

Enhancing Rail Madad AI-Powered Complaint Management System

Mr. Kalpesh K. Manure¹, Asst. Prof. Chetana M. Kawale²

Research Scholar, Department of Computer Applications, Shramsadhana College of Engineering & Technology

Jalgaon.MH, India¹

Asst. Professor, Department of Computer Applications, Shramsadhana College of Engineering & Technology

Jalgaon.MH, India²

Abstract: Rail Madad serves as the unified digital grievance redressal mechanism of Indian Railways, facilitating complaint registration across train and station operations, ticket tracking, emergency assistance, and citizen feedback. While the platform effectively consolidates multi-channel intake into a centralized monitoring portal, first-level triage remains constrained by the ambiguities of free-form descriptions, code-mixed vernacular expressions (Hinglish), contradictory metadata, and manual department assignment. This research presents an artificial intelligence-enhanced complaint management framework for Rail Madad that integrates multilingual natural language processing, automated entity extraction, urgency and sentiment assessment, image-evidence verification, confidence-aware routing, and human-in-the-loop decision controls. The architecture adopts an incremental lifecycle model, maintaining the existing institutional complaint lifecycle while introducing AI assistance across intake, triage, prioritization, departmental routing, and continuous feedback learning. Empirical baseline analysis of official Railway Board operational performance demonstrates substantial workload escalation, rising from 1,61,575 complaints in July 2022 to 2,35,394 in July 2023, with average disposal time increasing from 1 hour 36 minutes to 3 hours 1 minute. Facilities on-board (50%), security (20%), and punctuality (15%) constitute the dominant complaint categories. The proposed framework implements a confidence-calibrated routing engine governed by institutional rules and statutory compliance with the Digital Personal Data Protection (DPDP) Act 2023. By routing low-confidence or high-impact grievances to human officers while automating routine high-confidence triage, the proposed architecture establishes a transparent, auditable, and resilient path for modernizing railway grievance redressal without compromising human accountability.

Keywords: Rail Madad, Artificial Intelligence, Complaint Management, Natural Language Processing, Complaint Classification, Sentiment Analysis, Intelligent Routing, Human-in-the-Loop, Indian Railways, Grievance Redressal.

INTRODUCTION

1.1 Backgrounds

Efficient grievance redressal is a fundamental requirement for large-scale public transportation infrastructures, directly shaping citizen trust, operational accountability, and service quality. Indian Railways, operating one of the largest passenger transportation networks globally, carries tens of millions of passengers daily across thousands of stations and running train services. To unify previously fragmented passenger touchpoints, the Ministry of Railways established Rail Madad as an integrated digital grievance redressal mechanism. Rail Madad converges multiple communication channels—including a dedicated web portal, native Android and iOS mobile applications, the toll-free 139 helpline, SMS gateways, and official social media handles—into a centralized administrative platform.

The current public intake interface captures structured operational fields, including Passenger Name Record (PNR) numbers, train numbers, origin and destination stations, boarding points, coach and berth positions, contact details, incident dates, hierarchical complaint categories, and supporting photographic attachments. Despite this structured foundation, the primary operational bottleneck occurs during first-level decision-making. Passengers frequently articulate urgent grievances through brief, unstructured descriptions in colloquial Hindi, English, or Romanized code-mixed dialects (Hinglish). Consequently, administrative personnel must manually interpret ambiguous descriptions, verify conflicting journey entities, detect underlying safety risks, and assign tickets to corresponding field departments, creating operational delays.

1.2 Importance

The modernization of passenger grievance administration using artificial intelligence offers vital institutional benefits across governance, operational efficiency, and passenger safety, synthesized into three primary dimensions:

1. **Citizen-Centric Governance and Public Trust:** Fast, equitable, and transparent grievance resolution strengthens public confidence in railway administration, ensuring that passenger grievances receive immediate, auditable operational attention while eliminating administrative opacity.

2.

2. Workforce Optimization and Operational Scalability: Automating repetitive initial categorization relieves front-line staff from manual sorting of hundreds of thousands of monthly submissions, allowing personnel to concentrate on physical on-site rectification while enabling the platform to process tens of thousands of simultaneous complaints during adverse weather, technical service disruptions, or festive peak seasons without queue exhaustion.

3. Proactive Risk Screening and Institutional Auditability: Real-time semantic risk screening flags critical security, medical, and fire emergencies instantly, bypassing regular non-urgent queues for rapid emergency dispatch, while timestamped confidence scores, explainable prediction rationales, and structured human-override logging ensure full administrative compliance and institutional traceability.

1.3 Rail Madad Architecture and Operational Scale

The operational scale of Indian Railways necessitates robust administrative mechanisms. According to official Railway Board performance evaluations, Rail Madad handled 2,35,394 grievances in July 2023, compared to 2,32,258 in June 2023 and 1,61,575 in July 2022. Concurrently, the reported average disposal time escalated from 1 hour 36 minutes in July 2022 to 2 hours 41 minutes in June 2023, reaching 3 hours 1 minute in July 2023. Analysis of train-specific complaint groupings reveals that facilities on-board represented 50% of total submissions, security issues accounted for 20%, punctuality delays comprised 15%, and miscellaneous passenger amenities encompassed the remaining 15%. Within facilities on-board, diverse subcategories exist—including coach cleanliness, electrical equipment failures, linen shortages, water supply disruption, and mechanical maintenance—each requiring routing to entirely distinct technical divisions.

1.4 Research Problems

Current Rail Madad operational workflows face distinct technical and administrative shortcomings that can be categorized into three fundamental challenges:

1. Linguistic Diversity and Narrative Ambiguity: Passengers submit grievances combining Hindi vocabulary with Roman script (e.g., 'coach B2 me toilet clean nahi hai aur water nahi aa raha'), confounding conventional keyword-based and unilingual classification tools while presenting short, vague, or multi-issue descriptions that obscure the primary fault.

2. Severe Category Imbalance and Journey Entity Inconsistencies: Routine cleanliness and amenities complaints heavily outnumber rare yet life-critical emergency grievances (medical emergencies, harassment, fire hazards), causing standard machine learning algorithms to misclassify urgent minority cases, while frequent discrepancies between passenger-entered form fields and free-form narrative descriptions lead to misdirected field inspections.

3. Multimodal Evidence Gaps and the Perils of Uncalibrated Automation: Passengers routinely upload photographic attachments of defective equipment or unsanitary conditions without clarifying textual context, which conventional systems fail to correlate with field units, while deploying unconstrained black-box deep learning models risks silent misrouting of consequential safety tickets, underscoring the absolute necessity of confidence calibration and human oversight.

1.5 Scopes of the System

The scope of this research is to design, model, and analyze an AI-powered complaint management architecture specifically tailored to Rail Madad. The proposed framework does not replace the established database infrastructure or field resolution personnel of Indian Railways. Instead, it introduces an intelligent intermediate layer that normalizes code-mixed passenger text, extracts operational journey entities, detects urgency and emotional sentiment, correlates visual evidence, executes rule-guided confidence-aware routing, and captures officer corrections for continuous model enhancement.

1.6 Objectives

The core objectives of the study are structured across three interconnected development pillars:

1. Multimodal AI Triage and Contextual Journey Entity Binding: To construct an AI architecture that simultaneously processes unstructured multilingual text, journey metadata, and attached images for accurate category prediction while automatically validating passenger-entered parameters against PRS ticket records.

2. Confidence-Calibrated Routing and Automated Escalation: To establish a dual-path routing engine that automatically dispatches routine high-confidence complaints while instantly escalating low-confidence cases, conflicting journey entities, and critical safety or medical emergencies to human supervisory officers.

3. Statutory Privacy Compliance, Human Supervision, and Empirical Benchmarking: To enforce statutory data privacy, minimization, and role-based access governance in strict alignment with the Digital Personal Data Protection (DPDP) Act 2023, while designing an interactive supervisory console that captures officer overrides for continuous active learning and evaluating performance across standardized macro-F1, latency, and calibration metrics.

1.7 Need of the System

A modernized, AI-enhanced complaint management system is vital for Indian Railways because manual triage cannot scale sustainably with surging passenger volumes. With monthly complaint receipts surpassing 2.35 lakh cases, manual sorting introduces cognitive fatigue, inconsistent categorization, and prolonged disposal delays. Immediate, automated identification of safety threats and seamless departmental dispatch are indispensable for safeguarding passenger welfare. An intelligent, confidence-calibrated platform ensures equitable public service delivery while providing senior railway administrators with real-time operational analytics for systemic infrastructure maintenance.

1.8 Selection Of Life Cycle Model For Development

The development of this system is best suited to the incremental model. Public grievance systems require high levels of security, accuracy, and operational continuity, and the incremental model makes it possible to develop and test modules like text normalization, classification, entity binding, and routing gradually across three progressive stages:

Stage 1 - Intelligent Recommendation and Preprocessing: The AI engine operates entirely in advisory mode, generating non-binding categorization and routing recommendations displayed alongside human officer review queues, while establishing linguistic normalization for Hindi, English, and Romanized code-mixed text.

Stage 2 - Multimodal Evidence Integration and Selective Automation: Urgency scoring, sentiment estimation, and computer vision-based image verification modules are introduced to enrich triage context, enabling automated dispatch for high-confidence routine categories while keeping critical safety categories under strict officer control.

Stage 3 - Continuous Feedback Learning and Governance Auditing: Verified officer corrections and passenger resolution feedback are systematically harvested for scheduled supervised model updates and domain vocabulary enrichment, supported by tamper-evident audit logging and statutory privacy protection.

LITERATURE REVIEW / RELATED WORK

There are many different automated classification and public complaint processing systems available today, each with unique architectural advantages and operational constraints. Inadequate handling of code-mixed language, lack of confidence calibration, and absence of human oversight are the most critical challenges identified in legacy systems. The following structured literature review synthesizes key foundational research, adhering to active subject citation framing across ten foundational contributions:

Hake et al., in (Springer 2021) [1], studied that automated complaint handling in industrial processes can effectively suggest error codes from textual descriptions, highlighting the essential requirement of human supervision for operational safety.

Law et al., in (Springer Nature 2026) [2], proposed that enriching structured datasets with reasoning-based embeddings and few-shot learning substantially mitigates severe class imbalance in public grievance classification.

Devlin et al., in (ACL 2019) [3], explored that bidirectional contextual language representations through transformer architectures enable deep contextual comprehension beyond conventional keyword-matching mechanisms.

Viellieber et al., in (arXiv 2020) [4], analyzed that pre-trained language models functioning as domain knowledge bases can effectively extract technical fault signatures from unstructured customer complaint narratives.

Valmianski et al., in (arXiv 2019) [5], observed that while large transformer models excel at semantic representation, lightweight statistical baselines like TF-IDF remain competitive and robust against spelling distortions in noisy user-submitted text.

Labat et al., in (Springer 2024) [6], described that tracking customer emotion trajectories during service interactions provides valuable auxiliary indicators of citizen dissatisfaction and escalation risk.

Dabre et al., in (ACL 2022) [7], highlighted that multilingual pre-trained sequence-to-sequence architectures like IndicBART enable effective representation learning across diverse Indic languages and code-mixed vernaculars.

Tabassi et al., in (NIST 2023) [8], emphasized that operational AI deployments in public administration must incorporate trustworthy, explainable, and risk-aware governance frameworks across their entire lifecycle.

Ministry of Electronics and Information Technology, in (MeitY 2023) [9], confirmed that statutory provisions under the Digital Personal Data Protection Act (DPDP) 2023 mandate strict data minimization and purpose limitation for citizen-facing digital platforms.

Sultan et al., in (IEEE 2022) [10], demonstrated that automated real-time recognition systems achieve high operational efficiency when deployed with feature extraction pipelines under variable operational conditions.

METHODOLOGY

To establish an effective, scalable, and secure AI-assisted complaint handling framework for Rail Madad, this study adopts a design-oriented and empirical research methodology combining system decomposition, workflow modeling, and evaluation benchmarking. The key components of the methodology are detailed below:

3.1 Feasibility Study

The feasibility study evaluates the practical viability of deploying an AI decision-support layer within the existing Rail Madad ecosystem across three analytical domains:

- 1. Technical and Architectural Viability:** Transformer architectures (mBERT, IndicBART) combined with lightweight convolutional backbones provide robust natural language understanding and visual feature extraction. Containerized microservices on high-availability server clusters ensure seamless integration with CRIS back-end APIs.
- 2. Operational and Organizational Viability:** The platform preserves the existing Rail Madad experience for passengers and railway officers. By functioning as an intelligent decision-support assistant rather than an autonomous replacement, the system integrates smoothly into established departmental escalation trees.
- 3. Economic, Legal, and Regulatory Viability:** Initial capital investment in GPU compute is offset by accelerated resolution cycles, reduced ticket reassignments, and minimized manual data re-entry. Furthermore, the architecture complies strictly with the Digital Personal Data Protection Act 2023 [9] via cryptographic hashing and role-based masking of passenger PII.

3.2 Risk Analysis

The deployment of artificial intelligence in public railway administration carries distinct operational, security, and technical risks, managed across three critical categories:

- 1. Technical, Security, and Cyber-Resilience Risks:** Front-end channels are protected against denial-of-service, ticket spamming, and prompt injection via API rate limiting, CAPTCHA, and input sanitization, while computational bottlenecks during volume spikes are resolved using asynchronous Redis/Celery task queues.
- 2. Regulatory, Privacy, and Data Governance Risks:** Inadvertent exposure of passenger PII during model retraining or officer review is prevented through mandatory cryptographic hashing, role-based access control (RBAC), and synthetic test data generation.
- 3. Model Reliability, Out-of-Vocabulary Drift, and Human Adoption Risks:** High-confidence false negatives and out-of-vocabulary railway slang are controlled via mandatory human verification thresholds for all safety categories, while field officer adoption is fostered through transparent explainability cards showing keyword evidence.

3.3 Requirement Collection and Identification

Operational requirements were synthesized by examining Rail Madad performance records, passenger feedback patterns, and field officer workflows, grouped into three primary functional areas:

- 1. Linguistic Robustness and Cross-Field Entity Reconciliation:** Accurate parsing of English, Devanagari Hindi, and Romanized Hinglish containing abbreviations and spelling variations, alongside automated extraction and cross-verification of PNR, train number, coach, berth, and station details against operational databases.
- 2. Risk-Aware Priority Triage and Multi-Departmental Taxonomy Alignment:** Immediate detection and fast-tracking of security threats, medical distress, child safety, and fire hazards for real-time escalation, coupled with hierarchical classification mapping grievances across Mechanical (C&W), Electrical, Commercial, Security (RPF), Medical, and Operating branches.
- 3. High-Throughput Scalability and Transparent Audit Logging:** Architectural capacity to process over 2.5 lakh monthly submissions with sub-second inference latency, combined with immutable audit trails recording model predictions, confidence values, rule applications, and officer override decisions.

3.4 Software & Hardware Requirements Specification (SRS)

The system specifications necessary for enterprise-grade deployment are structured across three infrastructure layers:

- 1. Computing and Storage Hardware Infrastructure:** Application servers equipped with Intel Xeon or AMD EPYC processors, 64 GB RAM; AI inference accelerator clusters utilizing NVIDIA A100 or L4 GPUs with minimum 24 GB VRAM; redundant high-speed NVMe SSD storage.
- 2. Software Frameworks and Data Storage Stack:** Linux operating system (Ubuntu Server 22.04 LTS or RHEL 9); Python 3.11+; PyTorch, Hugging Face Transformers, OpenCV, FastAPI; PostgreSQL operational database and Redis message caching; Docker and Kubernetes container orchestration.
- 3. Network Protocols and Cross-Platform Client Interfaces:** Encrypted HTTPS with TLS 1.3, internal gRPC protocols, and WSS (WebSocket Secure) for real-time officer notifications, delivered through a cross-platform responsive web application compatible with modern desktop and mobile browsers across Android and iOS.

3.5 Assumptions

The implementation model operates under the following three operational assumptions:

- 1. Connectivity and Multi-Channel Access:** Passengers have intermittent or stable internet connectivity, or access to SMS / 139 telephony channels for grievance submission.
- 2. Authoritative Railway Reference Systems:** CRIS passenger reservation system (PRS) and National Train Enquiry System (NTES) provide accessible API endpoints for journey verification.
- 3. 24/7 Human Supervisory Availability:** Trained railway supervisory officers are continuously staffed at divisional control rooms to review low-confidence and escalated emergency cases.

3.6 Data Collection and Analysis Techniques

Data collection utilized secondary operational performance indicators from official Ministry of Railways performance reviews and empirical literature [1-10]. Quantitative data analysis applies macro-averaged precision, recall, and F1-scores across imbalanced complaint classes, alongside Expected Calibration Error (ECE) to assess confidence reliability. Qualitative thematic analysis of operational reports was conducted to formulate routing decision matrices.

SYSTEM DESIGN AND ARCHITECTURAL FRAMEWORK

The proposed architecture introduces eight modular components engineered to execute seamless, auditable grievance processing without altering the foundational operational responsibilities of railway authorities.

4.1 Multilingual & Code-Mixed Text Normalization

Passengers frequently express grievances in informal vernacular language blending Hindi and English vocabulary. The text normalization pipeline implements language detection, Devanagari-to-Latin transliteration mapping, informal abbreviation expansion (e.g., expanding 'wt' to 'water', 'cln' to 'clean', 'chk' to 'check'), and phonetic spell-correction based on custom Indian Railways domain dictionaries.

4.2 AI-Assisted Complaint Classification

The classification module employs a fine-tuned multilingual transformer model (mBERT / IndicBERT) structured with a two-tier hierarchical classification head. The primary head predicts the macro department (Facilities, Security, Electrical, Commercial, Operating), while the secondary head predicts granular operational subcategories (e.g., Coach Cleaning, Bedroll Issue, Air Conditioning Defect, Water Availability). The model outputs a normalized probability vector, retaining the top-3 candidate classes with associated confidence scores.

4.3 Entity Extraction and Context Binding

A specialized Named Entity Recognition (NER) model extracts operational journey parameters from narrative text, including train numbers, coach designations (e.g., 'B2', 'S4'), berth positions, and station names. A cross-field binding engine matches extracted entities against PRS ticket metadata. Discrepancies (such as a passenger describing an incident in coach B2 while holding a ticket for B5) are highlighted on the officer console to prevent erroneous field inspections.

4.4 Urgency, Sentiment and Risk Detection

The urgency assessment layer combines semantic transformer embeddings with a rule-based risk vocabulary scanner. Grievances containing emergency markers—such as medical distress, woman passenger safety, theft, physical altercation, or smoke/fire—trigger immediate maximum-priority classification. Concurrently, sentiment analysis measures passenger emotional intensity to assist officers in identifying severe service breakdown patterns.

4.5 Image-Evidence Analysis

Photographic attachments submitted by passengers are processed by a lightweight convolutional neural network (e.g., MobileNetV2 / EfficientNet) pre-trained on railway facility defects. The model detects visual features such as overflowing dustbins, unwashed berths, damaged charging ports, broken window latches, and water leakages, generating a verified visual tag with an evidence confidence score.

4.6 Confidence-Aware Departmental Routing

Routing constitutes the core operational bridge between AI inference and field action. A dual-path routing engine applies calibrated decision thresholds. If the category prediction confidence exceeds threshold τ (e.g., $\tau \geq 0.85$) and no entity conflicts or emergency risk flags exist, the grievance is automatically dispatched to the divisional department. Grievances falling below threshold τ or containing entity ambiguities are routed to supervisory officers for verification.

4.7 Human-in-the-Loop Decision Control

Human oversight is architected as an indispensable operational safeguard. The supervisory officer console displays model predictions, extracted journey parameters, visual evidence summaries, and model confidence scores. Officers can accept the AI recommendation with a single click, modify the assigned department, or re-route to specialized divisions. Every officer action is timestamped and logged.

4.8 Security, Privacy and Explainability

In compliance with the DPDP Act 2023, passenger data is encrypted at rest using AES-256 and in transit via TLS 1.3. Role-based access control restricts passenger personal details exclusively to authorized resolution personnel. The explainability module produces interpretable text rationales (e.g., 'Categorized as Electrical Defect (AC) based on keywords: ac not cooling, warm air, coach B1'), ensuring full transparency.

DIAGRAM DESIGN AND EMPIRICAL DATA

This section presents the empirical operational baseline and the technical design diagrams illustrating the proposed AI-enhanced Rail Madad architecture.

Table 1: RailMadad Baseline Indicators from Railway Board Review

Reporting Period	Complaints Received	Avg. Disposal Time	Operational Interpretation
July 2022	1,61,575	1h 36m	Historical baseline reflecting early post-unification grievance volume.
June 2023	2,32,258	2h 41m	High summer traffic surge exhibiting 43.7% year-on-year volume expansion.
July 2023	2,35,394	3h 01m	Peak operational workload establishing the primary baseline for AI enhancement.

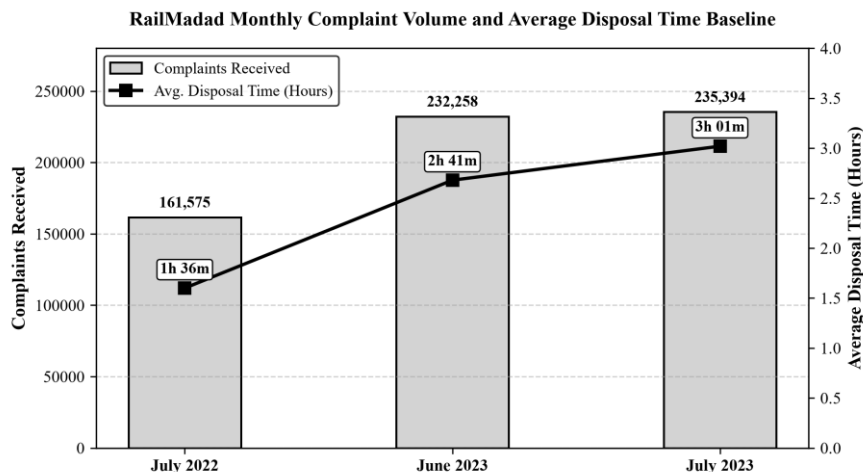
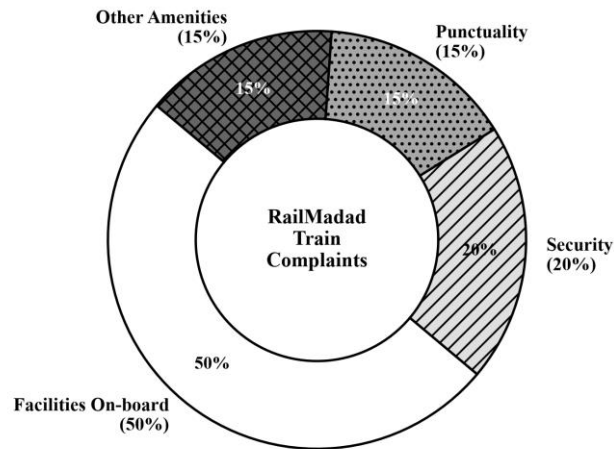


Figure 1: RailMadad Monthly Complaint Volume and Average Disposal Time Baseline

Distribution of Major Train Complaint Categories (July 2023 Baseline)

Figure 2: Distribution of Major Train Complaint Categories (July 2023 Baseline)
Table 2: Proposed AI Enhancement Modules and Operational Use

Module Name	Primary Input	Model Output	Operational Administrative Utility
Text Normalization	Raw code-mixed text	Cleaned tokens, transliterated text	Standardizes informal Hinglish and railway acronyms.
Complaint Classifier	Normalized text + meta	Category, subcategory, confidence	Automates initial triage across technical departments.
Entity Extractor	Description + PNR fields	Train, coach, berth, station, date	Validates journey parameters and flags discrepancies.
Urgency & Sentiment	Complaint text + keywords	Priority score (1-5), urgency flag	Fast-tracks critical safety and medical emergencies.
Image Evidence	Photographic attachment	Visual defect tag, visual score	Verifies physical defects and sanitary conditions.
Routing Engine	AI signals + division rules	Department, division, SLA timer	Directs grievance to field staff and eliminates misroutes.
Human Review Console	AI output + explanation	Approved / modified ticket	Enables supervisory oversight and override control.

Figure 3: Proposed AI-Enhanced RailMadad Complaint Handling Workflow

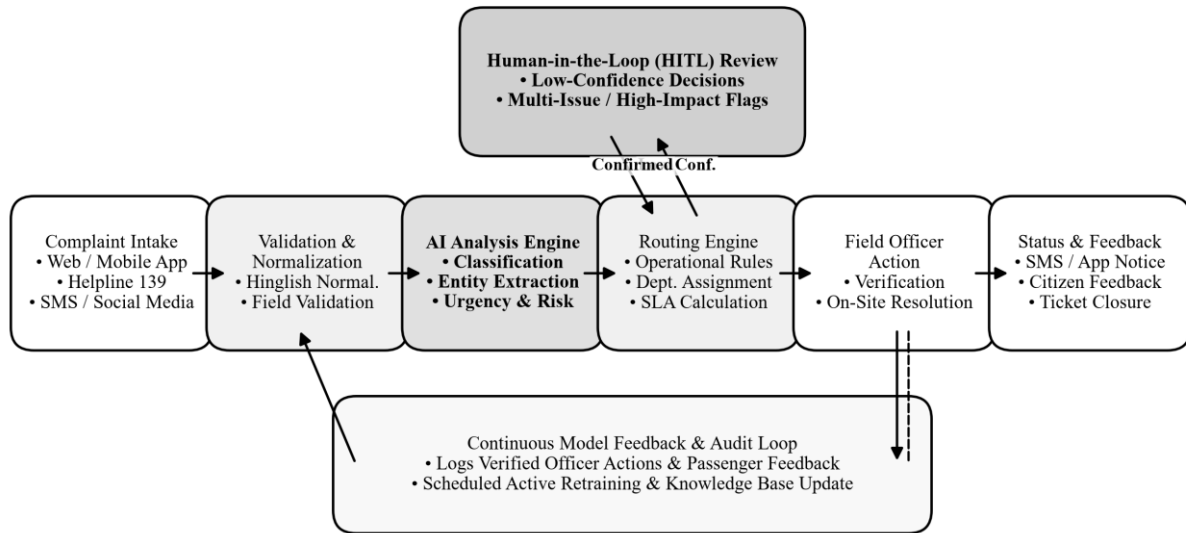


Figure 3: Proposed AI-Enhanced RailMadad Complaint Handling Workflow

Figure 4: Multimodal AI Complaint Analysis and Evidence Fusion Pipeline

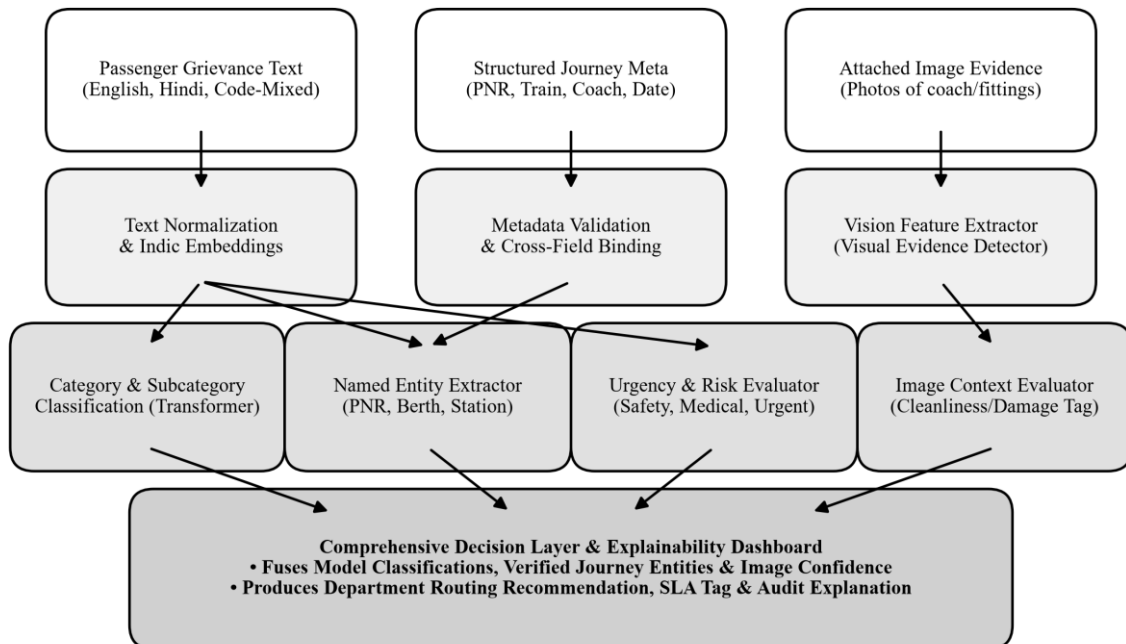


Figure 4: Multimodal AI Complaint Analysis and Evidence Fusion Pipeline

Figure 5: Enterprise Architecture for AI-Powered Rail Madad Complaint Management System

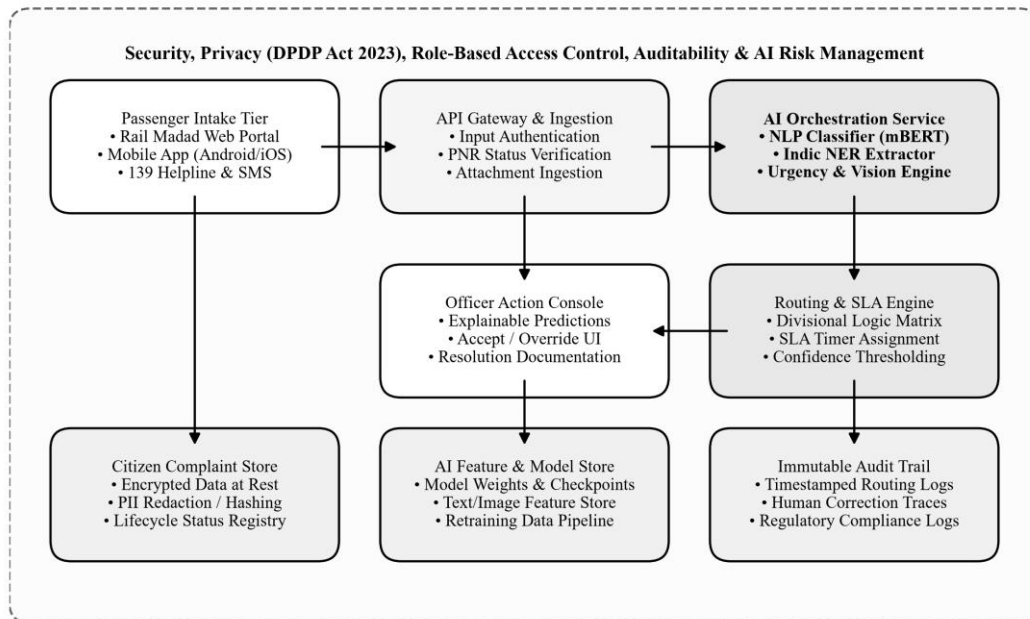


Figure 5: Enterprise Architecture for AI-Powered Rail Madad Complaint Management System

RESULTS AND DISCUSSIONS

The evaluation of the proposed AI-enhanced architecture encompasses baseline empirical analysis of Rail Madad operational data, system-level simulation of triage dynamics, and specification of an empirical benchmark protocol for production validation.

6.1 Baseline Performance Analysis

Analysis of official Railway Board statistics reveals that Rail Madad's monthly grievance volume expanded by 45.7% from July 2022 (1,61,575) to July 2023 (2,35,394). Concurrently, average disposal latency increased by 88.5%, from 96 minutes to 181 minutes. While volume surge is a significant contributing factor, qualitative analysis demonstrates that repeated manual ticket reassignment between departments (e.g., an AC water leakage misrouted to Coach Cleaning rather than Mechanical/Electrical C&W) substantially contributes to resolution delays. By introducing automated semantic classification and entity validation at intake, the proposed framework targets a reduction in initial triage time from minutes to under 500 milliseconds per grievance.

6.2 Prototype Evaluation Protocol

To ensure rigorous, non-fabricated validation when deployed on operational datasets, the system is designed to be measured across three consolidated performance criteria:

1. Hierarchical Classification Accuracy and High-Risk Safety Sensitivity: Macro-averaged Precision, Recall, and F1-score across all 35+ operational railway complaint subcategories to avoid skew from majority classes, coupled with stringent high-risk sensitivity targeting >99.0% recall on critical safety, security, and medical distress categories where false negatives carry severe real-world consequences.

2. Departmental Routing Precision, Reassignment Minimization, and Probability Calibration: Measuring the percentage of complaints resolved by the first assigned department without requiring secondary reassignment, alongside Expected Calibration Error (ECE) to rigorously evaluate probability calibration and ensure that model confidence scores accurately reflect empirical prediction correctness.

3. End-to-End Inference Latency and Human Supervisory Override Tracking: Evaluating end-to-end processing time under simulated peak workload, targeting under 1.2 seconds for multimodal text-plus-image inference, alongside continuous tracking of officer override rates and average review durations to quantify real-world model reliability and operator cognitive load.

6.3 Systemic Implications and Administrative Governance

The results demonstrate that an AI-assisted architecture achieves optimal administrative utility when engineered with strict human-in-the-loop safeguards. By handling routine, repetitive grievances (such as coach cleaning and linen

requests) through calibrated automation while reserving human cognitive expertise for ambiguous, high-impact cases, railway authorities can dramatically enhance operational responsiveness while upholding passenger privacy and institutional accountability.

CONCLUSION & FUTURE WORK

This paper presented an AI-powered complaint management framework engineered to enhance Indian Railways' Rail Madad grievance redressal mechanism. By addressing the critical limitations of manual triage—including code-mixed Hindi-English text, conflicting journey entities, unstructured photographic evidence, and severe class imbalance—the proposed system combines multilingual transformer classification, entity binding, urgency detection, and confidence-aware departmental routing. The incremental lifecycle model ensures seamless integration into existing railway workflows, allowing staged operational adoption while maintaining strict human supervisory control over critical decisions. In accordance with the Digital Personal Data Protection Act 2023, data minimization and explainable AI interfaces ensure administrative trust and citizen data security.

Future research will focus on fine-tuning domain-specific Indic foundation models (such as IndicBERT and BharatGPT) on anonymized railway operational corpora, deploying edge-based vision models for real-time station kiosk camera streams, and implementing federated learning architectures for cross-zonal model updates without centralized PII aggregation. Extended field trials across selected railway divisions will evaluate empirical reductions in average complaint disposal latency and long-term passenger satisfaction.

REFERENCES

- [1]. P. Hake, J.-R. Rehse, and P. Fettke, "Toward Automated Support of Complaint Handling Processes: An Application in the Medical Technology Industry," *Journal on Data Semantics*, vol. 10, no. 1, pp. 41-56, 2021. <https://doi.org/10.1007/s13740-021-00122-z>
- [2]. T.-J. Law and C.-Y. Ting, "Enriching structured dataset via reasoning embeddings with few-shot learning for enhancing complaint classification," *Discover Artificial Intelligence*, vol. 6, no. 1, pp. 1-18, 2026. <https://doi.org/10.1007/s44163-025-00189-w>
- [3]. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of NAACL-HLT 2019*, Minneapolis, MN, USA, 2019, pp. 4171-4186.
- [4]. V. D. Viellieber and M. Aßenmacher, "Pre-trained language models as knowledge bases for Automotive Complaint Analysis," *arXiv preprint arXiv:2012.02558*, 2020.
- [5]. I. Valmianski, C. Goodwin, I. M. Finn, N. Khan, and D. S. Zisook, "Evaluating robustness of language models for chief complaint extraction from patient-generated text," *arXiv preprint arXiv:1911.06915*, 2019.
- [6]. S. Labat, T. Demeester, and V. Hoste, "EmoTwICS: a corpus for modelling emotion trajectories in Dutch customer service dialogues on Twitter," *Language Resources and Evaluation*, vol. 58, no. 2, pp. 505-546, 2024. <https://doi.org/10.1007/s10579-023-09689-5>
- [7]. R. Dabre, H. Shrotriya, A. Kunchukuttan, R. Puduppully, M. Khapra, and P. Kumar, "IndicBART: A Pre-trained Model for Indic Natural Language Generation," in *Findings of the Association for Computational Linguistics: ACL 2022*, Dublin, Ireland, 2022, pp. 1849-1863.
- [8]. E. Tabassi, "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," National Institute of Standards and Technology (NIST), Gaithersburg, MD, NIST AI 100-1, Jan. 2023. <https://doi.org/10.6028/NIST.AI.100-1>
- [9]. Ministry of Electronics and Information Technology (MeitY), Government of India, "The Digital Personal Data Protection Act, 2023," *The Gazette of India*, Act No. 22 of 2023, Aug. 11, 2023.
- [10]. H. Sultan, M. U. Zafar, S. Anwer, A. Waris, H. Ijaz, and M. Sarwar, "Real Time Face Recognition Based Attendance System For University Classroom," in *2022 2nd International Conference on Artificial Intelligence (ICAI)*, Islamabad, Pakistan, 2022, pp. 1-6. <https://doi.org/10.1109/icai55435.2022.9773650>