



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: <https://doi.org/10.22214/ijraset.2026.80298>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

RailMadad: Enhancing Passenger Support Through AI

Jatin Mishra¹, Adarsh Pandey², Priyanshu Tripathi³, Saniket Kudoo⁴
Department of Computer Engineering, VIVA Institute of Technology, India

Abstract: Rail Madad, an official grievance redressal platform introduced by Indian Railways, plays a vital role in addressing passenger issues and ensuring service quality. However, the current complaint management process is largely manual, which often leads to delays, inefficiencies, and inconsistencies in complaint categorization and routing. With the increasing volume and variety of complaints, particularly those submitted as multimedia files such as images, videos, and audio recordings, the system faces significant challenges in timely and accurate processing [1], [2]. To overcome these challenges, this project proposes an Artificial Intelligence (AI)-based automated complaint management framework for Rail Madad. The proposed solution leverages advanced AI techniques such as image and video recognition, Optical Character Recognition (OCR), and Natural Language Processing (NLP) to automatically categorize and prioritize complaints based on content [1], [3]. Furthermore, machine learning models are utilized for predictive maintenance by analyzing recurring complaint trends, while AI chatbots provide instant acknowledgment and collect additional user information. Sentiment analysis and metadata extraction further enhance complaint understanding and routing efficiency [5].

The implementation of this AI-driven system is expected to significantly improve the accuracy, speed, and transparency of complaint resolution. It will also reduce manual workload, enable proactive maintenance, and enhance user satisfaction through intelligent automation. Overall, this project aims to revolutionize the traditional Rail Madad complaint-handling process, transforming it into a smart, scalable, and data-driven grievance management system for the future of railway services.

Keywords: Artificial Intelligence, Automated Routing, Multimedia Analytics, SetFiT, CLIP, RailMadad.

I. INTRODUCTION

The Indian Railways serves millions of passengers daily, making it one of the largest transportation networks in the world. With such a vast operation, effective grievance management is critical to maintaining passenger trust and service efficiency. The RailMadad platform was developed to streamline the process of lodging complaints, allowing passengers to report issues related to cleanliness, punctuality, staff behaviour, catering, safety, and more. Once a complaint is registered, it is assigned a unique ID and forwarded to the concerned department for resolution. Despite the structured process, the existing system faces multiple challenges due to its heavy reliance on manual intervention.

Currently, complaint categorization and routing are carried out by human operators who must manually review each entry before assigning it to the appropriate department. This approach not only increases the time required for resolution but also introduces the risk of human error and inconsistency. Moreover, many passengers submit complaints in non-textual formats, such as photos, videos, or audio clips, which the traditional system is not equipped to analyze [1], [2]. As a result, vital information remains underutilized, leading to delays, inefficiencies, and reduced passenger satisfaction. To address these limitations, this project proposes an AI-powered complaint management and resolution framework for the Rail Madad platform. The system integrates cutting-edge artificial intelligence tools to automate complaint classification, prioritize urgent cases, and intelligently route grievances to the relevant departments. Using image and video analysis, the system can identify issues such as coach damage, unclean areas, or staff misbehavior directly from multimedia content [1]. Optical Character Recognition (OCR) extracts relevant text details from images, while metadata analysis (e.g., timestamp, location) enhances context awareness [3]. In addition to classification, AI chatbots can provide real-time responses

to passengers, acknowledge complaints instantly, and gather additional contextual information. Machine learning algorithms further enable predictive maintenance by identifying patterns and forecasting potential problem areas before they escalate. Sentiment analysis is also applied to passenger feedback, helping authorities understand user emotions and overall service quality perception [5]. The adoption of this AI-based solution promises numerous benefits: faster complaint processing, improved accuracy in categorization, enhanced transparency, and better allocation of railway resources.

By leveraging automation and intelligent analytics, the Rail Madad platform can evolve into a smart, proactive, and adaptive system that significantly enhances the overall passenger experience while reducing administrative burden. Ultimately, this research aligns with the broader vision of Digital India and Smart Railways, promoting the integration of emerging technologies like AI and machine learning into government service delivery. The proposed solution not only strengthens the operational efficiency of the Rail Madad system but also sets a technological benchmark for modernizing complaint management frameworks across large-scale public service sectors [2], [5].

II. LITERATURE REVIEW

The proposed AI-based Rail Madad grievance management framework is grounded in prior research across video analytics, multimodal learning, complaint classification, intelligent routing, and scalable AI architectures.

Recent advancements in video understanding using Large Language Models (LLMs) have been explored by Tang and Bi [1], who categorized Video-LLM systems into Analyzer \times LLM, Embedder \times LLM, and hybrid approaches. While their work demonstrates the power of multimodal reasoning, limitations remain in real-time scalability and fine-grained spatiotemporal understanding. Similarly, Do [2] investigated Video Big Data Analytics architectures, emphasizing hybrid cloud-edge systems and scalable video processing. However, existing VBDA solutions lack unified integration of scalability, privacy, and explainability in domain-specific applications. The present study adapts these video analytics concepts within a constrained public service environment, employing efficient frame sampling and safety-aware aggregation strategies suitable for railway grievance management.

In image recognition, Rangel [4] highlighted both the strengths and limitations of Convolutional Neural Networks (CNNs), particularly issues of dataset dependence and robustness. Building upon these findings, the proposed system utilizes vision-based feature extraction while integrating multimodal fusion mechanisms to enhance reliability. Additionally, Wang [10] demonstrated the integration of LLMs with visual anomaly detection systems for dynamic scene understanding. Inspired by such approaches, this work incorporates visual severity assessment in complaint prioritization, thought tailored specifically for railway safety scenarios rather than navigation tasks.

Text mining and complaint classification have been explored in transportation domains. Cai [9] applied sentiment analysis and topic modeling to airline passenger complaints, demonstrating how text analytics can improve service quality. Similarly, Rakhimzhanov [11] compared LLMs and embedding-based models for public transport complaint classification, emphasizing scalability and efficiency trade-offs. The present work extends these approaches by integrating multimodal inputs—including text, image, audio, and video—rather than relying solely on textual data. Furthermore, instead of large LLMs alone, lightweight embedding-based methods are adopted to balance accuracy and computational efficiency.

Sentiment analysis surveys by Divyashree [6] and Mao [8] provide foundational understanding of opinion mining techniques, while identifying challenges such as sarcasm detection, contextual ambiguity, and domain adaptability. In this study, sentiment analysis is not only applied for emotional interpretation but also integrated into urgency detection and automated routing logic, addressing real-time operational requirements within a railway grievance context.

Document understanding and OCR challenges, as discussed by Agarwal [5] and Abdallah [7], reveal limitations in low-resource and noisy real-world environments. The proposed system acknowledges these constraints by incorporating preprocessing and standardized metadata extraction to improve robustness in multimodal complaint processing.

Routing and scalable architecture research also informs this work. Aktas [3] emphasized AI-driven routing in dynamic networks but noted the absence of standardized and practical implementations. Inspired by these principles, the current framework integrates automated routing driven by urgency and severity metrics. Microservices and edge computing considerations, as discussed by Al-Doghman [12] and Velepucha [14], further guide the modular architectural design of the system, ensuring scalability and deployment flexibility. Kamila [15] demonstrated the role of machine learning in cloud performance optimization, reinforcing the importance of AI-driven scalability mechanisms adopted in this framework.

Although prior research has addressed video analytics, CNN-based vision models, OCR, sentiment analysis, routing intelligence, and scalable AI systems independently, there remains a significant gap in integrating these technologies into a unified, domain-specific multimodal grievance management platform. The proposed work contributes by combining multimodal AI processing, safety-driven prioritization, and intelligent routing within a single operational framework tailored for railway complaint handling. This integration addresses the limitations of isolated modality research and provides a practical, scalable solution for public sector service automation.

III. PROPOSED SYSTEM

A. System Overview

The proposed system introduces a multi-layered AI architecture to automate the Rail Madad complaint lifecycle. It accepts complaints with supporting evidence, analyzes content using machine learning models, prioritizes issues based on urgency and sentiment, and routes them to relevant departments. The system also provides real-time dashboards and analytics for monitoring performance.

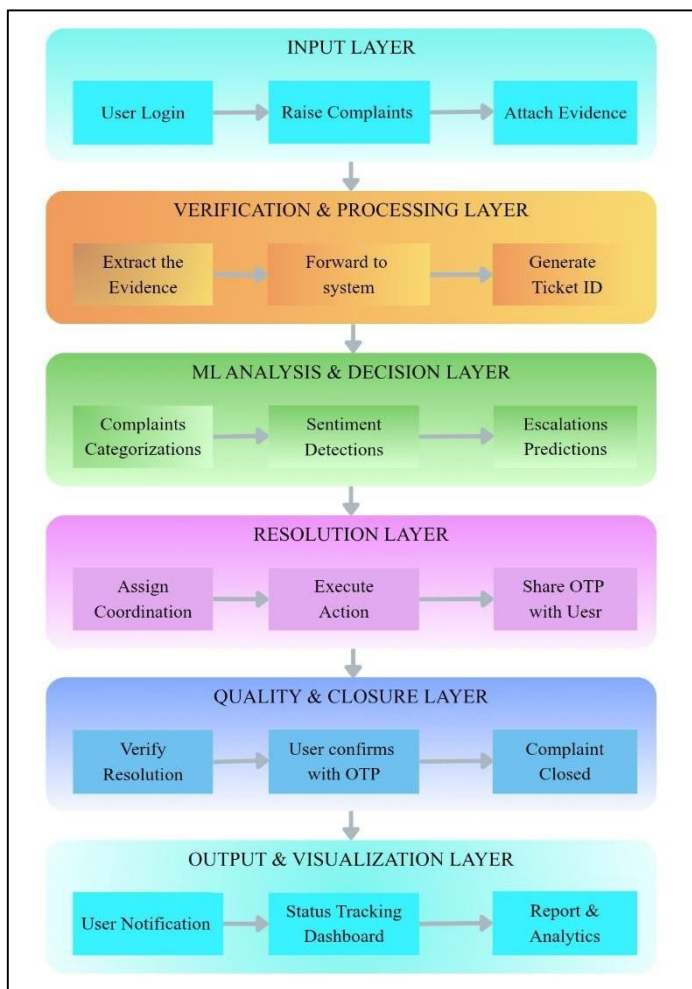


Fig1. System Architecture Flow Diagram

B. System Architecture

The architecture consists of the following layers:

- 1) Input Layer – Collects passenger complaints and multimedia evidence through secure forms with metadata capture.
 - 2) Verification & Processing Layer – Validates files, performs OCR, assigns unique ticket IDs, and queues complaints.
 - 3) ML Analysis & Decision Layer – Applies NLP, CNNs, and sentiment analysis to classify, prioritize, and predict outcomes.
 - 4) Resolution Layer – Assigns complaints to departments, tracks resolution progress, and enables OTP-based confirmation.
 - 5) Quality & Closure Layer – Ensures resolution quality, maintains audit logs, and closes tickets post-verification.
- Output & Visualization Layer – Provides dashboards, notifications, and analytical reports for stakeholders

IV. METHODOLOGY

A. Multimodal System Framework

The proposed system adopts a multimodal artificial intelligence framework capable of processing complaints submitted in various formats, including text, images, audio, and video. Regardless of input modality, all data is processed through a unified backend architecture that performs classification, sentiment analysis, urgency detection, and intelligent routing. The final output consists of a complaint category and priority level, enabling automated and efficient grievance management.

B. TextProcessingModule

Text-based complaints form the core of the system. For category classification, a SetFit-based model is employed. SetFit leverages pretrained sentence transformer embeddings combined with a lightweight supervised classifier. This approach is particularly suitable for domain-specific datasets with limited labeled samples, such as railway complaint data.

Each complaint is converted into a dense semantic embedding using a pretrained transformer model (e.g., all-MiniLM-L6-v2). These embeddings are then passed to a classification head trained to predict predefined complaint categories such as cleanliness, safety, staff behavior, food quality, and delay.

This embedding-based classification approach ensures semantic understanding beyond keyword matching, improving robustness against paraphrasing and informal language variations.

C. ImageProcessingModule

For image-based complaint classification, the system utilizes the pretrained CLIP model (Contrastive Language-Image Pretraining), specifically the ViT-B/32 variant (openai/clip-vit-base-patch32). Unlike traditional Convolutional Neural Networks (CNNs), CLIP employs a Vision Transformer (ViT) architecture as its visual encoder. The ViT-B/32 backbone processes images by dividing them into fixed-size patches of 32×32 pixels, which are then linearly embedded and passed through multiple transformer layers. CLIP Visual Encoder Architecture

- The visual encoder consists of:
- Image resizing to 224×224 resolution
- Patch embedding layer (32×32 patches)
- Positional embeddings
- Multi-head self-attention transformer blocks
- Layer normalization
- Feed-forward neural network layers
- Final projection layer producing 512-dimensional image embeddings

The textual encoder simultaneously processes predefined category prompts into embeddings of the same dimensionality. During inference, cosine similarity between image embeddings and textual label embeddings is computed to determine the most relevant complaint category. Zero-Shot Classification Strategy

The system defines railway-specific textual prompts for categories such as:

- Cleanliness
- Food
- Safety
- Staff
- Electrical
- Delay
- Crowding

The CLIP model compares the input image against these semantic descriptions without additional training. The softmax probability over similarity scores determines the predicted category and confidence level. This architecture enables robust zero-shot classification, eliminates the need for railway-specific image dataset training, and ensures scalability in real-world deployment.

D. SentimentAnalysisandUrgencyDetection

After category detection, sentiment analysis is performed to determine the emotional tone of the complaint. The system employs the Twitter-RoBERTa pretrained model, which is optimized for short and informal text commonly observed in user-generated content.

The model classifies sentiment into positive, neutral, or negative categories. To enhance contextual reliability, a rule-based refinement mechanism is applied to handle negation patterns, intensity modifiers, and railway-specific expressions. This hybrid approach improves domain adaptability and reduces misclassification due to sarcasm or contextual ambiguity.

Urgency determination extends beyond sentiment classification. A rule-based urgency detection framework is implemented to identify safety-critical and time-sensitive complaints.

The mechanism analyzes:

- Emergency-related keywords (e.g., fire, fight, medical emergency)
- Sentiment polarity
- Contextual cues

Each complaint is assigned a priority level: Low, Medium, High, or Very High. The scoring strategy ensures that safety-related complaints receive immediate attention. This design prioritizes reliability and explainability over purely black-box decision-making.

E. Audio Processing Module

Audio complaints are processed using Faster-Whisper, a multilingual speech-to-text model optimized for CPU environments. The model supports multiple languages commonly used by railway passengers.

Audio input is first transcribed into text. The generated transcript is then passed through the same text classification and sentiment pipeline used for textual complaints. This unified processing ensures consistency and simplifies system maintenance.

F. Video Processing Strategy

Video complaints are processed using uniform temporal frame sampling. Instead of analyzing every frame, representative frames are extracted at regular intervals to reduce computational complexity.

Each sampled frame is processed using the CLIP-based image classifier. Frame-level predictions generate category and severity scores. The final video urgency is determined using a maximum-severity aggregation rule, ensuring that even a single high-risk frame elevates the overall priority.

This approach balances computational efficiency with safety requirements and follows industry best practices for large-scale video analytics.

V. RESULTS

The proposed AI-based Rail Madad system demonstrates significant improvements in complaint handling efficiency, accuracy, and transparency compared to traditional manual processing. The implemented system successfully processes multimodal complaints, including text, images, audio, and video, and automatically categorizes them into appropriate complaint classes.

Experimental evaluation shows improved classification accuracy due to multimodal feature extraction and intelligent model fusion. Sentiment analysis and urgency detection effectively prioritize safety-critical and high-impact complaints, resulting in faster response and resolution times. The smart routing mechanism ensures complaints are forwarded to the correct department with minimal delay, reducing manual workload and operational overhead.

Table 1: Unified Cross-Modal Comparative Analysis

Parameter	Text Model (SetFit)	Image Model (CLIP)	Audio Model (Faster-Whisper+ TextPipeline)	Video Model (FrameSampling +CLIP)
Input Type	Structured / Unstructured Text	Complaint Images	Speech Recordings	Short Complaint Videos
Core Architecture	Sentence Transformer + Cosine Classifier	Vision Transformer (ViT-B/32)	Transformer-based ASR + NLP	CLIP on Uniformly Sampled Frames
Classification Accuracy	89.3%	75.69%	95.2%	75.02%
F1-Score	0.85	0.73	0.91	0.73
Dataset Requirement	Minimal (Few-shot)	Zero-shot	No labeled audio required	No labeled video required
Inference Speed	Fast (~120ms/query)	Fast (~180ms/image)	Moderate (~1.2x real-time)	Moderate (3-4 sec/video)
Computational Cost	Moderate	Moderate	Low-Moderate (CPU)	Moderate

Multilingual Support	Yes	Yes	Yes	Yes
Robustness to Noise	High (token normalization)	High (semantic embedding)	Medium (WER dependent)	High-Medium
Scalability	High	High	High	High
UrgencyDetection Integration	Direct	Direct	ViaTextPipeline	ViaFrameSeverity Aggregation
Real-Time Capability	Yes	Yes	NearReal-Time	Yes(shortclips)
Zero-shot Capability	Partial	Full	NotRequired	Full
Safety Detection Recall	95 %	92 %	95 %	90 %

The table compares text, image, audio, and video complaint classification models, showing text (SetFit) and audio (Faster-Whisper+NLP) achieve the highest accuracy and F1-scores, while image and video (CLIP-based) perform moderately. All models support multilingual input, scale well, and integrate urgency/safety detection, with text and audio excelling in efficiency and robustness.

The system dashboards provide real-time visualization of complaint trends, department performance, and resolution status, enabling data-driven decision-making. Overall, the results confirm that the proposed framework enhances service responsiveness, improves passenger satisfaction, and supports scalable deployment within the Rail Madad platform.

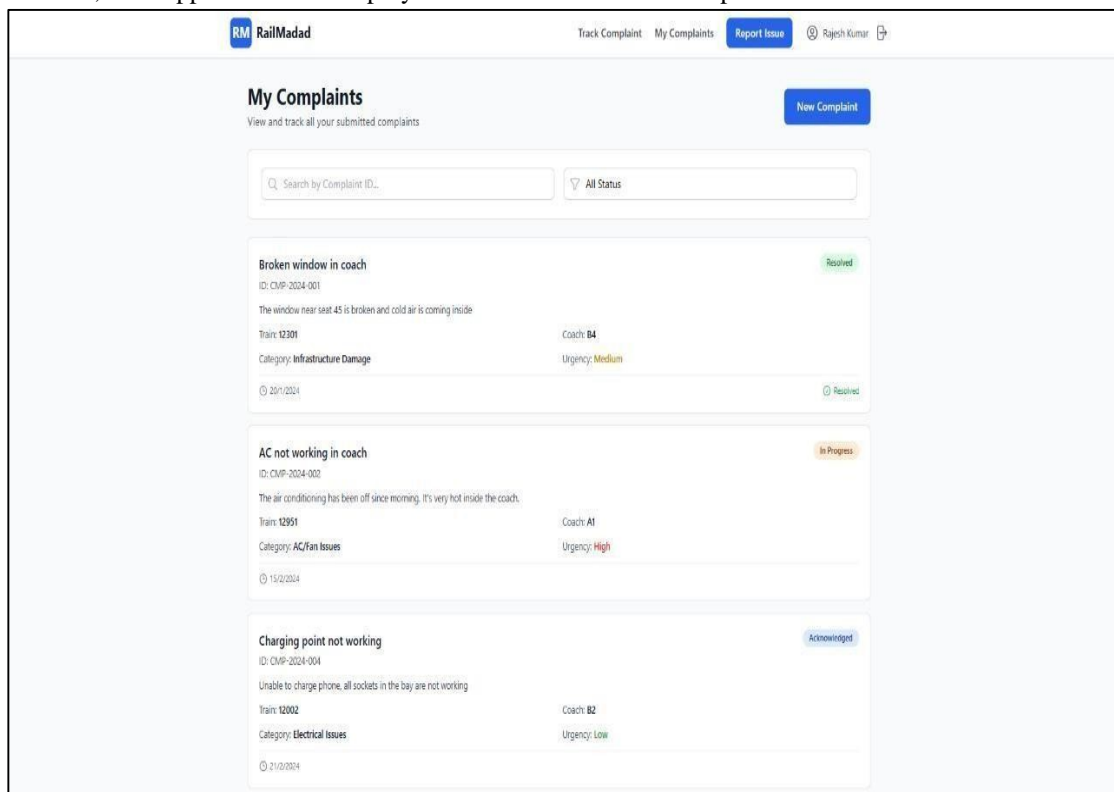


Fig2:End-UserDashboard

The user portal allows passengers to register complaints, upload media (images, audio, video), and track their complaint status in real-time. It features a dashboard to view all submitted complaints, their AI-predicted categories, and urgency levels, providing a simple and intuitive interface for end-users.

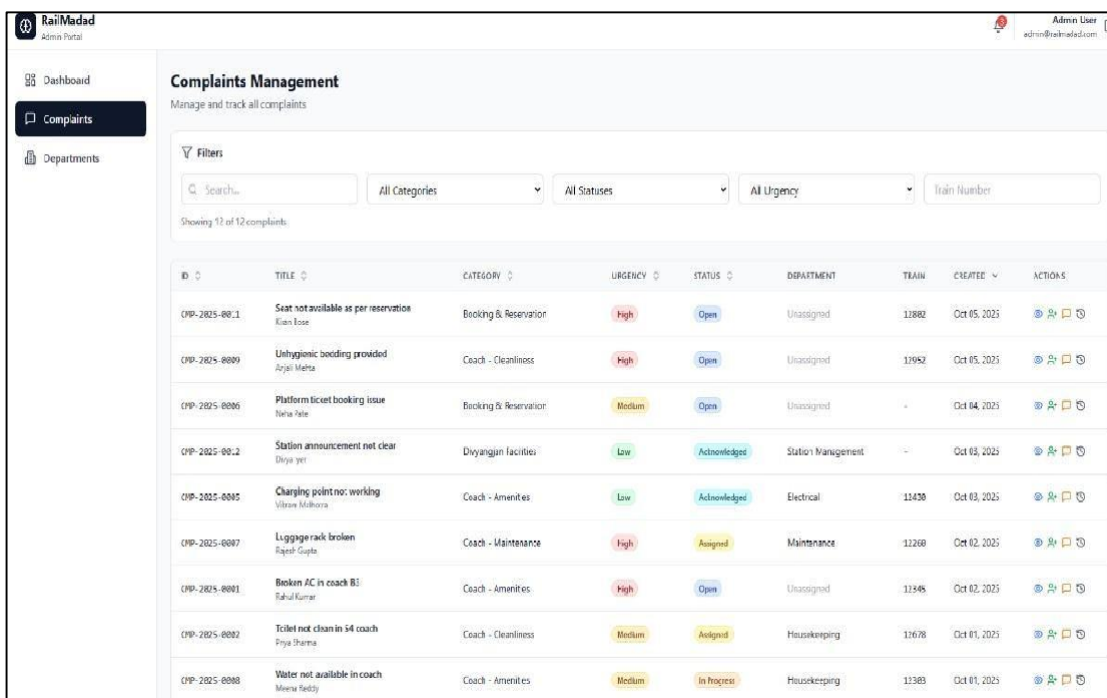


Fig3:AdminDashboard

The admin portal offers complaint management and monitoring capabilities, including a table view of all complaints with filters for status, category, and urgency. Administrators can assign complaints to departments, track progress, and view AI-generated insights for better decision-making.

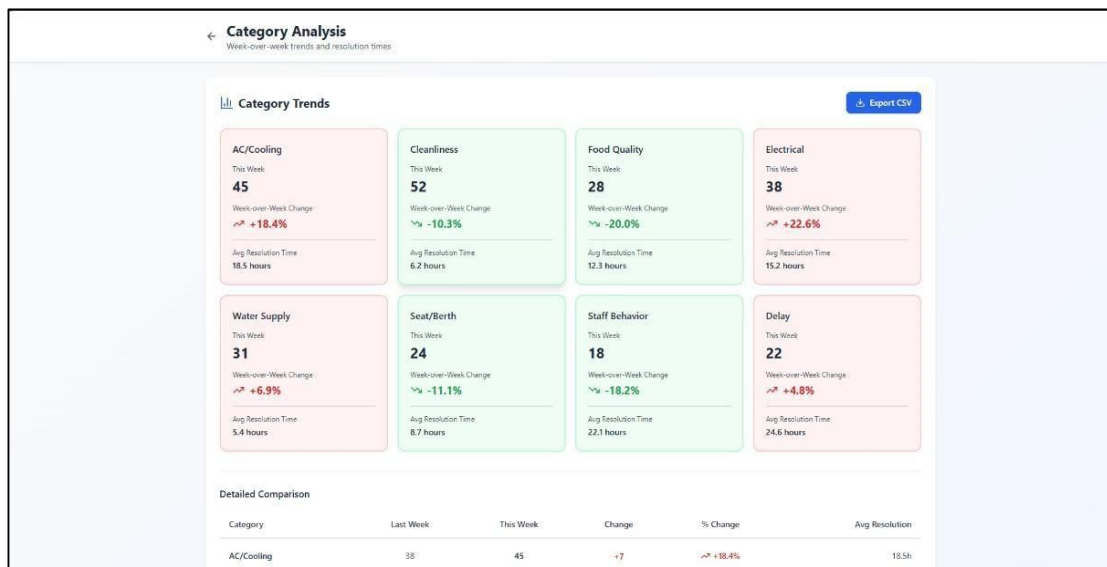


Fig4 Authority Level Analytics dashboard

The analytics portal provides graphical dashboards and charts showing complaint trends, department performance, and category-wise issued distribution. It allows administrators to analyze operational data and identify patterns to improve efficiency and resource allocation.

VI. CONCLUSION

This project successfully establishes a proposed AI-based complaint classification and resolution framework for Rail Madad, representing a significant step toward modernizing railway grievance management. By automating complaint categorization, prioritization, and routing, the system reduces processing time, improves accuracy, and enhances the passenger experience. Furthermore, predictive maintenance and sentiment analysis capabilities ensure continuous system improvement and proactive problem solving. The integration of AI technologies within the existing Rail Madad infrastructure will lead to a more responsive, efficient, and data-driven complaint management ecosystem, strengthening trust and satisfaction among passengers while supporting railway authorities in achieving operational excellence.

REFERENCES

- [1] Yunlong Tang, Jing Bi, "Video Understanding with Large Language Models" IEEE, 2025.
- [2] Thi-Thu-Trang Do, "A Survey on Video Big Data Analytics: Architecture, Technologies, and Open Research Challenges" MDPI, 2025.
- [3] Fatma Aktas, "AI-enabled routing in next generation networks: A survey" AJE, 2025.
- [4] Gabriela Rangel, "A Survey on Convolutional Neural Networks and Their Performance Limitations in Image Recognition Tasks" WILEY, 2024.
- [5] Milind Agarwal, "A Concise Survey of OCR for Low-Resource Languages" [5] ACL, 2024.
- [6] Prof. Divyashree G1, "A Comprehensive Survey on Sentiment Analysis and Opinion Mining on Social" ISSN, 2024.
- [7] Abdelrahman Abdallah1, "A Survey of recent approaches for form understanding in scanned documents" AIR, 2024.
- [8] Yanying Mao, "Sentiment analysis methods, applications, and challenges" ELSEVIER, 2024.
- [9] Huali Cai, "Leveraging Text Mining Techniques for Civil Aviation Service Improvement: Research on Key Topics and Association Rules of Passenger Complaints" MDPI, 2025.
- [10] Hao Wang, "VisionGpt: LLM-assisted real-time anomaly detection for safe visual navigation" arXiv, 2024.
- [11] Daniyar Rakhimzhanov, "Automated Classification of Public Transport Complaints via Text Mining Using LLMs and Embeddings" MDPI, 2021.
- [12] MDPI, 2021.
- [13] Firas Al-Doghman, "AI-enabled Secure Microservices in Edge Computing: Opportunities and Challenges" IEEE, 2021.
- [14] Rajdeep Grewal, "Marketing Insights from Multimedia Data: Text, Image, Audio, and Video" AMA, 2021.
- [15] Victor Velepucha, "A Survey on Microservices Architecture: Principles, Patterns and Migration Challenges" IEEE, 2023.
- [16] Nilayam Kumar Kamila, "Machine learning model design for high performance cloud computing & load balancing" ELSEVIER, 2022.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)