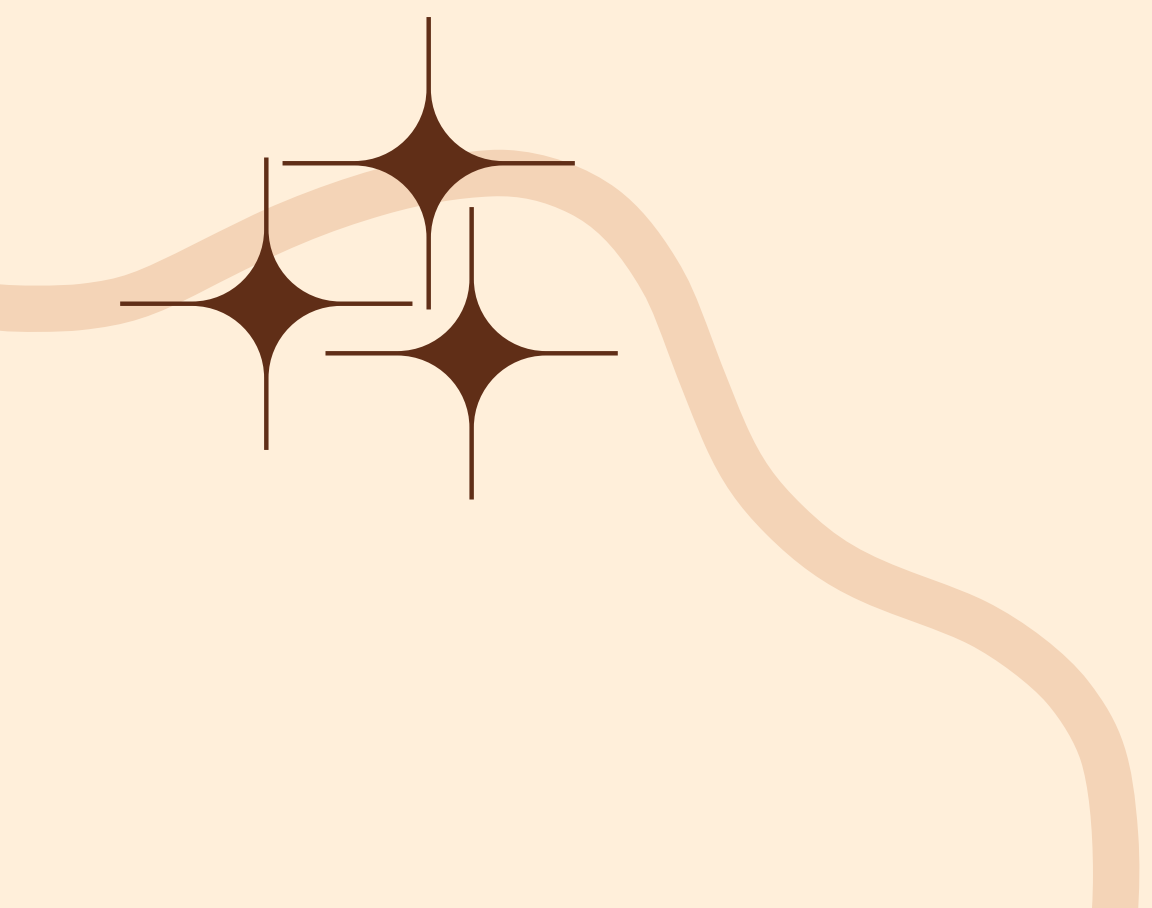





PROJECT: MULTIMODAL PRODUCT REVIEW ANALYZER



Maryam Ihsan
Muhammad Haaris
Zuha Aqib
Maham Junaid





THE PROBLEM

Businesses Are Drowning in Reviews



OUR SOLUTION

Input → AI Engine → Insights

MACHINE LEARNING MILESTONES 1

- **ML** REVIEW PROCESSING
PIPELINE

WHAT ARE WE ANALYZING?

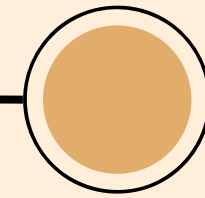
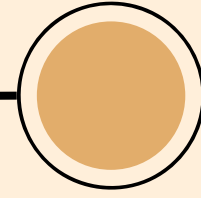
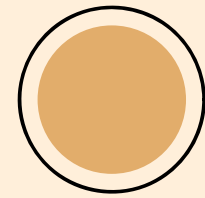
INPUTS

- Text reviews, star ratings (1-5), customer-uploaded images, structured metadata,

OUTPUTS

- Sentiment scores, trend visualizations, flagged issues, Brand/Product line classification

PROCESS



Sentiment Analysis (NLP): Use pre-trained BERT/VADER for polarity (positive/negative/neutral) on text; fine-tune on labeled subsets.

Trends & Issues: Topic modeling to extract themes (e.g., "battery life" issues); time-series analysis on ratings for trends (e.g., declining sentiment post-update).

Weighted sentiment score =
 $f(\text{text_sentiment}, \text{rating})$; classify issues (e.g., Random Forest on keywords/ratings).

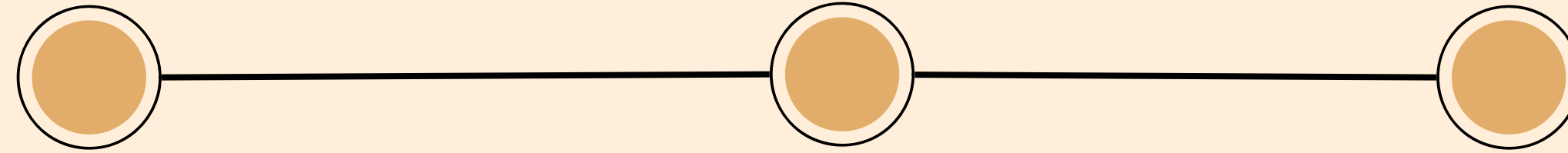
DIRECTING OUR ACHIEVABLES

This processed output creates a refined knowledge base, enabling our LLM in Milestone 2 to generate accurate, relevant, and brand-specific summaries and answers

LLM PART

- An LLM with retrieval-augmented generation produces balanced, product-specific pros/cons summaries grounded in real reviews.
- There is also text generation which provides a suitable response to the user based on what the RAG provides.

WHY DATASET MATTERS

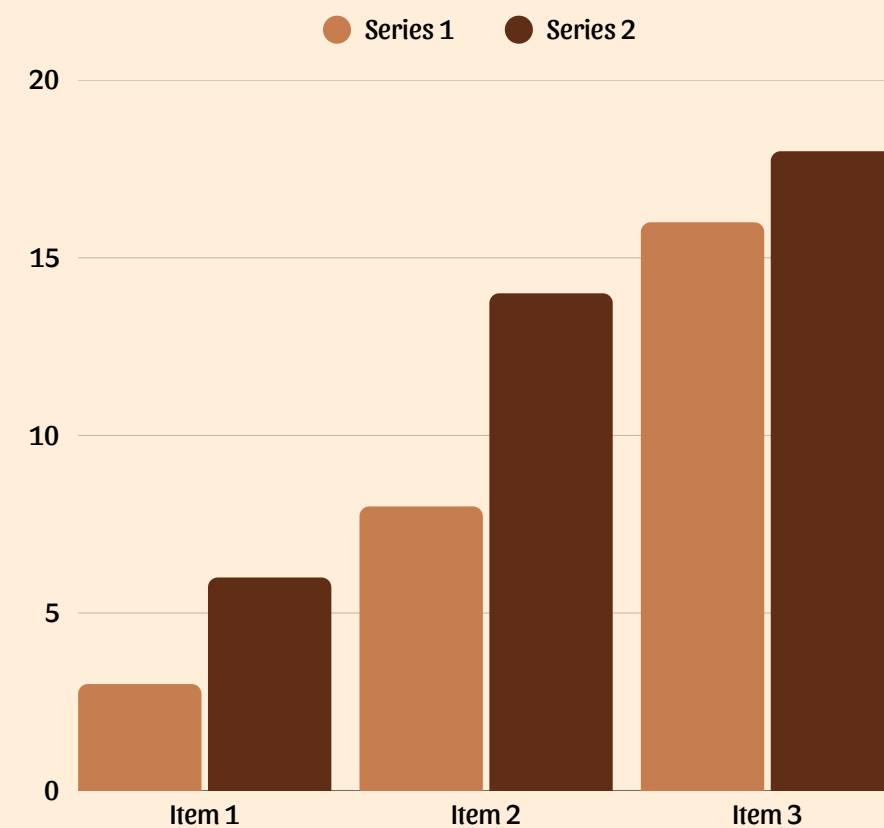


Rich, diverse review data is the backbone of our project.

Text reviews → for sentiment + aspect extraction.

Ratings & metadata (brand, category, model) → for filtering by product/domain.

AMAZON REVIEWS DATASET (UCSD, JULIAN MCAULEY)



Scale: Millions of product reviews across multiple domains.

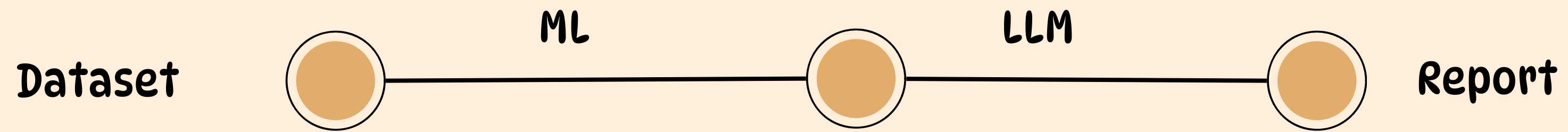
Features:

- Review text + star rating.
- Product metadata (brand, category, model).
- Images (for some products).

Why it fits our project:

- Covers laptops, smartphones, and more.
- Provides structured fields for brand filtering.
- Rich text for sentiment + LLM summarization.

HOW IT FITS OUR PIPELINE



Step 1 (ML):

- Filter reviews by category/brand (Dell, HP, etc.).
- Sentiment + aspect analysis from text + ratings.

Step 2 (LLM):

- RAG-powered pros/cons summaries grounded in reviews.

THANK
YOU

