Explainable Complaint Severity & Escalation Risk Modeling

Why this project
Organizations receive large volumes of guest feedback, yet not all negative comments represent the same level of operational risk. Traditional sentiment analysis captures overall positivity or negativity but does not distinguish between minor dissatisfaction and high-severity issues that may escalate into safety, service, or reputational concerns. This project addresses that gap by focusing on prioritization rather than averages.

What is being analyzed
The analysis uses only raw guest comments, reflecting real-world scenarios where downstream labels such as refunds, escalations, or revenue impact are unavailable. Linguistic and contextual signals are extracted from text, including sentiment intensity, emotional language, urgency indicators, and the level of specificity in complaints. These signals capture how guests express dissatisfaction, not just what they complain about.

How the analysis is performed
An explainable, weakly supervised modeling approach is applied. Interpretable NLP features are engineered from raw text and combined into a transparent proxy severity score that reflects escalation risk. Machine learning models are then trained to generalize these patterns across new comments while preserving interpretability. Model outputs are validated through qualitative review and feature importance analysis to ensure alignment with business logic.

What problem this solves
This project enables organizations to move from sentiment monitoring to operational triage by identifying which complaints require immediate attention. It supports early risk detection, prioritization of response efforts, and explainable decision-making for operations teams. Ultimately, the analysis shifts the question from “What are guests unhappy about?” to “Which complaints matter most right now, and why?”