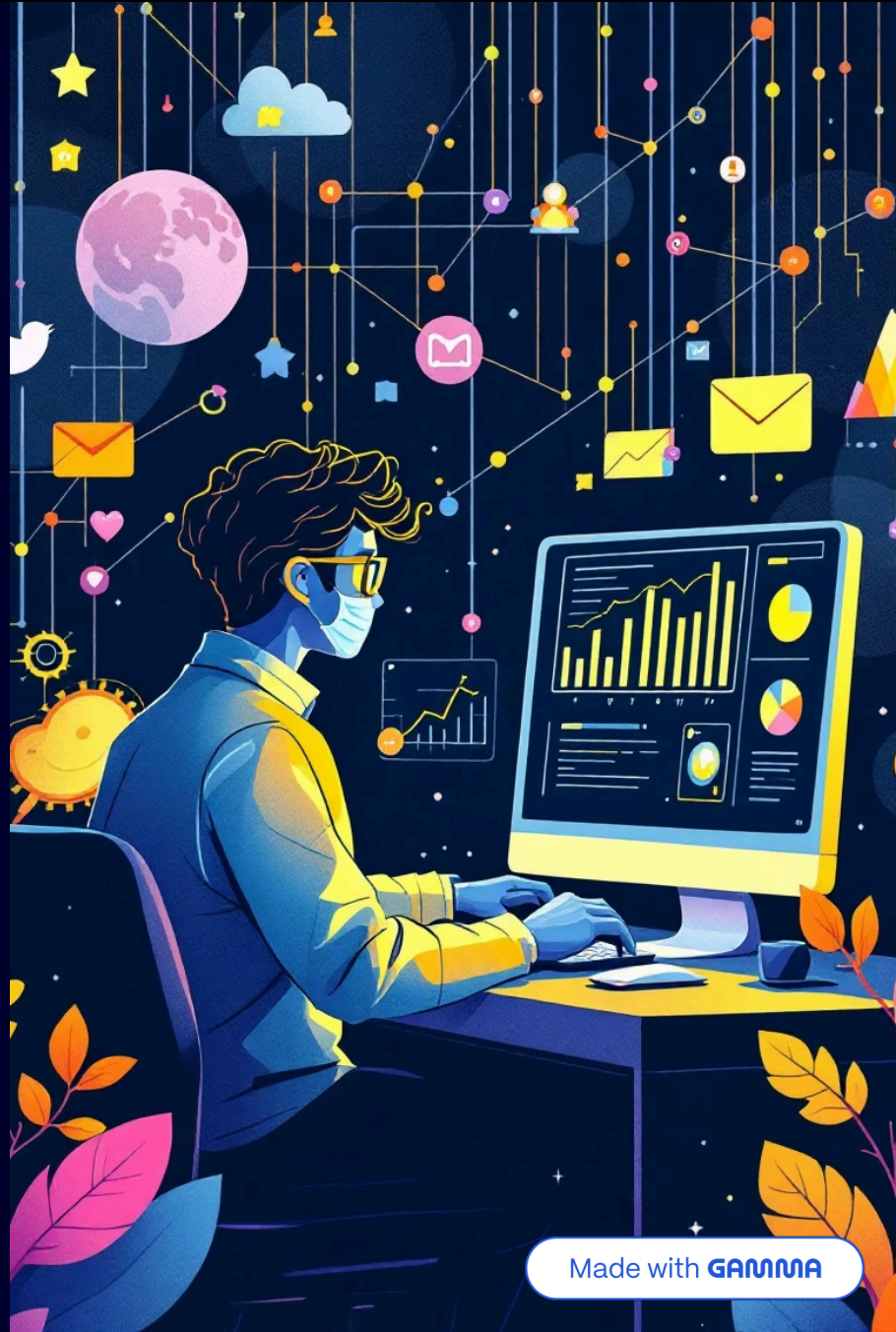
Project SocioScope

Aspect-Based Sentiment Analysis for Socio-Economic Tweets

Students: Ofir Fichman, Pavel Fadeev, Israel Peled



Motivating Use Case: The Pulse of Public Opinion

Why It Matters?

Understanding real-time public reactions to socio-economic policies helps guide better decisions and prevent unrest, offering faster, large-scale insights than traditional surveys.

Defining the Project Task

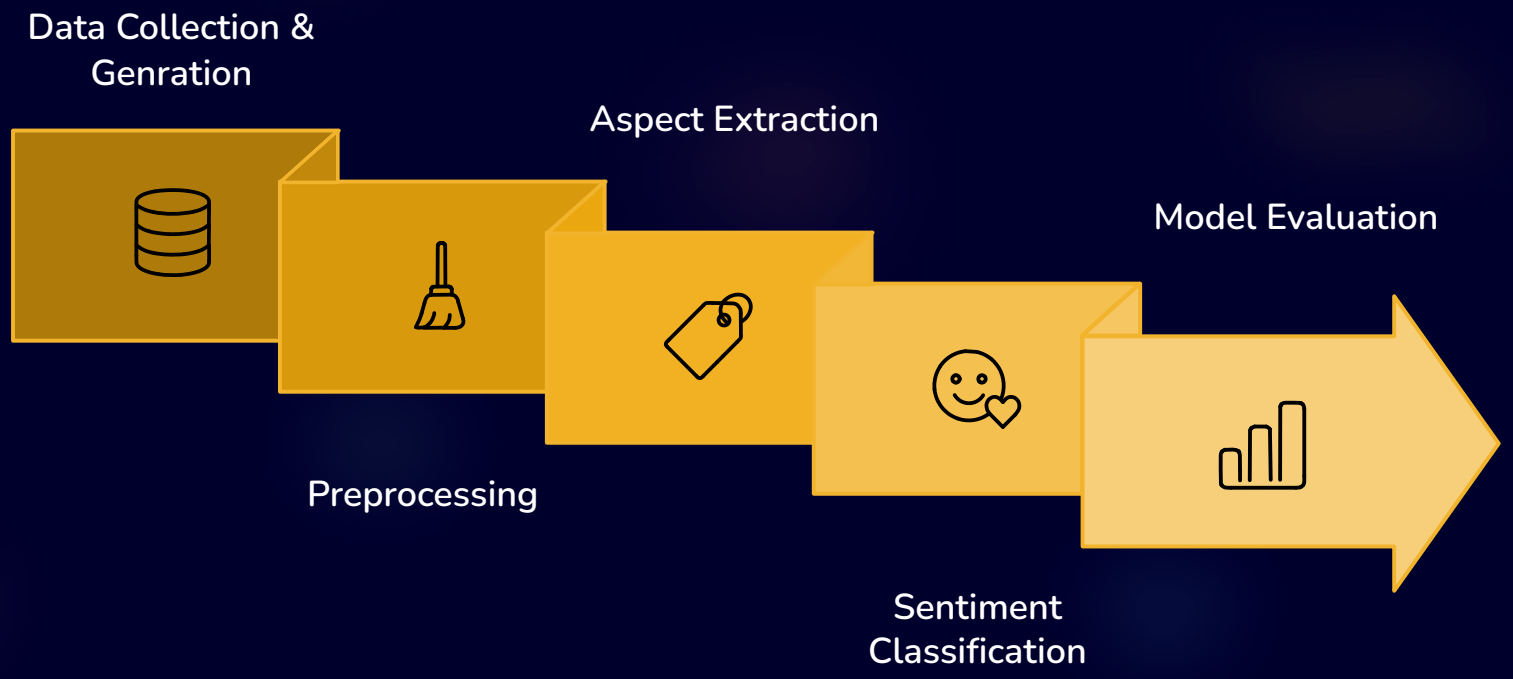
Formal Problem Statement

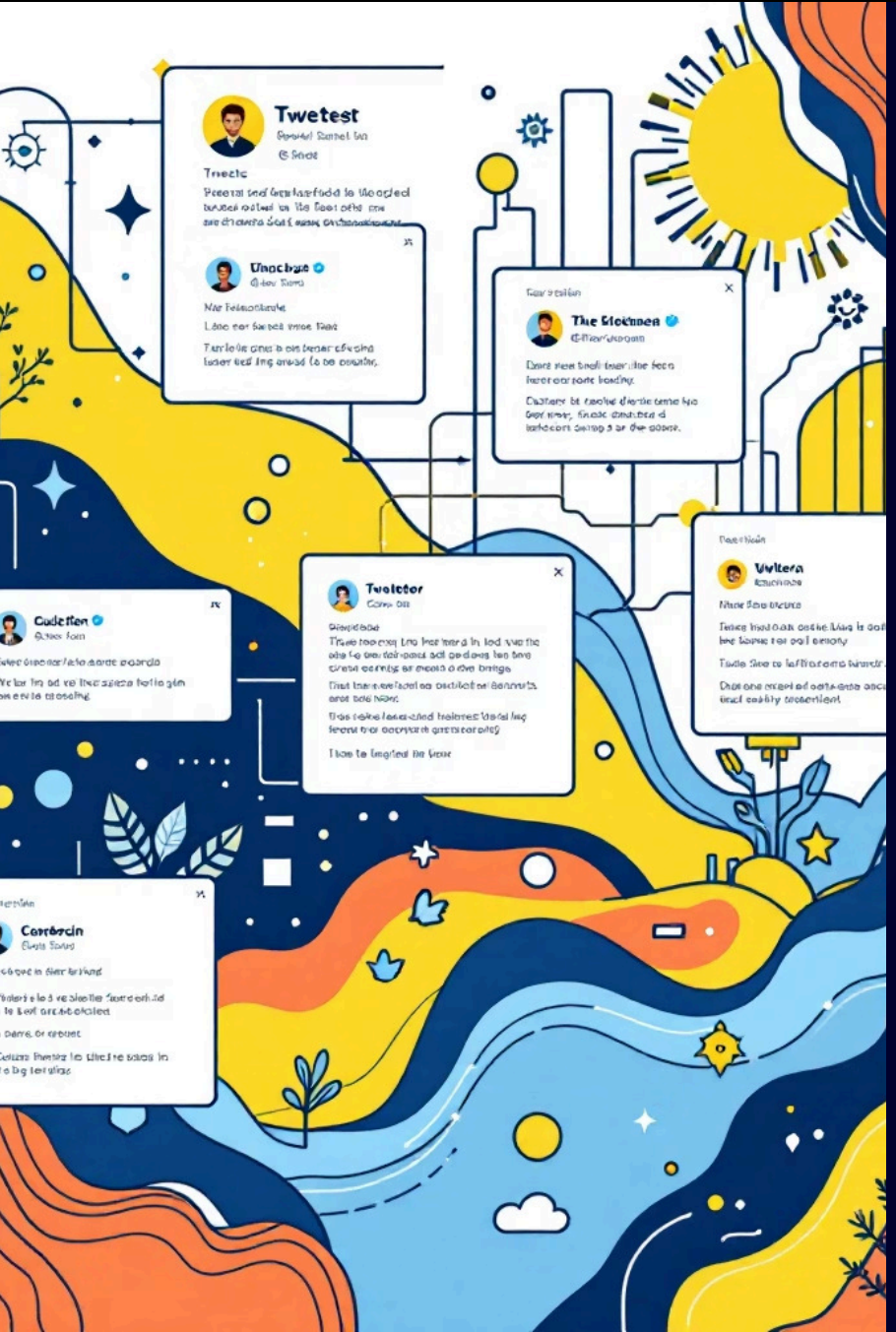
The task is to detect socio-economic aspects mentioned (e.g. healthcare, education) in tweets and classify the sentiment expressed toward each aspect.

Inputs and Outputs

- **Input:** Raw tweet text.
- **Output:** A vector where each index represents a specific socio-economic aspect, and its value indicates the corresponding sentiment score.

Models and Methods: Our Analytical Pipeline



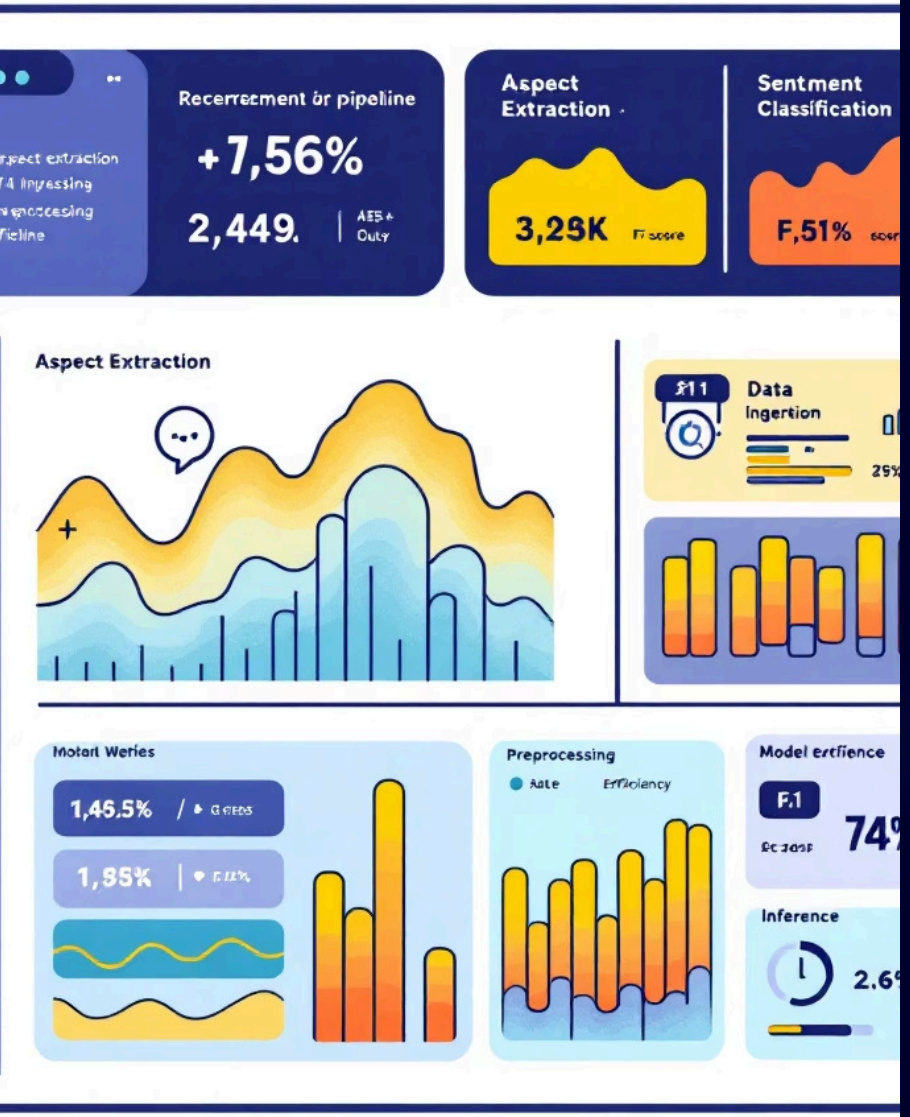


Data Specification and Generation

We will use both scraped real tweets and LLM-generated synthetic tweets.

No manual labeling is performed—aspect and sentiment annotations for the synthetic data are produced automatically by an LLM.

Combining real and synthetic tweets provides a more diverse and robust dataset for model training and testing.



Metrics and KPI

Key metrics such as Precision, Recall, and F1-score will be used to evaluate how well the system identifies the correct socio-economic aspects and assigns the appropriate sentiment. These metrics provide a clear measurement of accuracy, and the overall reliability of the model's predictions.