HOTEL GUEST REVIEW SENTIMENT ANALYZER

AUTOMATICALLY DETECT POSITIVE AND NEGATIVE EXPERIENCES AND UNDERSTANDS RECCURING COMPLAINTS.

TEAM MEMBERS

- AP23110010022 - K Meghana

- AP23110010190 - G Pushpa

- AP23110010503 - M Reshma

- AP23110010725 - M Vyshnavi

- AP23110011462 - G Mohana krishna

PROBLEM STATEMENT

Hotels receive a large number of guest reviews daily

Manual analysis of these reviews is slow and inefficient.

Staff cannot quickly identify whether feedback is positive ,negative, or neutral.

Recurring issues (cleanliness, food, staff behavior, amenities)often go unnoticed.

Lack of timely insights affects decision-making and service improvement

OBJECTIVES

- Automatically identify sentiment (positive/negative/neutral) from hotel guest reviews.

- Detect recurring issues such as cleanliness, food, staff behavior , and amenities.

- Provide quick and accurate insights to support hotel management decisions.

- Reduce manual effort in reading and analyzing large volumes of feedback.

- Improve guest satisfaction through data-driven service improvements.

SYSTEM WORK FLOW
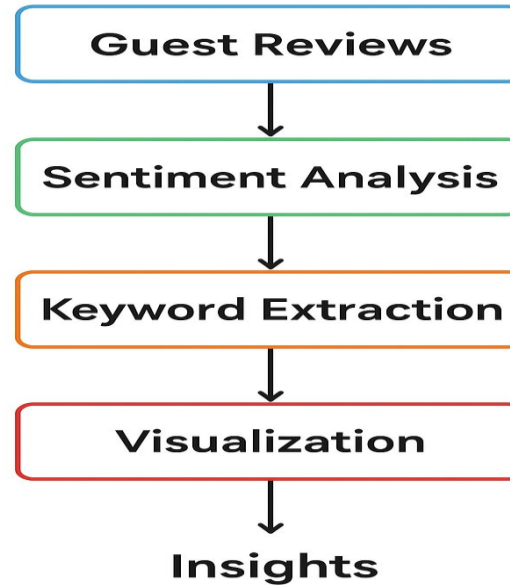
DATA &DEEP LEARNING APPROACH

DATASET OVERVIEW

•SOURCE: COMPREHENSIVE TRIPADVISOR HOTEL REVIEWS.
•SIZE: A ROBUST DATASET OF 20,491 REVIEWS.
•FEATURES: RAW REVIEW TEXT AND ASSOCIATED 1–5 STAR RATINGS.
•LABELS : BINARY CLASSIFICATION (POSITIVE: RATING > 3, NEGATIVE: RATING<3).

 METHODOLOGY: FROM TEXT TO PREDICTION

•LOWERCASE CONVERSION
•SPECIAL CHARACTER REMOVAL
•STOP WORD REMOVAL (PRESERVING NEGATIONS)
•TOKENIZATION (10,000 VOCABULARY)
•SEQUENCE PADDING

DATA &DEEP LEARNING APPROACH

- Model Architecture:

- Embedding Layer (128 dimensions)

- Bidirectional GRU (128 units)

- Global Average Pooling• Dropout (0.3) for regularization

- Dense Layer (64 units, ReLU activation)

- Output Layer (Sigmoid activation for binary classification)

- Training Regimen:

- Data Split: 70% Training, 30% Testing

- Optimizer: Adam for efficient gradient descent

- Loss Function: Binary Cross entropy

- Regularization: Early Stopping to prevent overfitting

Our Tech Stack: Powering the Application

Frontend

React 19
Vite
Lucide React
Modern CSS (responsive design)

Backend

python 3.10+
Fast API
SQL Alchemy ORM
SQLite Database

Development Tools

Google colab
VS code
Git& GitHub

Machine Learning

TensorFlow / Keras
NLTK
Pandas & NumPy
Scikit-learn

PERFORMANCE & INSIGHTS MODEL PERFORMANCE METRICS

KEY FINDINGS:

• BIDIRECTIONAL GRU EFFECTIVELY CAPTURES CONTEXTUAL NUANCES FROM BOTH TEXT DIRECTIONS.

• STRATEGIC PRESERVATION OF NEGATION WORDS IMPROVES ACCURACY.

• A 10,000-WORD VOCABULARY ENSURES COVERAGE WITHOUT OVERFITTING.

• STRONG PERFORMANCE ON BOTH SHORT AND LONG REVIEWS.

THANK YOU