Seneca

Sentiment & Topic Analysis:
MKBHD's Cyberstruck Comments

Course: Sentiment Analysis & Text Mining

Professor William Dick

Seneca

Project Motivation & Questions:

01 — What do YouTube commenters think of the Cybertruck?

02 — Are highly-liked comments mostly positive or negative?

03 — What topics dominate the discussion?

➤ — *Business Need: Fast insight into public perception using text mining.*

Data & Tools

Data

YouTube comments from MKBHD Cybertruck Video: Likes, Texts, Publication Time & Author

Collection

YouTube Comment Downloader: Tool to extract, collect, and analyze YouTube comments

Libraries

pandas, NLTK for NLP, TextBlob and VADER for sentiment analysis, BERTopic for topic modeling, matplotlib visualization

Methodology Snapshot

Preprocessing

Case normalization with removal of punctuation, stopwords, emojis enhances text preprocessing accuracy significantly.

Sentiment Analysis

The internal limitations or deficiencies that may hinder the organization's progress include resource constraints, outdated systems, poor communication.

Topic Modeling

The external market conditions or circumstances that present favorable prospects for growth include emerging technologies and untapped customer segments.

Visualization

External factors or risks that may harm the business include competition, regulations, economic shifts, and disruptions.
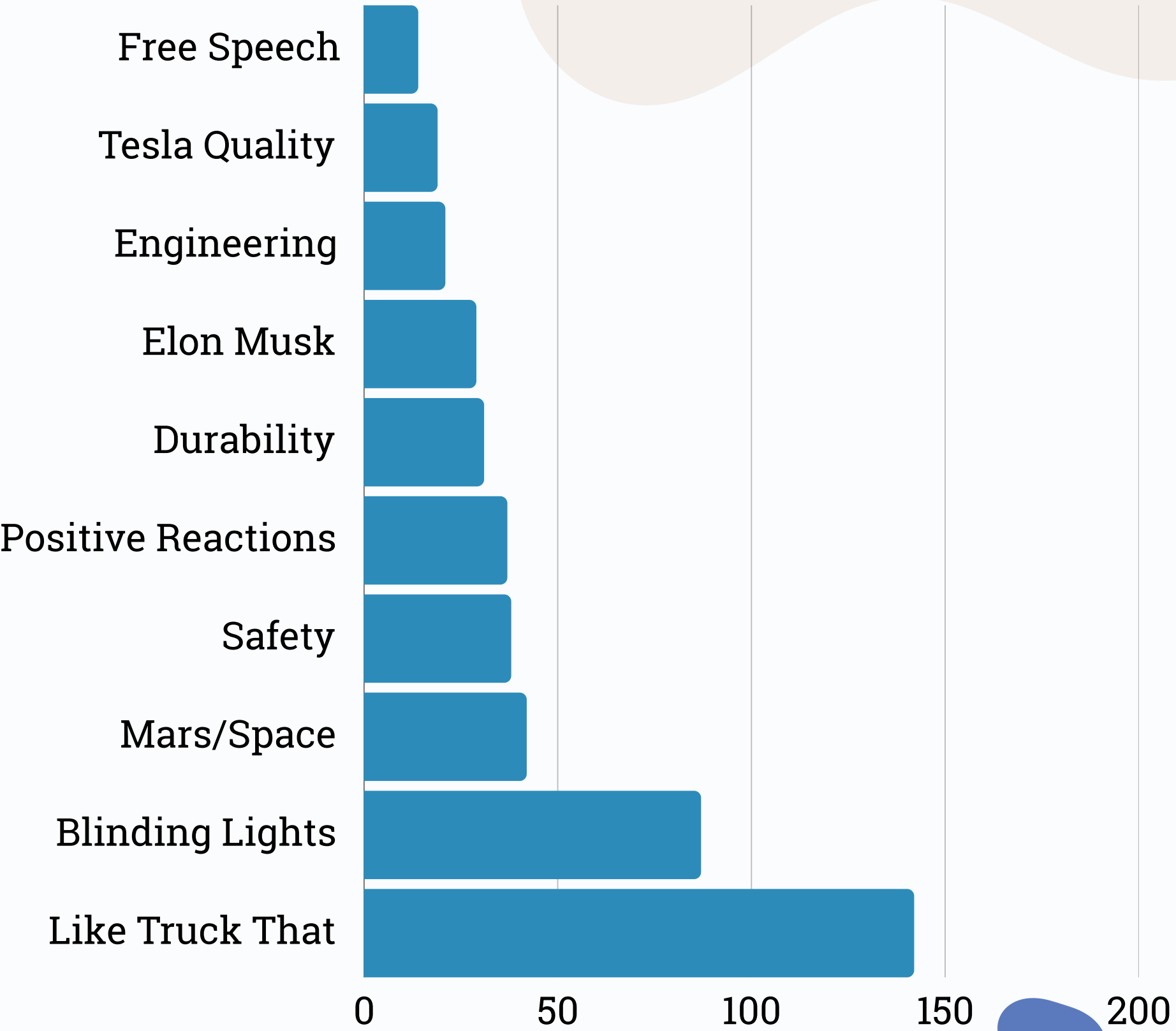
Seneca

Seneca

Sentiment Distribution of Cybertruck Comments

Finding: Slight positive bias overall, significant negativity

Neutral
26.7%

Positive
40%

Negative
33.3%

Seneca
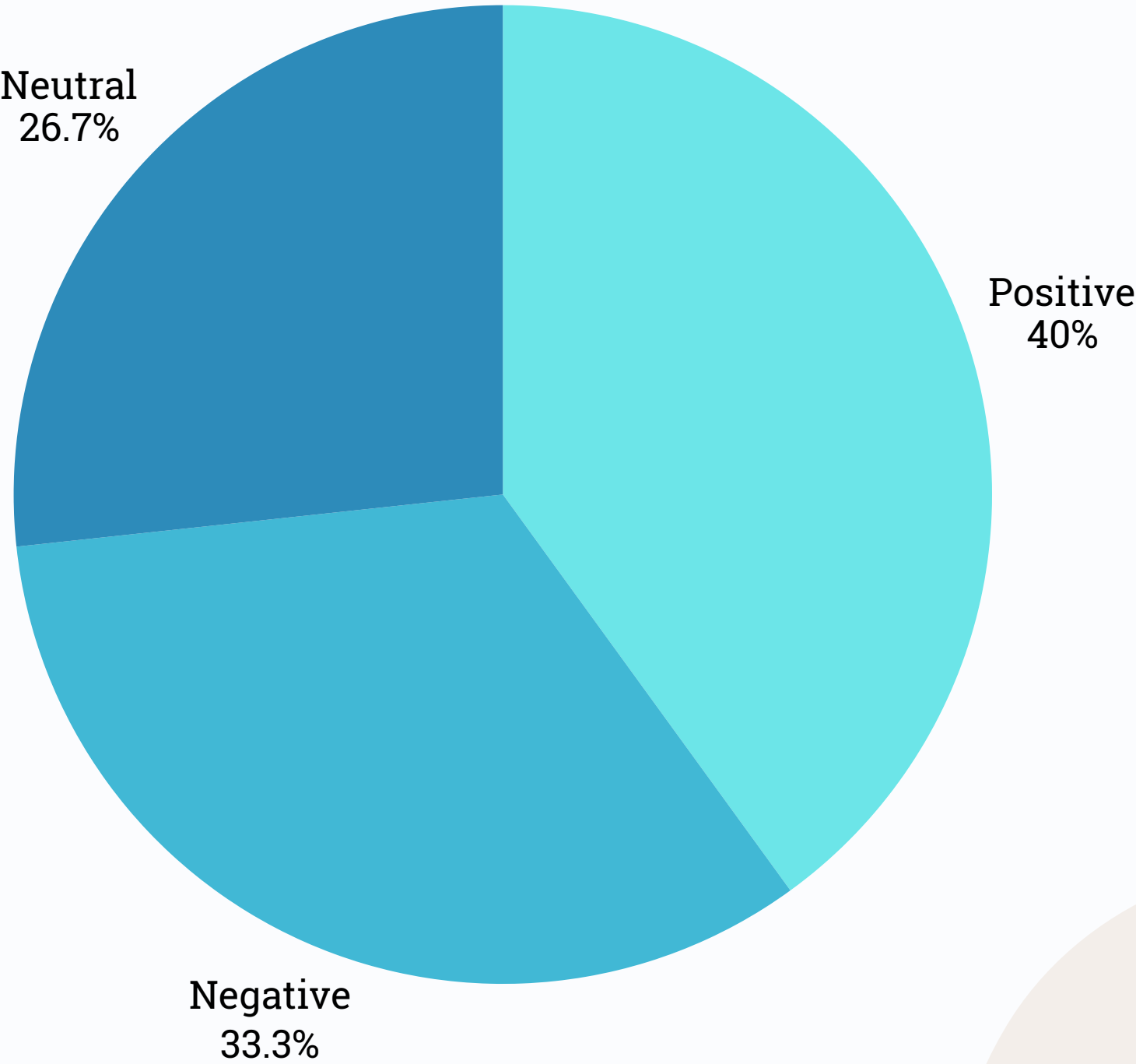
Topic Distribution in MKBHD Cybertruck Comments

Most comments: "General truck opinions"
then "headlights/visibility" and "pedestrian safety"

Seneca

Key Findings & Implications

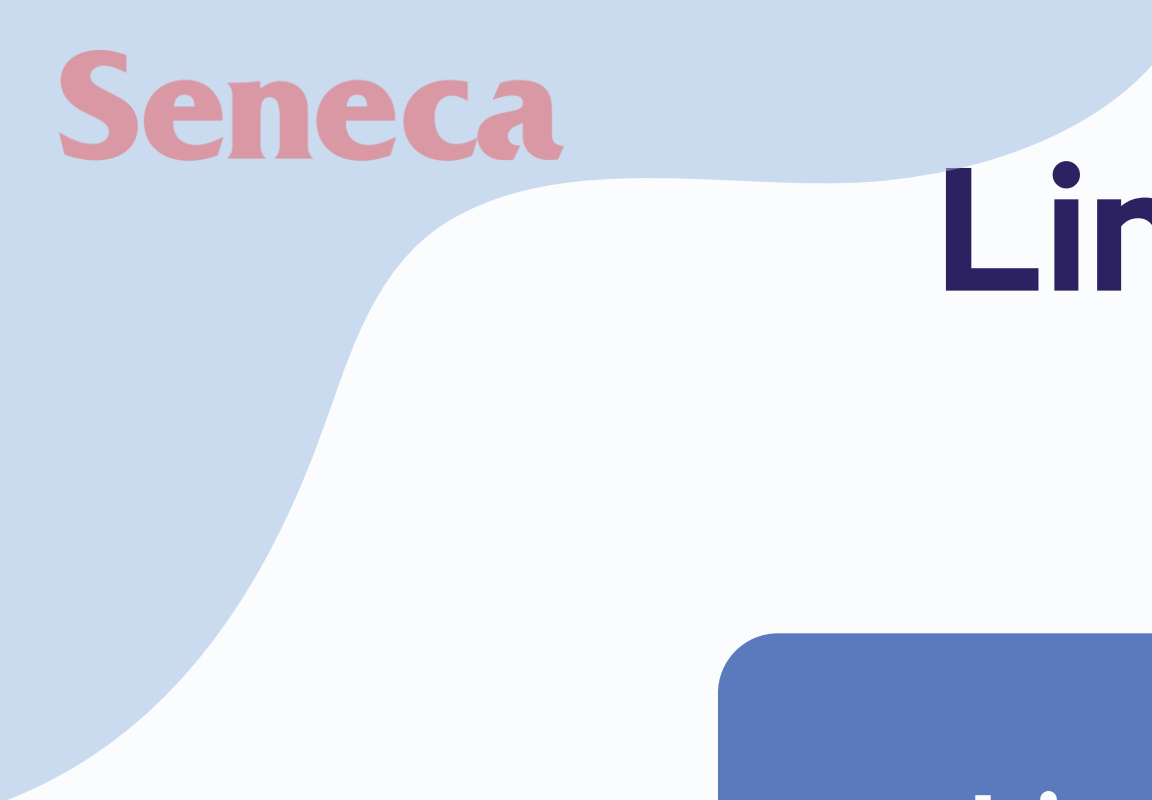**01**   **Reception is polarized: strong opinions both ways**

02   **Safety, design, and Musk are core debates**

03   **Negative/controversial reactions boost engagement**

➤   *Business Takeaways: Focus messaging on safety and design issues; controversy can fuel buzz but needs careful handling*
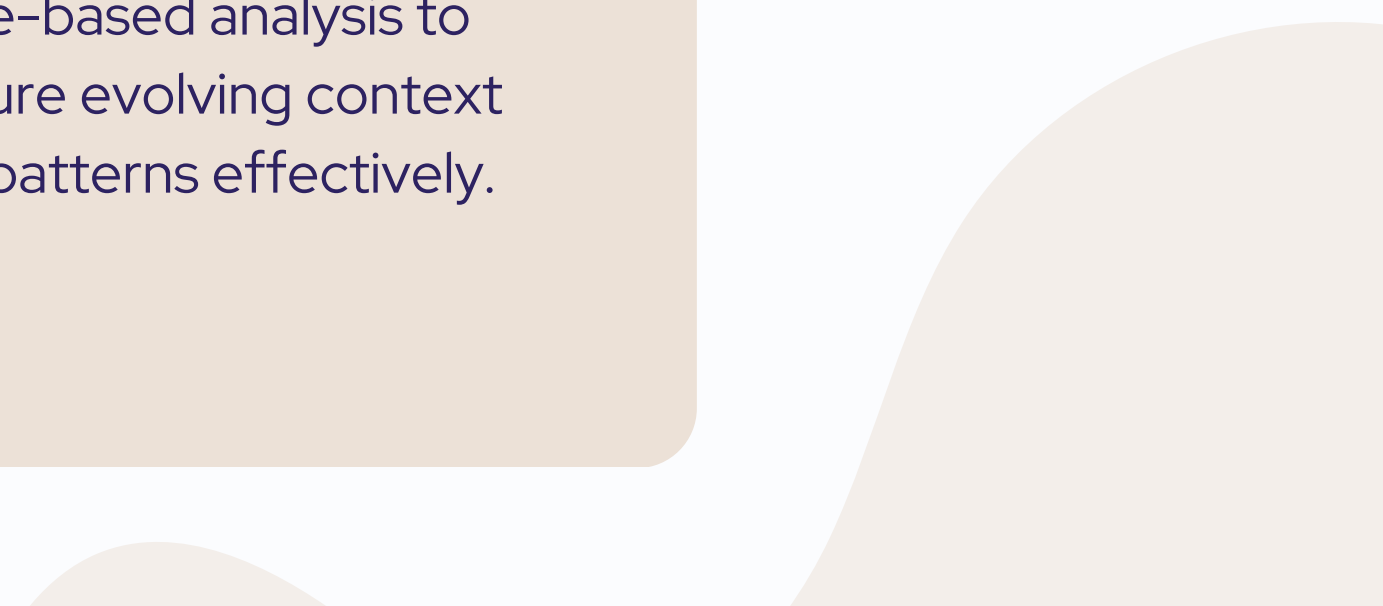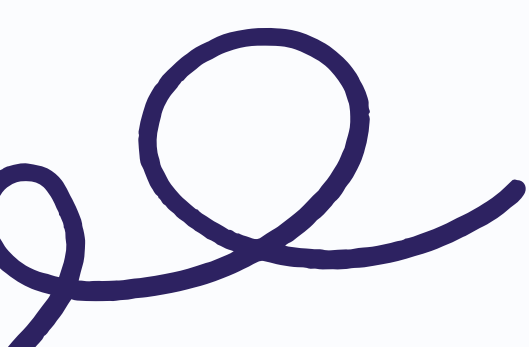
Limitation & Next Steps

Limitation

Analyzing only one video without time-series data makes it difficult to accurately detect sarcasm, tone, or contextual meaning.

Next Steps

To improve accuracy, analyze multiple sources or channels and incorporate time-based analysis to capture evolving context and patterns effectively.

Thank You!

Created by:
Arpit Trivedi | Danish Memon | Dhairya Gada
Kareena Dhamija | Ryan Lobo | Samarth Chugh