gAI Text Classification Project – Final Project

Group : has_*marketing_urgency_B*

Members :

- *Nadya Oktania*
- *Nhat Le*
- *Shreya Billipalli*
- *Kseniia Khludeeva*

youtube link: https://www.youtube.com/watch?v=FP8DXjr6ueE

Introduction

This project explores the use of **generative AI (LLMs)** and **dictionary-based methods** to classify marketing tactics in Instagram posts, with a focus on two key strategies:

- **Personalized products/services**
- **Urgency marketing**

We built multiple tools using **Claude** and **Streamlit**, crafted and tested structured prompts, refined human-labeled ground truth, and compared model outputs across different AI systems.

This hands-on process helped us evaluate the **performance, strengths, and limitations** of AI models in real-world marketing analysis.

Background

Text classification enables businesses to understand how marketing language shapes consumer perception and behavior.

By labeling social media captions based on tactics like personalization or urgency, we gain insight into:

- Which tactics are most commonly used
- How different language styles affect engagement

Our team tested both:

- **Rule-based (dictionary) methods** for speed and precision
- **Generative AI models (LLMs)** for flexible, nuanced interpretation

Later, we merged tactic predictions with engagement data (likes, comments) to analyze **real performance impact**.

Applications

Below is the list of applications along with a demonstration of how to use them.

1. CSV Text Transformer to Sentence-Level Format

2. Marketing Tactic Classifier

3. Classifier Word Metrics

4. Join table app

1. CSV Text Transformer to Sentence-Level Format

Purpose: Breaks down longer posts or chat logs into structured sentences with clear context and speaker labels.

Who can use it:

- **Data analysts** preparing textual data for NLP or classification.
- **Researchers** analyzing customer-agent conversations.
- **Product teams** needing chat/post-level features for modeling or dashboards.

Flow:

1. Upload a CSV with text data.
2. Select columns for ID and context.
3. Choose options for how context and statements should be defined (e.g., sentence vs. turn, rolling vs. full).
4. Download the transformed CSV for use in downstream tasks like classification or summarization.

---

2. Marketing Tactic Classifier

Purpose: Allows users to manage keyword sets for different marketing tactics and classify text based on keyword presence.

Who can use it:

- **Marketing teams** defining and refining tactic taxonomies.

- **Data labelers** who want a semi-automated way to assign labels.
- **Analysts** validating text-based marketing content.

Flow:

1. View and edit tactic definitions (each with a list of keywords).
2. Add new keywords or tactics as needed.
3. Classify input statements to see which tactic(s) match.
4. Export or save updated keyword sets and classification results.

---

3. Classifier Word Metrics

Purpose: Measures the alignment of each statement with tactic-specific keywords for interpretability.

Who can use it:

- **Model evaluators** needing transparency into text classification.
- **Marketing strategists** exploring how closely content aligns with target themes.
- **Analysts** assessing keyword coverage or bias.

Flow:

1. Input a list of sentences (or use results from the Text Transformer).
2. Compare each sentence against each tactic's keyword list.
3. Output includes a score or count per tactic.
4. Use these metrics for ranking, filtering, or visualization.

---

4. CSV File Joiner

Purpose: Combines two datasets (e.g., sentence-level output + classification results) based on shared keys.

Who can use it:

- **Analysts and engineers** merging pre- and post-classification outputs.

- **QA teams** validating labeled vs. raw data.
- **Anyone** doing structured data prep without writing code.

Flow:

1. Upload two CSV files.
2. Choose the columns to join on and the type of join (inner, outer, left, right).
3. Preview and download the merged result.
4. Use the final file for analysis, modeling, or reporting.

Overall reflection

Across all stages, we found that:

- **AI is fast and scalable**, but misses nuance.
- **Prompt design** is everything. Even small changes can cut model disagreement rates drastically.
- **Human judgment** remains essential for gray areas.
- **Dictionary methods** had higher precision, but missed subtle personalization.
- **gAI models** showed high recall, useful for trend detection.

We experimented with fuzzy matching, built reusable tools, and developed a strong understanding of how to balance automation and manual review in real-world classification projects.
