GenAI-Powered Food Delivery Sentiment Analysis

Business Applications of Generative AI

Dean Ahlgren

November 2025

Contents / Agenda

Business Context & Problem Overview

Business Context:

- Food delivery companies process thousands of customer reviews daily
- Manual analysis is infeasible and introduces human inconsistency
- Unstructured text contains sarcasm, mixed emotions, and cultural idioms

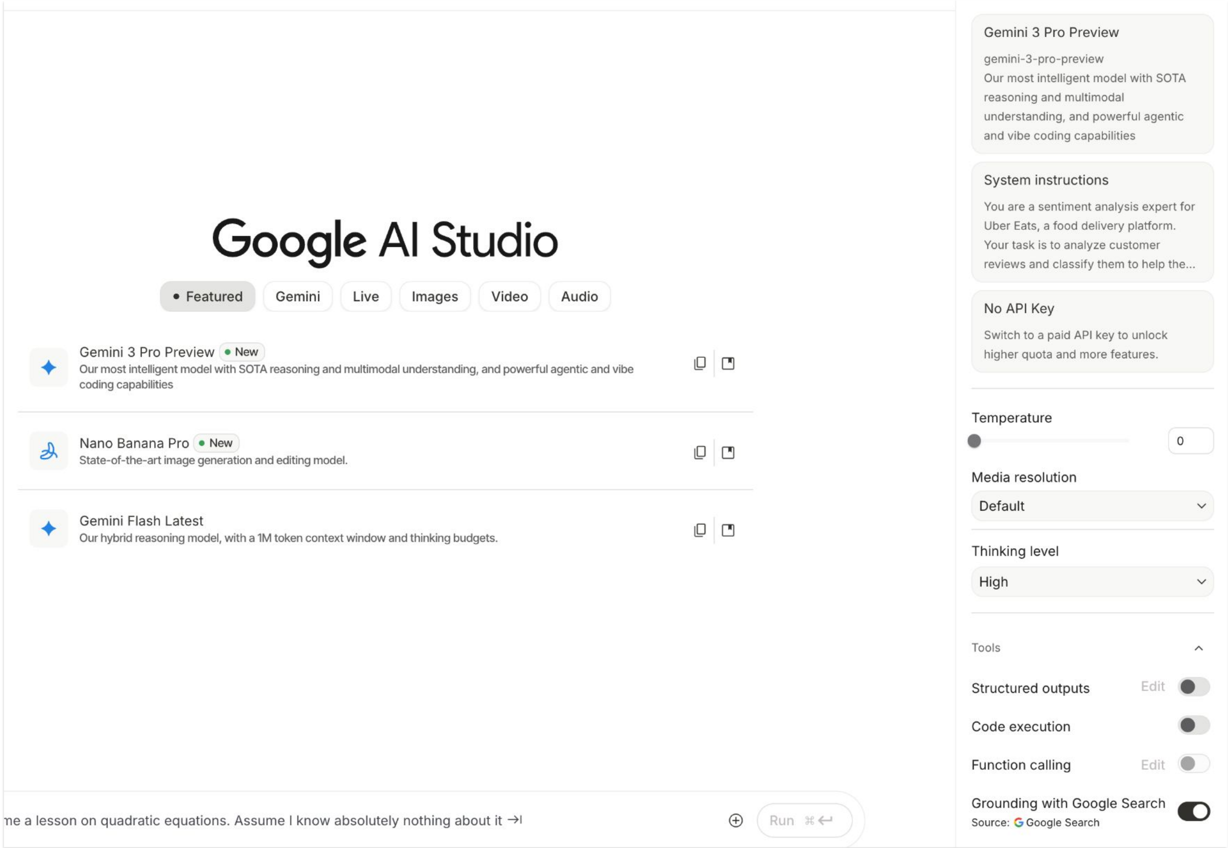
Solution Approach:

- Use Generative AI (Gemini 3 Pro) to automate sentiment classification
- Classify reviews as Positive, Negative, or Neutral
- Generate tags: Delivery Time, Food Quality, Price, Packaging, Overall Experience
- Provide actionable business recommendations for each review

Prompt Engineering - Platform Configuration

All prompts were executed in Google AI Studio using Gemini 3 Pro Preview, selected for its state-of-the-art reasoning capabilities.

The temperature parameter was set to 0 to ensure deterministic, consistent outputs suitable for business classification tasks.



Prompt Engineering - Zero-Shot Prompt Design

Concept: Zero-Shot prompting provides task instructions without examples. The model must infer the classification criteria purely from the system instructions.

Design Rationale: The prompt establishes model identity ("You are a sentiment analysis expert"), defines the exact output format, constrains valid categories and tags, and includes rules to prevent hallucination.

Category: Positive
Tags: Food Quality, Delivery Time, Overall Experience
Suggested Action: Highlight this feedback in local advertising to showcase the service's efficiency and speed.

Zero-Shot Prompt

You are a sentiment analysis expert for Uber Eats, a leading food delivery platform. Analyze customer reviews to help the business team respond effectively.

For each review provided, output:

1. Category: Classify as exactly one of: Positive, Negative, or Neutral
2. Tags: Select all applicable from: Delivery Time, Food Quality, Price, Packaging, Overall Experience
3. Suggested Action: Provide one specific recommendation for the business team

Output Format:
Category: [classification]
Tags: [comma-separated tags]
Suggested Action: [recommendation]

Rules:
- Base your analysis only on information explicitly stated in the review
- Do not invent or assume details not present
- If sentiment is mixed but leans one direction, classify by the dominant sentiment

Prompt Engineering - Few-Shot Prompt Design

Concept: Few-Shot prompting provides 2-4 example input-output pairs before the actual task. The model learns the expected pattern from these demonstrations.

Design Rationale: Three examples were included—one Positive, one Negative, and one Neutral—to calibrate the model's classification thresholds. Each example demonstrates the complete output format including tags and suggested actions.

Category: Positive
Tags: Delivery Time, Food Quality, Overall Experience
Suggested Action: Send a thank-you notification and suggest adding the restaurant to "Favorites" to facilitate quick future orders.

Few-Shot Prompt

You are a sentiment analysis expert for Uber Eats, a leading food delivery platform. Analyze customer reviews to help the business team respond effectively.

For each review, output:
1. Category: Exactly one of: Positive, Negative, or Neutral
2. Tags: Select applicable tags from: Delivery Time, Food Quality, Price, Packaging, Overall Experience
3. Suggested Action: One specific business recommendation

Example 1:
Review: "Delivery was super fast and the food was still hot! Will definitely order again."
Category: Positive
Tags: Delivery Time, Food Quality, Overall Experience
Suggested Action: Send thank-you message and offer loyalty discount to encourage repeat orders.

Example 2:
Review: "Waited over an hour for cold pizza. Packaging was crushed. Very disappointed."
Category: Negative
Tags: Delivery Time, Food Quality, Packaging
Suggested Action: Escalate to logistics team, issue refund, and flag delivery partner for review.

Example 3:
Review: "Food was okay. Nothing special but nothing wrong either.

Prompt Engineering - Chain-of-Thought Prompt Design

Concept: Chain-of-Thought prompting instructs the model to reason through the problem step-by-step before providing the final answer. This explicit reasoning process improves accuracy on complex classification tasks.

Design Rationale: The prompt defines a 5-step reasoning process: (1) Identify key phrases, (2) Assess sentiment indicators, (3) Identify relevant topics, (4) Determine overall sentiment, and (5) Recommend action. An example demonstrates the complete reasoning chain.

Customer_ID	City	Delivery_Time	Category	Tags	Suggested Action
C001	New York	28	Positive	Food Quality, Delivery Time, Overall Experience	Send appreciation message to reinforce customer loyalty.
C002	Los Angeles	95	Negative	Delivery Time, Food Quality, Overall Experience	Issue full refund and apologize for severe service failure.

Chain-of-Thought (CoT) Prompt

You are a sentiment analysis expert for Uber Eats. Analyze customer reviews by thinking through your reasoning step-by-step before providing the final classification.

For each review, follow this process:

Step 1 - Identify Key Phrases: List the emotionally significant words or phrases in the review.

Step 2 - Assess Sentiment Indicators: Determine whether each phrase suggests positive, negative, or neutral sentiment.

Step 3 - Identify Relevant Topics: Determine which aspects are mentioned (Delivery Time, Food Quality, Price, Packaging, Overall Experience).

Step 4 - Determine Overall Sentiment: Based on the balance of indicators, decide the dominant sentiment.

Step 5 - Recommend Action: Based on the category and tags, suggest what the business should do.

Final Output Format:
Category: [Positive/Negative/Neutral]
Tags: [comma-separated]
Suggested Action: [recommendation]

Example:

Review: "Driver was friendly but the food was completely wrong. Got

Generated Output & Quality

Great Learning
POWER AHEAD

Final Output Table

Using the Chain-of-Thought prompt (selected as the best-performing technique), all 100 customer reviews were processed and classified. The complete output table is provided as a separate PDF submission.

Sentiment Distribution

The 100 customer reviews were classified as follows:

Category	Count	Percentage
Positive	50	50%
Neutral	30	30%
Negative	20	20%

Interpretation

Half of all customers reported positive experiences, indicating a solid baseline of service quality. However, 20% of customers had negative experiences requiring immediate service recovery to prevent churn.

Comparative Evaluation of Prompting Techniques

- **Classification Accuracy:** All three techniques achieved identical accuracy (validated by manually reviewing 20 sample classifications against expected sentiment)
- **Tag Precision:** CoT captured nuanced tags (Driver Behavior, Customer Service) that Zero-Shot and Few-Shot missed
- **Action Quality:** CoT recommended full refunds for severe cases vs. partial refunds from other techniques

Recommendation: Chain-of-Thought is the optimal technique because:

- Superior tag precision for complaint routing
- More contextually appropriate business actions
- Provides audit trail through explicit reasoning documentation

Side-by-Side Comparison: Output Differences

Example Review: “Driver was rude and shoved the bag at me. Food was cold.”

Zero-Shot Output:

- Category: Negative | Tags: Delivery Time, Food Quality | Action: Apologize and offer discount

Few-Shot Output:

- Category: Negative | Tags: Delivery Time, Food Quality | Action: Issue partial refund

Chain-of-Thought Output:

- Category: Negative | Tags: Delivery Time, Food Quality, **Driver Behavior** | Action: Escalate to driver management + full refund

Key Difference: CoT’s step-by-step reasoning explicitly analyzed “driver was rude” as a distinct service dimension, capturing the “Driver Behavior” tag that other techniques missed. This enables proper complaint routing to the driver management team.

Key Insights & Business Recommendations

Key Insights:

- Delivery time is the strongest predictor of customer sentiment (Positive: 32.6 min avg, Neutral: 54.4 min avg, Negative: 80.7 min avg)
- Food Quality and Delivery Time appear in 97-98% of all reviews
- CoT's explicit reasoning catches nuances that other techniques miss

Business Recommendations:

- Implement proactive alerts for orders exceeding 60 minutes
- Invest in insulated packaging to mitigate cold food complaints
- Route driver conduct issues separately using CoT-identified tags
- Engage neutral customers (30%) with targeted campaigns to prevent churn

APPENDIX

Dataset Overview

- Dataset: 100 anonymized Uber Eats customer reviews
- Fields: Customer_ID, City, Delivery_Time (minutes), Review (text)
- Cities: 100 unique US cities
- Platform: Google AI Studio with Gemini 3 Pro Preview

Happy Learning !