

Fynd AI Intern Assessment – Final Project Report

Project: AI-Powered Feedback Ecosystem (Task 1 & Task 2)

Author: Deepesh Yadav

GitHub Repository: <https://github.com/deepeshyadav760/AI-Feedback-System-web-based>

1. Overall Approach

I treated this project as the complete lifecycle of an AI feature. I began with empirical research (Task 1) to understand how Large Language Models interpret sentiment and star ratings. Once I identified the most reliable prompting strategy, I transitioned into systems engineering (Task 2) to build a production-ready environment. The guiding philosophy was 'AI-centric but Engineering-first', ensuring scalability, security, and robustness.

2. Design and Architecture Decisions

Decoupled Micro-Frontends: The User Dashboard and Admin Dashboard were developed as two independent React applications, allowing independent scaling and security boundaries.

Server-Side AI Orchestration: All LLM calls are handled in the Node.js/Express backend to protect the Groq API key and allow coordinated post-processing within a single request cycle.

Non-Relational Database (MongoDB): A document-based database was chosen to naturally store unstructured reviews and complex AI-generated metadata.

3. Prompt Iterations and Improvements

Iteration 1 – Zero-Shot: A fast baseline approach that occasionally missed nuanced sentiment.

Iteration 2 – Few-Shot: Provided calibration examples. While effective for large models like Llama-3 70B, the Llama-3.1 8B model frequently produced malformed JSON, reducing accuracy to 28%.

Iteration 3 – Chain-of-Thought (Winner): Forced structured reasoning across Service, Quality, and Value, achieving the highest accuracy of 75%.

4. Evaluation Methodology and Results (Task 1)

Dataset: Stratified sample of 100 Yelp reviews.

Model: Llama-3.1-8b-instant via Groq (selected due to free-tier token constraints).

Metrics: Accuracy, JSON Validity, and Reliability.

Results:

Zero-Shot: 70% accuracy (reliable but basic).

Few-Shot: 28% accuracy (formatting instability).

Chain-of-Thought: 75% accuracy (most consistent and nuanced).

5. System Behaviour, Trade-offs, and Limitations (Task 2)

System Behaviour: Each review submission triggers three parallel AI processes: an empathetic user response, a concise admin summary, and a three-step business action plan.

Trade-offs: The deployed system uses the Llama-3.3-70B-Versatile model, offering superior reasoning depth at the cost of higher latency. Free-tier Render hosting introduces cold-start delays, accepted for cost efficiency.

Limitations: The Admin Dashboard lacks authentication, and reviews are processed individually without long-term contextual memory. Addressing this would require authentication layers and a vector database (RAG).

Final Submission Links

GitHub Repository: <https://github.com/deepeshydav760/AI-Feedback-System-web-based>
User Dashboard: <https://fynd-user-feedback.vercel.app/>
Admin Dashboard: <https://fynd-admin-panel.vercel.app/>