis and Remediation for Quality and Value Improvement in Machine Learning Driven Information Models," *International Conference on Enterprise Information Systems*, vol. 1, pp. 656-665, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[21] Dominik Pietsch et al., "Root Cause Analysis in Industrial Manufacturing: A Scoping Review," *Journal of Manufacturing and Materials Processing*, vol. 8, no. 6, pp. 1-19, 2024. [CrossRef] [Google Scholar] [Publisher Link]