

# Executive Summary

## Food Delivery Sentiment Analysis Using Generative AI Prompt Engineering

**John Buczkowski**  
December 2025

### Objective

The goal of this project was to evaluate how different prompt engineering strategies affect the accuracy, consistency, and usability of sentiment classification using Generative AI. The task involved classifying customer food-delivery reviews into Positive, Negative, or Neutral categories, assigning relevant business tags, and generating concise suggested actions. The focus was on determining which prompting approach best aligned with a predefined annotation logic.

### Approach

Four prompting strategies were designed, tuned, and evaluated: Zero-Shot, Few-Shot, Zero-Shot Chain-of-Thought (CoT), and Few-Shot Chain-of-Thought (CoT). A manually labeled Ground Truth dataset of 36 reviews was created to serve as a benchmark. Each prompt type was evaluated against the Ground Truth for category alignment, tag alignment, and consistency. The two best-performing prompts were then applied to the full dataset of 100 reviews.

### Key Findings

**Few-Shot prompting achieved the highest alignment with Ground Truth.** The optimized Few-Shot prompt achieved perfect category and tag alignment on the benchmark dataset.

**Chain-of-Thought did not improve results for this task.** CoT prompts added reasoning detail but did not improve accuracy and often introduced extra or unnecessary tags.

**Prompt structure mattered more than temperature.** Once prompts were well-defined and constrained, temperature changes had minimal impact on results.

**Few-Shot prompts were more stable and reliable.** Few-Shot prompts produced more consistent outputs and adhered more reliably to formatting constraints across repeated runs.

### Final Recommendation

Few-Shot prompting is recommended over Chain-of-Thought for rule-based sentiment classification tasks. It produces more consistent and precise labels, is easier to maintain, faster and cheaper to run, easier to debug, and best matched the Ground Truth annotation logic in this analysis.

### Business Value

This approach enables scalable and consistent sentiment analysis of customer feedback without manual review. It supports faster identification of delivery issues, highlights positive customer experiences for marketing, and provides actionable insights to operations teams while keeping implementation complexity and costs low.