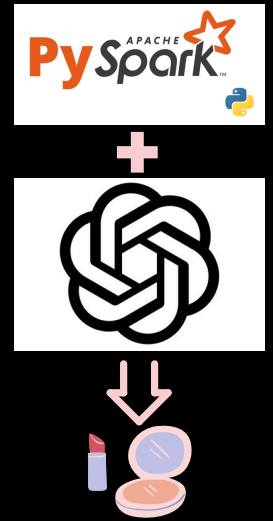TEAM 2

OpenAI-Driven Review Intelligence of the Beauty Industry with Spark
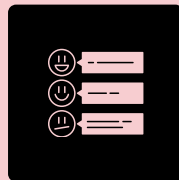
Crafting Tomorrow's Beauty Trends Today
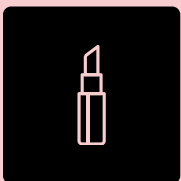
Are You Facing Business Issues Like...
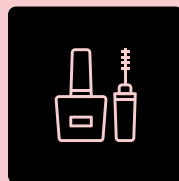
Identifying unmet needs for new product development?

Optimization of existing products based on countless consumer feedbacks?

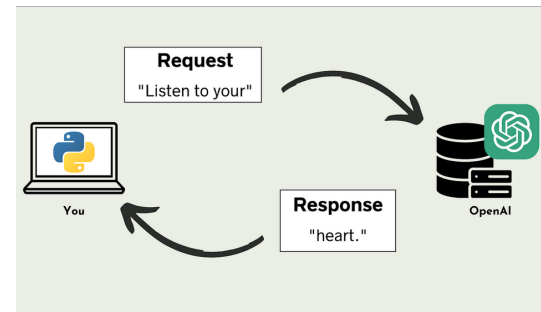**Risk of Product Failure?**

Enhancing Marketing Effectiveness

What is OpenAI ChatGPT API ?

- Provides access to advanced AI models like GPT-3 in Python environment
- Enables tasks such as text generation and language understanding.
- Customizable
- Scalable

Request "Listen to your"

Response "heart."

You

OpenAI

Business Values?

Customized Sentiment Analysis

- Enhanced insight into consumer preferences through customized sentiment categorization of reviews.

Strategic Decision Support

- Generates analytical results to guide product development & marketing strategies

Comprehensiveness

- Detailed dashboard that provides multi-level analysis

Industry Versatility

- Adaptable to various industries, requiring only the availability of review data.

Technical Highlights

Scalability in Handling Data

- Overcomes prompt token limit
- Handles large content sizes to generate topics across the full data spectrum.

Unsupervised Topic Generation

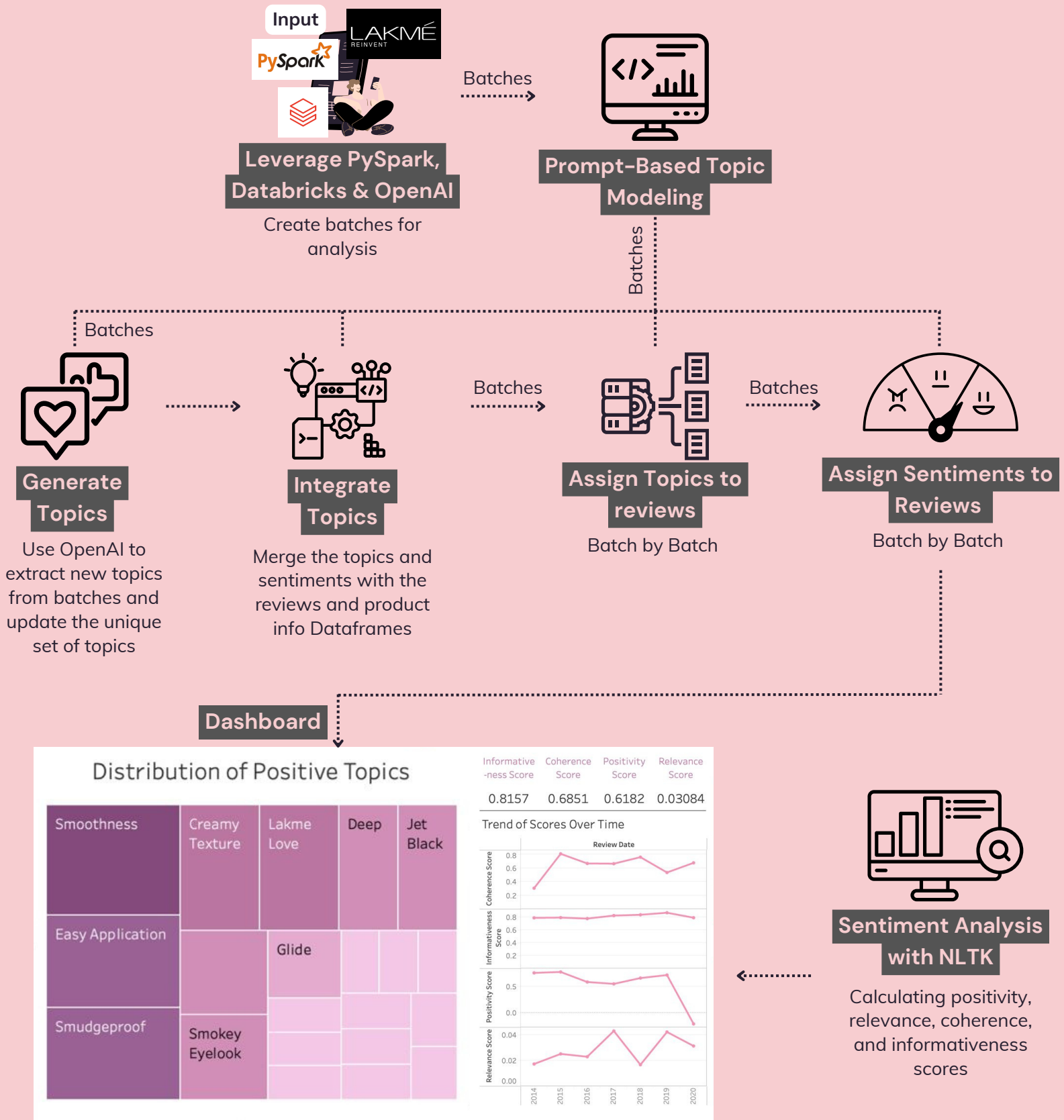- Generates topics autonomously

No Training Needed

- Pre-trained Transformer

Advanced Topic Identification

- Leverages its LLM training experience to identify topics, not relying solely on the current dataset.

How Does It Work?

Input

LAKMÉ REINVENT

PySpark

Leverage PySpark, Databricks & OpenAI

Create batches for analysis

Batches →

Prompt-Based Topic Modeling

Batches

Generate Topics

Use OpenAI to extract new topics from batches and update the unique set of topics

Batches →

Integrate Topics

Merge the topics and sentiments with the reviews and product info Dataframes

Batches →

Assign Topics to reviews

Batch by Batch

Batches →

Assign Sentiments to Reviews

Batch by Batch

Batches →

Dashboard

Distribution of Positive Topics

Smoothness	Creamy Texture	Lakme Love	Deep	Jet Black
Easy Application		Glide		
Smudgeproof	Smokey Eyelook			

Informative-ness Score	Coherence Score	Positivity Score	Relevance Score
0.8157	0.6851	0.6182	0.03084

Trend of Scores Over Time

Review Date

(Charts: Coherence Score, Informativeness Score, Positivity Score, Relevance Score over years 2014–2020)

Sentiment Analysis with NLTK

Calculating positivity, relevance, coherence, and informativeness scores

Contact Us

Praveen Kumar G — govin098@umn.edu
Anurag Wadhwa — wadhw052@umn.edu
Yumian Ding — ding0416@umn.edu
XIngyue Wang — wan01786@umn.edu
Claire Tsao — tsao0022@umn.edu
Paul Agbaje — agbaj004@umn.edu

GitHub QR Code