Lavazza Project

Table of Contents

Checkpoint 2

- Value Proposition
- Project Goal
- Data Overview
- Research Questions
- Pipelines
- Models
- Evaluation strategy

Project Value Proposition

For **marketers** seeking deeper **insights** on product launches, our **analytics software** transforms raw online comments and reviews into **actionable metrics**.

Project Goal

Developing an **application software** in order to understand the **real impact of a product launch**, in terms of **consumers sentiment** crossing social media **comments** and third-party **reviews**.

Data Overview

Data sources:

- **Wonderflow**
- **Digiming**
- **Instagram** and **Youtube**

Data Overview

568 Comments	**0%** Labelled Comments	**72** Tiny Eco Comments
4960 Reviews	**95%** Labelled Reviews	**130** Tiny Eco Reviews

Unbalanced and **not fully labelled** dataset

Data Overview

568 Comments	**0%** Labelled Comments	**72** Tiny Eco Comments
4960 Reviews	**95%** Labelled Reviews	**130** Tiny Eco Reviews

Unbalanced and **not fully labelled** dataset

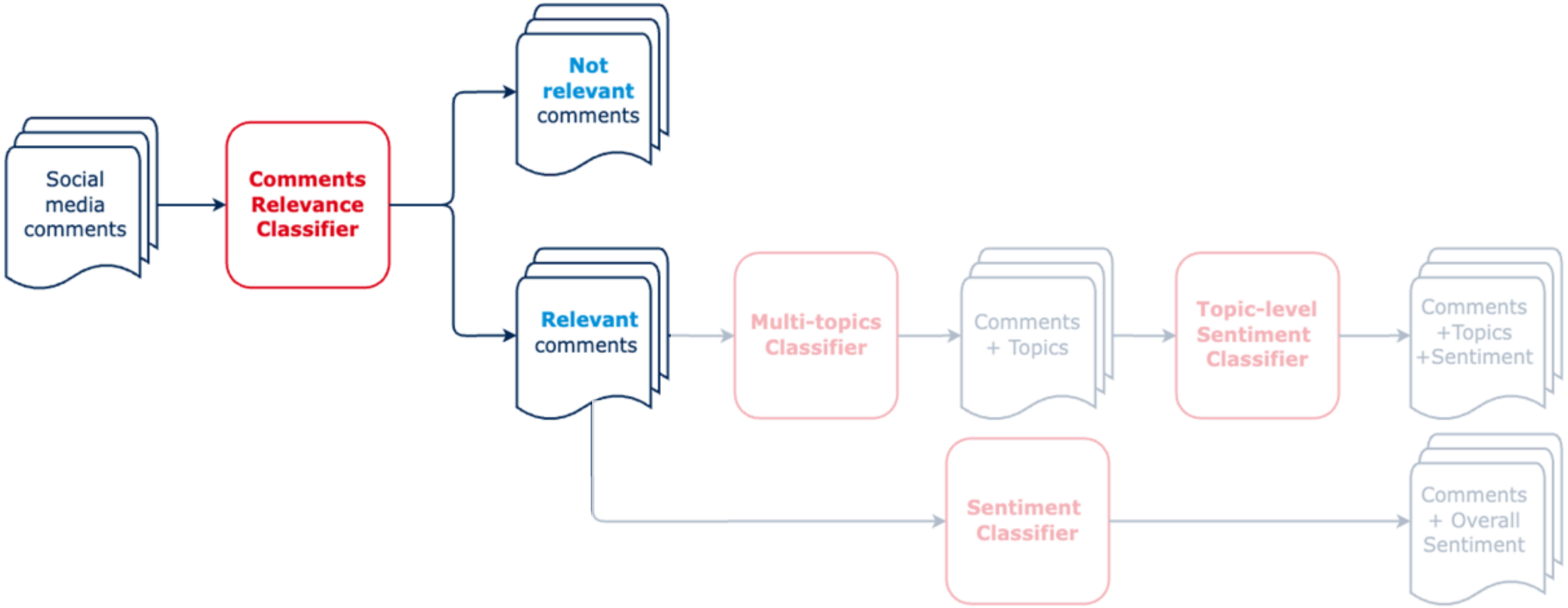
Research questions

Can **Large Language Models** be used to extract useful insights from customer comments and reviews to improve marketing strategies?

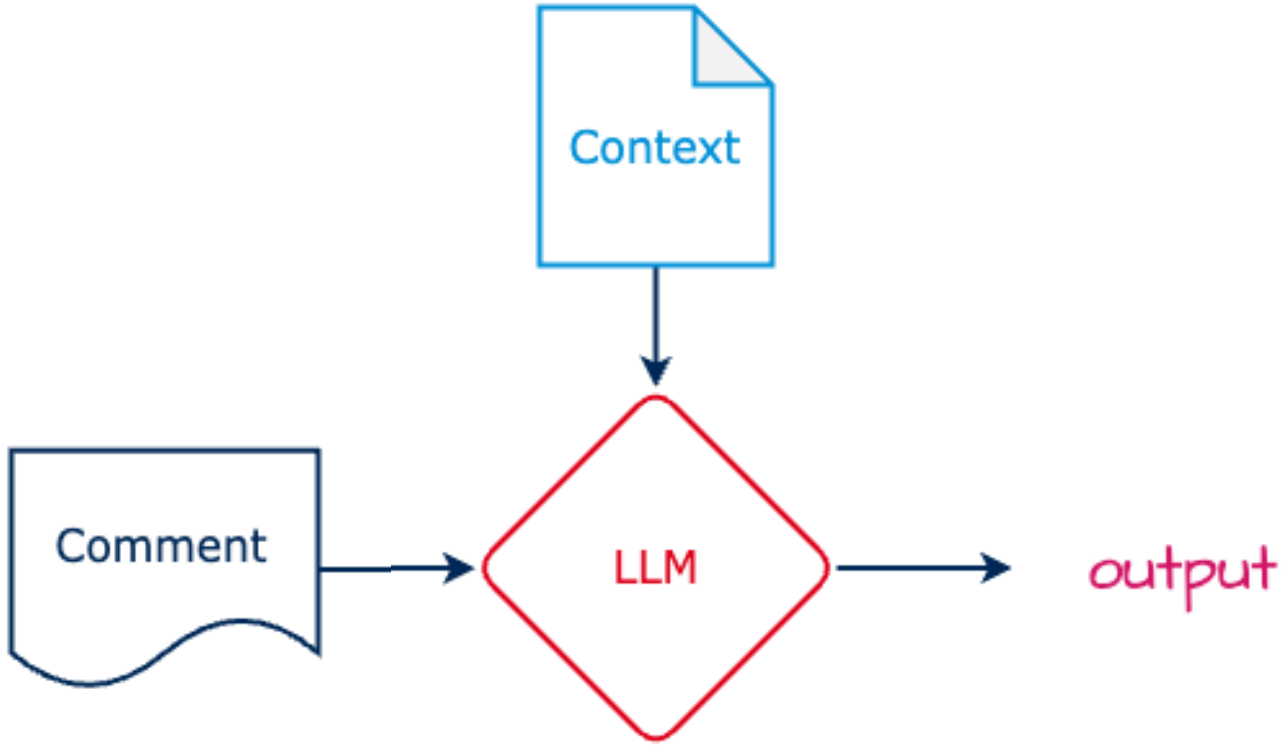
In particular:

1. How effectively can **LLMs** identify **product-relevant comments**?
2. To what extent can **LLMs** extract specific **product aspects** from reviews and social media comments?
3. How accurate are **LLMs** as **sentiment classifiers** when analyzing product-related aspects?

Classification Pipelines - Comments

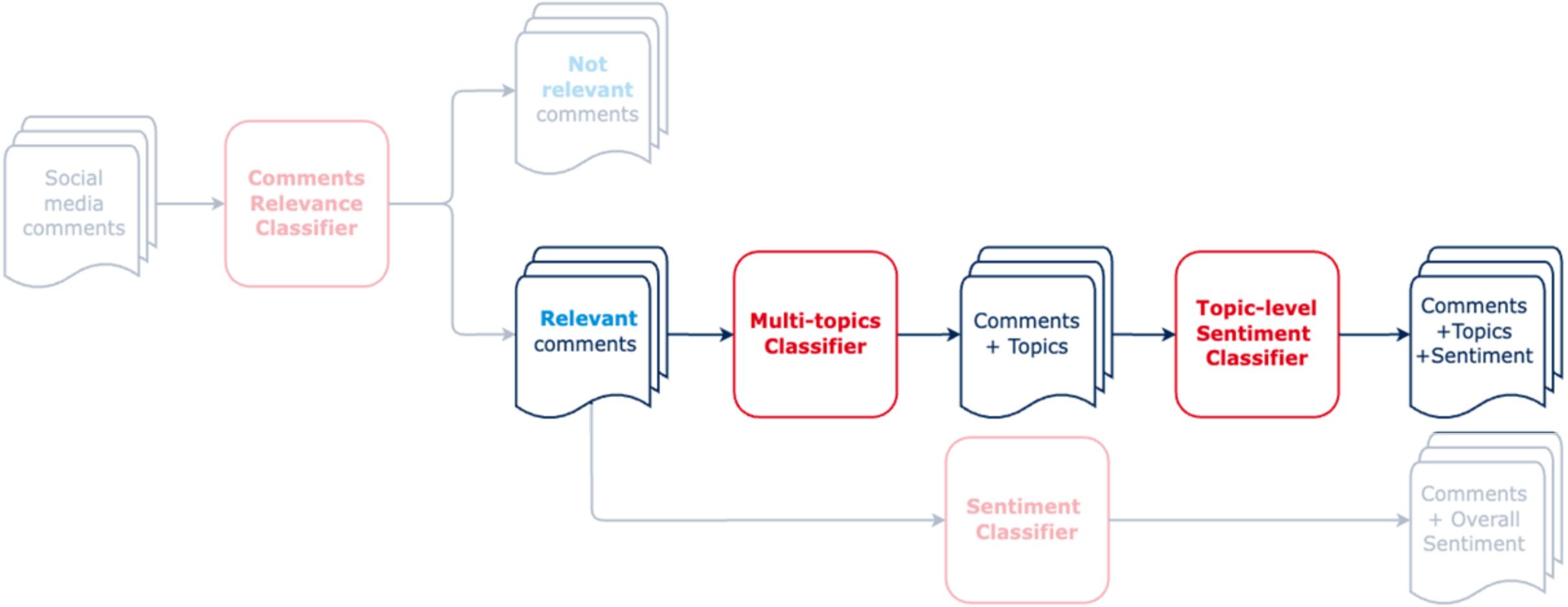


Models - Comments Relevance Classifier

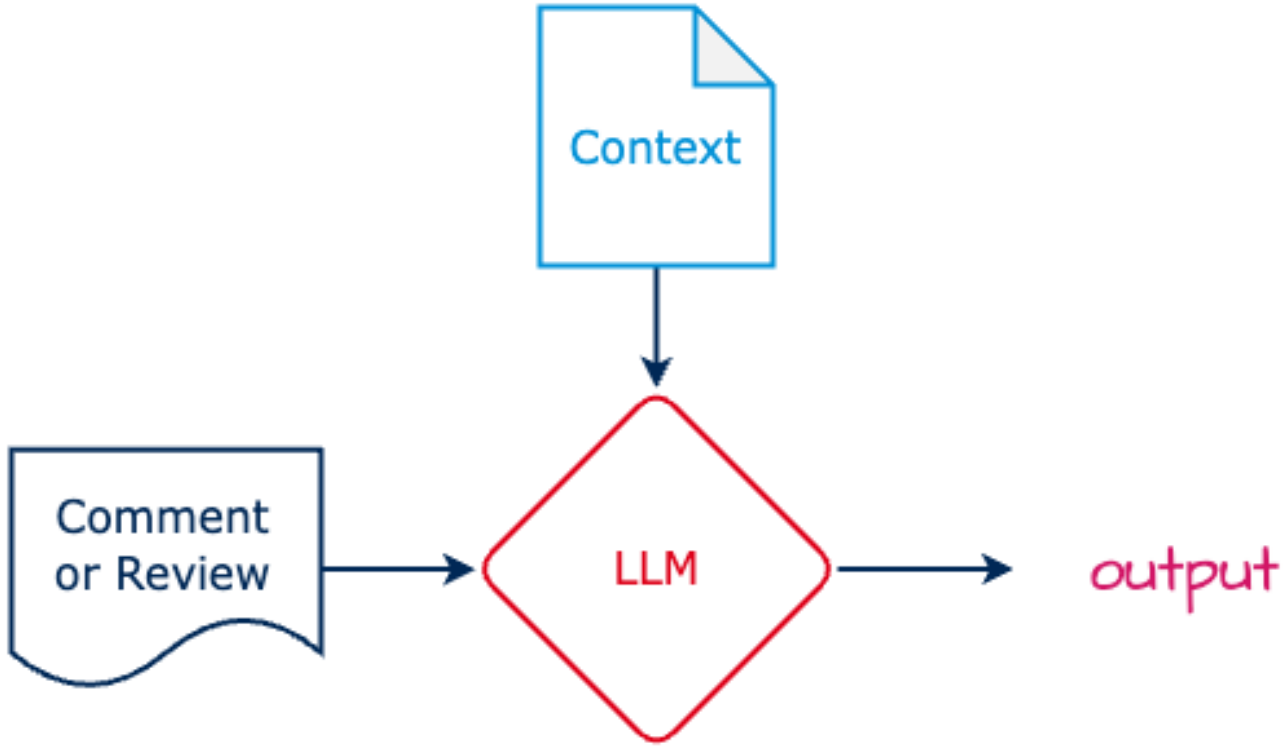


- **Context**: Detailed description of the product to which the comment should refer and few examples
- **Output**: Relevance label

Classification Pipelines - Comments

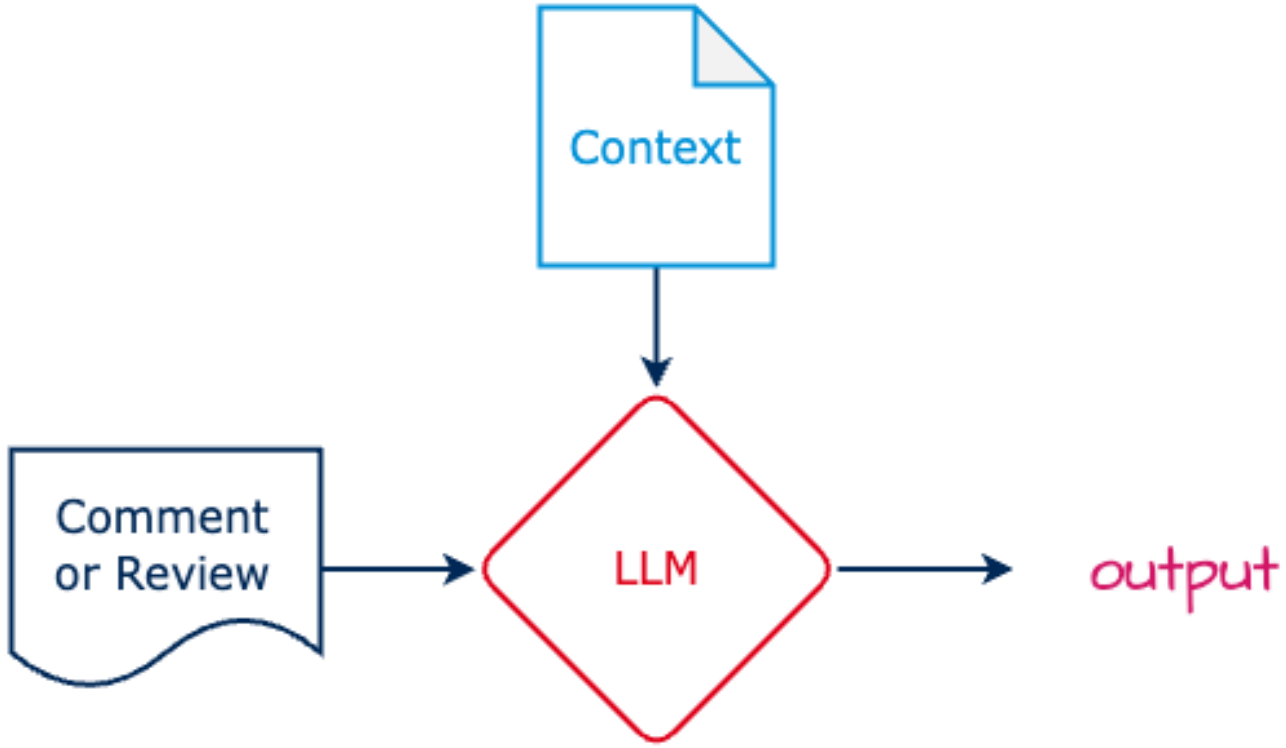


Models - Multi-topics Classifier



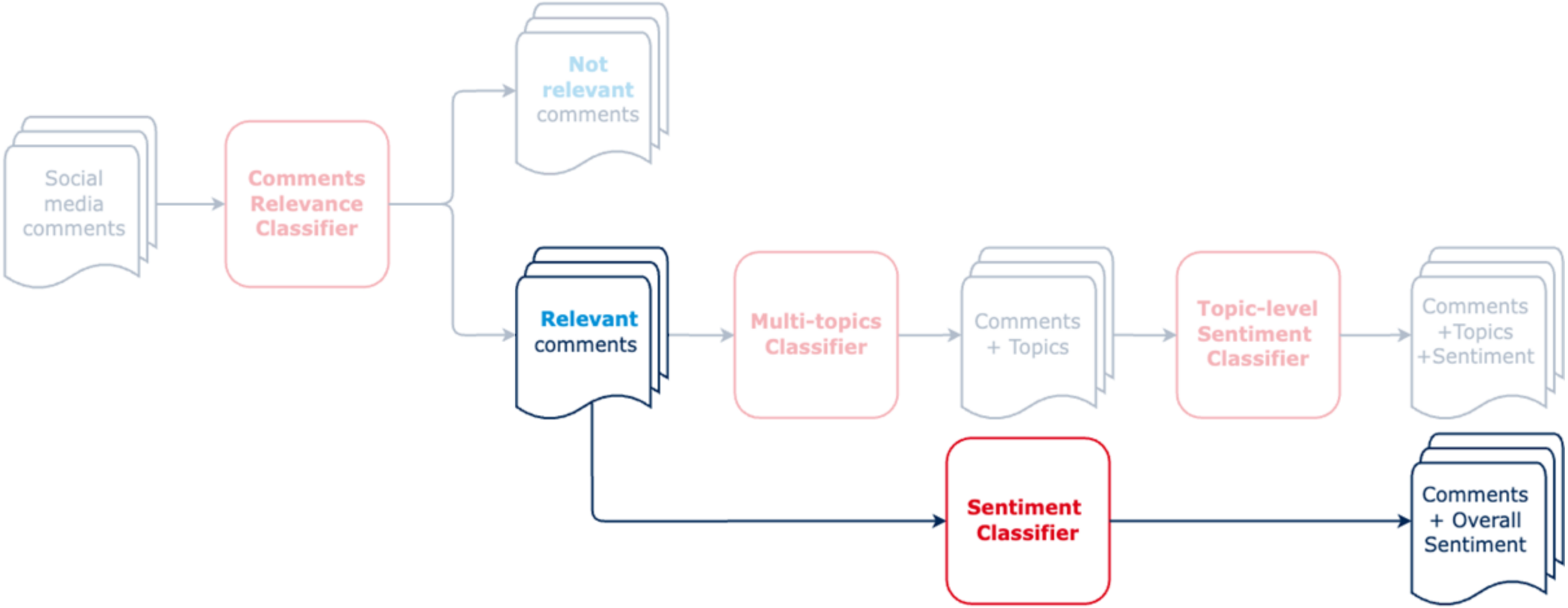
- **Context**: List of all the topics (product aspects) and few examples
- **Output**: List of topics related to the input comment

Models - Topic-level Sentiment Classifier

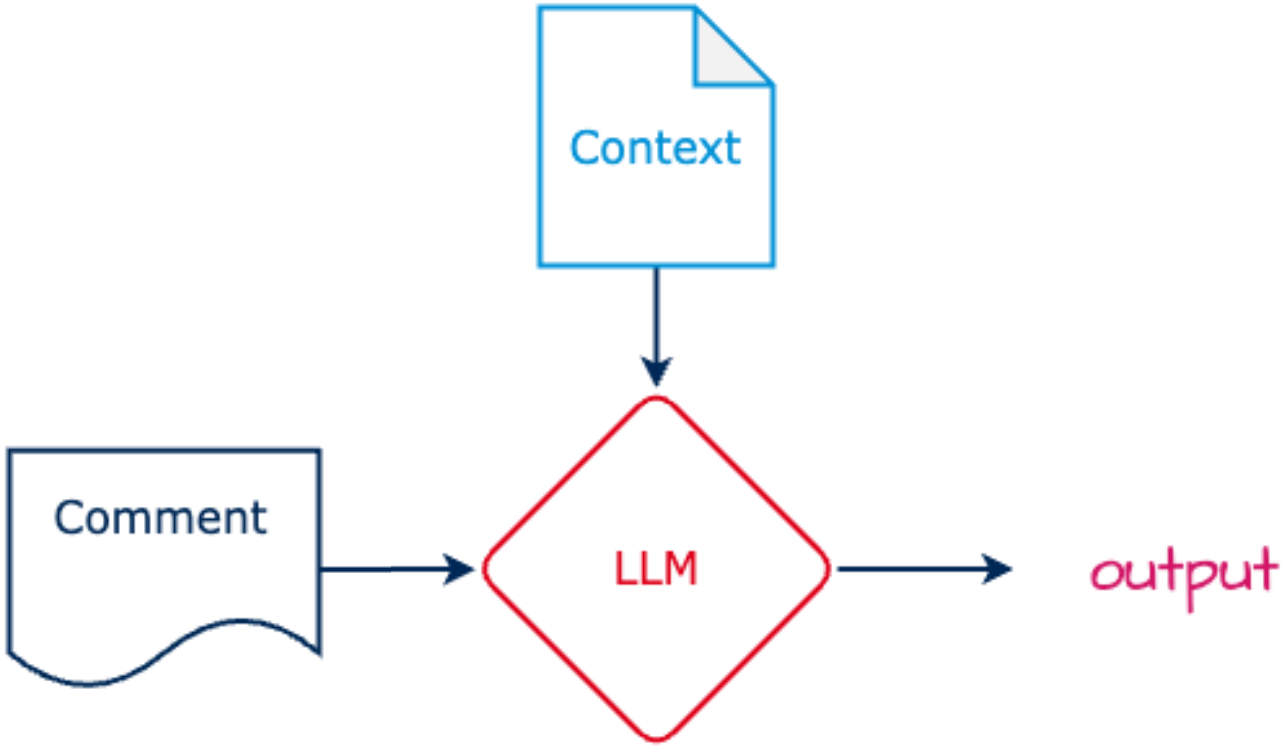


- **Context**: List of the topics related to the specific comment and few examples
- **Output**: List of topic-level sentiments

Classification Pipelines - Comments

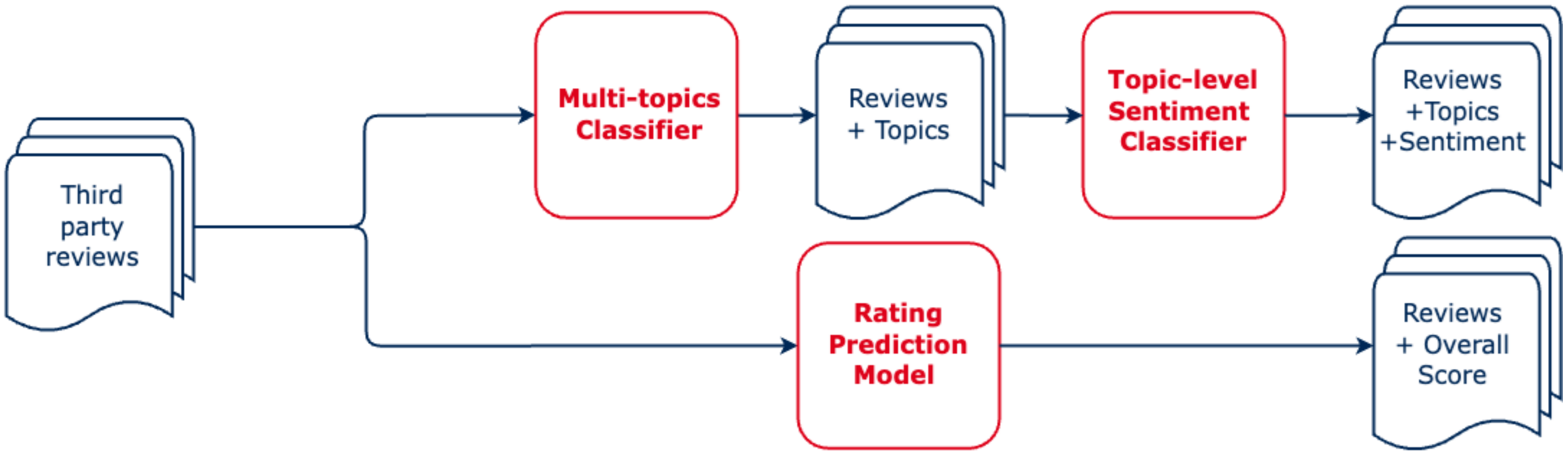


Models - Sentiment Classifier

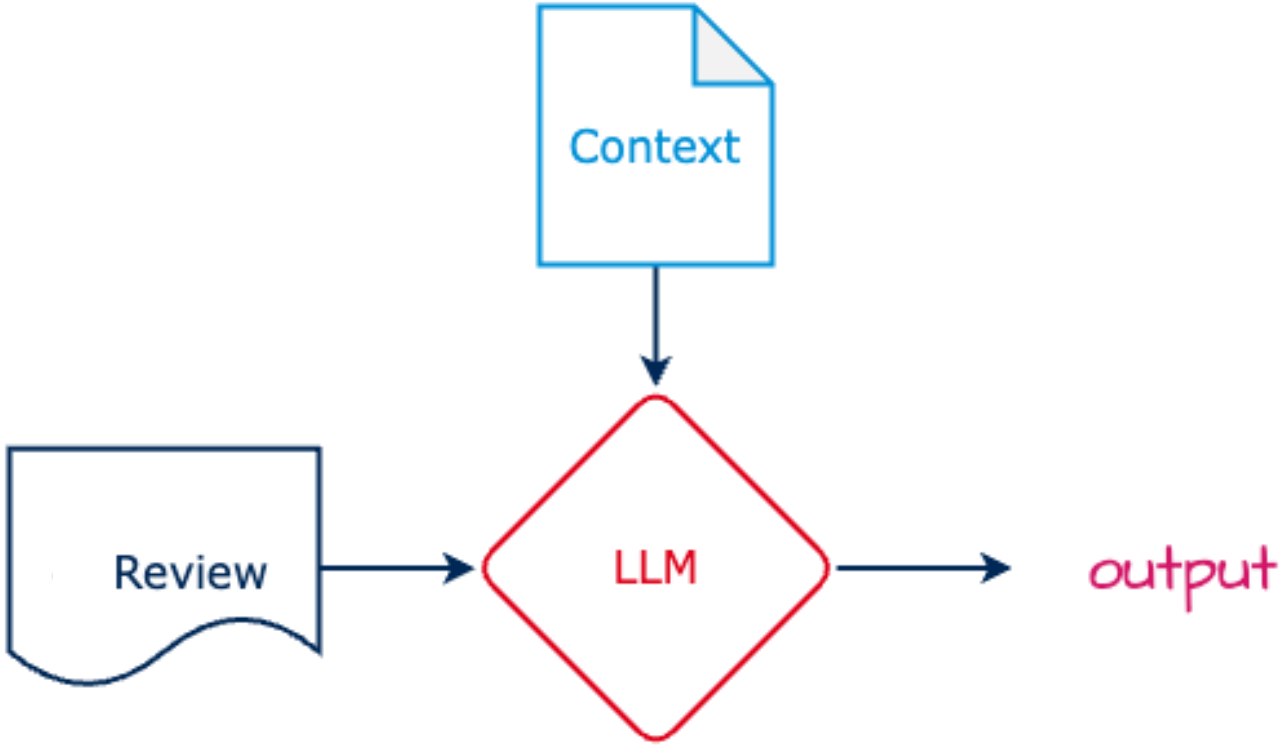


- **Context**: few positive and negative examples
- **Output**: Overall sentiment of the comment

Classification Pipelines - Reviews

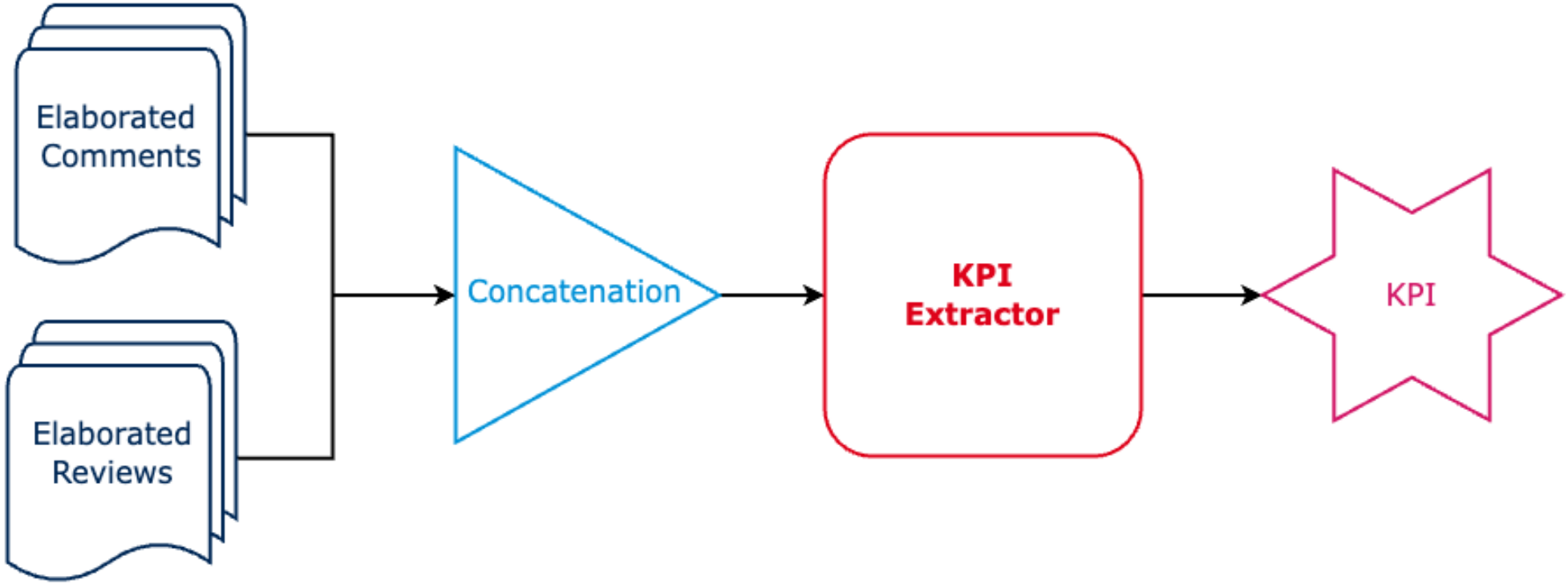


Models - Rating Prediction Model



- **Context**: Brief description of the stars and examples of rated reviews
- **Output**: Star rating for each review

KPI Extraction Pipelines



LLM

Google DeepMind

Gemma2: large language model from Google, different versions: 2B, 9B and 27B

Meta

Llama3: large language model from Meta, 8B, 70B and 405B

Evaluation strategy

- **Quantitative:**

  Performance of **classification models** and **topic modeling tasks** via leveraging synthetic data and our data

- **Qualitative:**

  Alignment with marketing objectives via meetings with stakeholders

Thank you

Alessio Gioè - Catalano Vincenzo - Tommaso Mazzarini