

ANALYSIS OF HOTEL REVIEWS

Transforming Customer Feedback to Actionable
Business Insights Using AI



TABLE OF CONTENT

01 PROJECT OVERVIEW

03 INSIGHT & FINDINGS

05 AI SUPPORT EXPLANATION

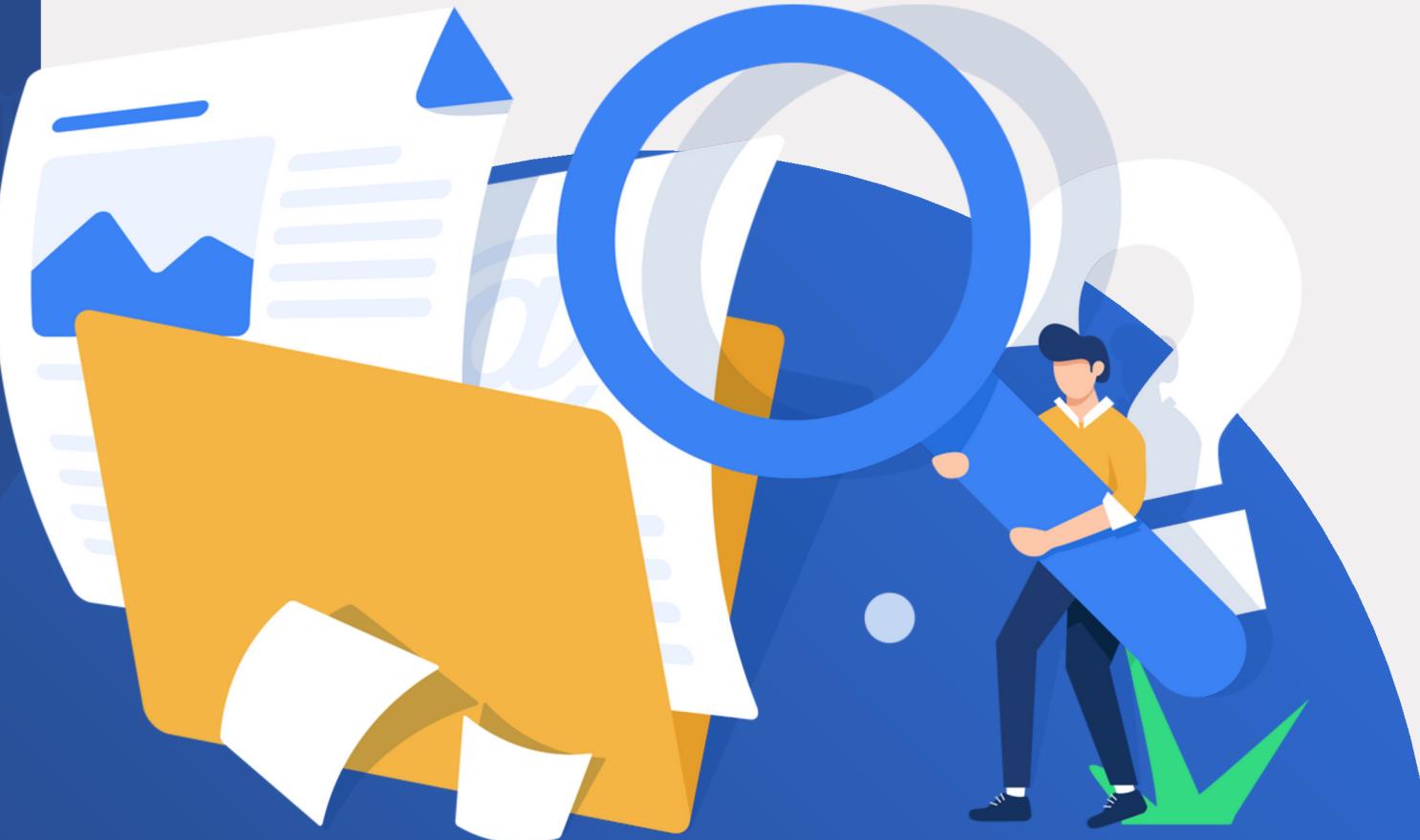
02 ANALYSIS PROCESS

04 CONCLUSION &
RECOMMENDATI
ONS

06 CLOSING



PROJECT OVERVIEW



Online customer reviews are a valuable source of insight, yet their high volume and unstructured format present a significant challenge for manual analysis. This project aims to develop an automated data analysis workflow that transforms thousands of qualitative reviews into a dynamic business intelligence dashboard. The ultimate goal is to deliver actionable recommendations to hotel management, enabling them to improve guest satisfaction and drive operational excellence.

 [Dataset link](#)

ANALYSIS PROCESS

► PREPROCESSING & EDA

Each review is professionally cleaned. The system then performs sarcasm detection and aspect (topic) extraction using rule-based linguistic patterns to ensure maximum accuracy.

► SENTIMENT ANALYSIS

A state-of-the-art AI model (RoBERTa) determines the baseline sentiment (Positive, Negative, Neutral). This output is then automatically adjusted if sarcasm was detected in the previous step.

► CONTEXT-AWARE SMART SUMMARIZATION

An AI Summarization model (BART-Large) generates a summary for each review. This process is context-aware, enriching its prompts with the outputs from the sentiment and sarcasm analyses.

► QUALITY VALIDATION & FALBACK

Every AI-generated summary is validated for quality. If a summary is deemed subpar (e.g., too generic or short), the system automatically switches to an extractive summarization method as a safety net.

► SYNTHESIS & VISUALIZATION

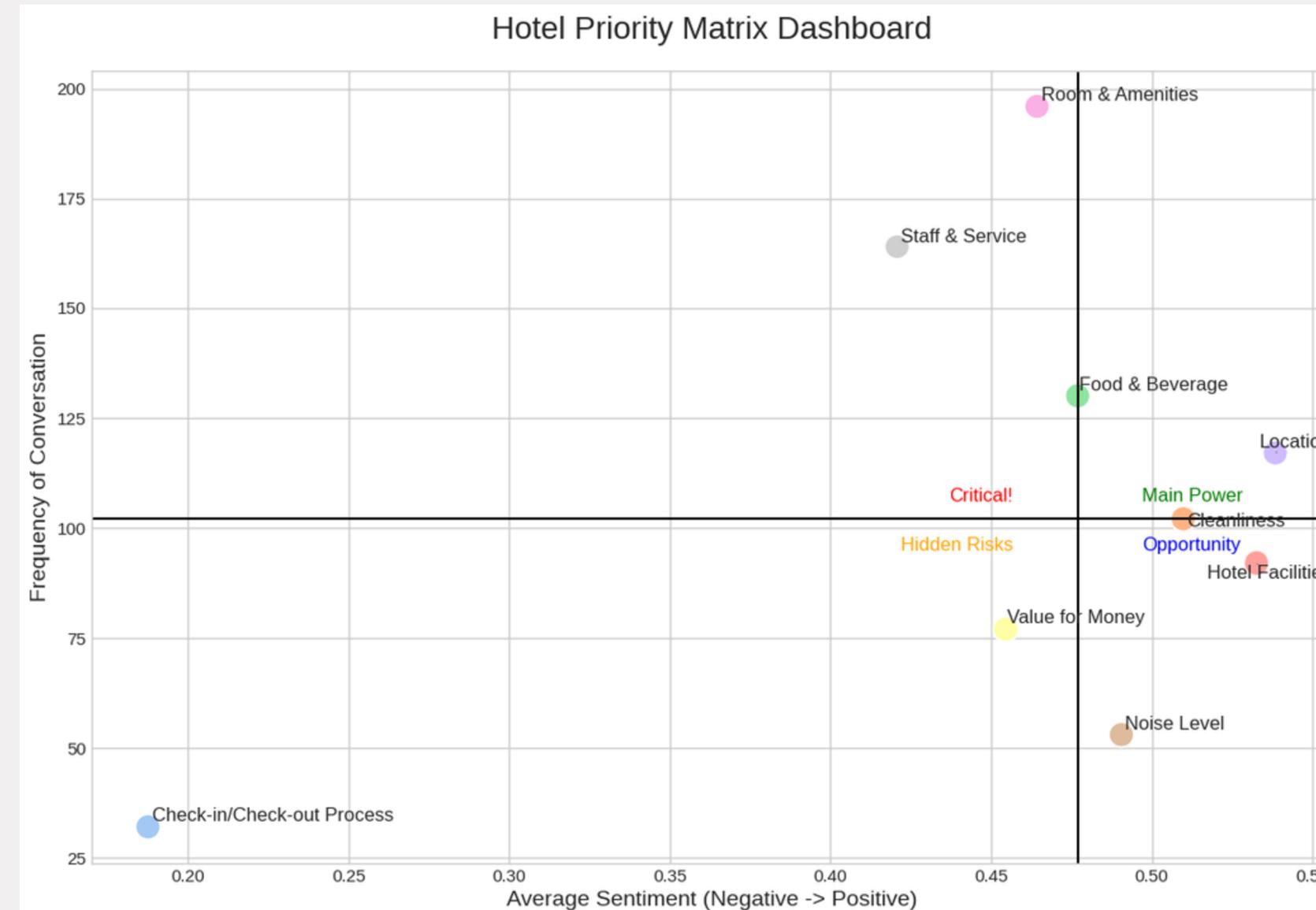
All validated analysis results are then consolidated to create the Priority Matrix Dashboard.



INSIGHTS & FINDINGS



Our central finding is visualized in the Priority Matrix, which maps key topics by their frequency of mention and average sentiment. This reveals four strategic quadrants:



INSIGHTS & FINDINGS



CORE STRENGTHS (HIGH FREQUENCY, HIGH POSITIVITY):

Location, Hotel Facilities, and Cleanliness are the pillars of the hotel's positive reputation, indicating high and consistent standards that should be maintained and highlighted.



CRITICAL ISSUES (HIGH FREQUENCY, HIGH NEGATIVITY):

Room & Amenities and Staff & Service are the most high-impact areas detracting from the guest experience. Their high frequency and relatively low sentiment scores make them the top operational priority to address.



HIDDEN RISKS (LOW FREQUENCY, HIGH NEGATIVITY):

Check-in/Check-out Process is the clearest "red flag." Although infrequently mentioned, it has the lowest sentiment score of all, indicating a fundamental issue that can be highly frustrating for guests.



OPPORTUNITIES (LOW FREQUENCY, FAVORABLE SENTIMENT):

Topics like Value for Money and Noise Level show favorable sentiment when discussed. There is potential to improve perception in these areas and elevate them into new strengths.

CONCLUSION & RECOMMENDATIONS



CONCLUSION

The hotel's strong foundation, built on its Location, Facilities, and Cleanliness, is currently undermined by critical issues in Room & Amenities and Staff & Service. Furthermore, a significant hidden risk in the Check-in/Check-out Process threatens to detract from an otherwise positive experience.



RECOMMENDATIONS

- Prioritize Room Renovation
Focus resources on renovating rooms and implementing staff retraining programs to address the two most critical issues.
- Leverage Location in Marketing
Make the hotel's prime Location and high standards of Cleanliness the central theme of marketing campaigns.
- Implement AI-Powered Reporting
Integrate the AI workflow to create a "daily digest" of negative reviews, allowing managers to resolve guest issues in near real-time.

AI SUPPORT EXPLANATION



Built within a structured Python Class, the system combines the strength of Rule-Based linguistic reasoning with the contextual understanding capabilities of Large Language Models (LLMs).

01

MULTI-MODEL ENGINE

Utilizes a robust multi-model architecture. For summarization, it uses `facebook/bart-large-cnn` with a fallback to a secondary model for stability. For sentiment analysis, it employs `nlptown/bert-base-multilingual-uncased-sentiment`, a model specifically trained to classify reviews on a 1-to-5-star scale.

02

RULE-BASED LINGUISTIC LOGIC

Sarcasm detection and topic extraction are not fully delegated to the AI. Instead, they use Regex rules and keyword dictionaries for maximum precision and control.

03

CONTEXT-AWARE & VALIDATED GENERATION

The summarization process is context-aware. Outputs from the sentiment and sarcasm analysis are used as inputs to enrich the AI prompts. The resulting summaries are then validated for quality before being presented, with an automatic fallback mechanism to guarantee reliability.



THANK YOU

For further discussion or collaboration, feel free to reach out.

 ahmadihsan506@gmail.com

 LinkedIn

 GitHub

