TABLE OF CONTENT:

OVERVIEW OF THE TASK

- Create and test prompts to classify the sentiment (positive, neutral, negative) of user reviews with high accuracy and clarity.
- Design a Set of Initial Prompts: Generate at least 5 initial prompt variations for a language model to classify sentiment in reviews.
- Fine-Tune Prompts for Edge Cases: Adjust and iterate on these prompts to handle challenging cases (e.g., sarcasm, mixed emotions).
- Experiment with Context: Add context, like “from a customer support perspective,” to see how it affects sentiment accuracy.

INTRODUCTION

What is Prompt Engineering?

Generative artificial intelligence (AI) systems are designed to generate specific outputs based on the quality of provided prompts. Prompt engineering helps generative AI models better comprehend and respond to a wide range of queries, from the simple to the highly technical.

The basic rule is that good prompts equal good results. Generative AI relies on the iterative refinement of different prompt engineering techniques to effectively learn from diverse input data and adapt to minimize biases, confusion and produce more accurate responses.

Prompt engineers play a pivotal role in crafting queries that help generative AI models understand not just the language but also the nuance and intent behind the query. A high-quality, thorough and knowledgeable prompt, in turn, influences the quality of AI-generated content, whether it's images, code, data summaries or text.

A thoughtful approach to creating prompts is necessary to bridge the gap between raw queries and meaningful AI-generated responses. By fine-tuning effective prompts, engineers can significantly optimize the quality and relevance of outputs to solve for both the specific and the general. This process reduces the need for manual review and post-generation editing, ultimately saving time and effort in achieving the desired outcomes.

Why Prompt Engineering?

Greater developer control

Prompt engineering gives developers more control over users' interactions with the AI. Effective prompts provide intent and establish context to the large language models. They help the AI refine the output and present it concisely in the required format.

They also prevent your users from misusing the AI or requesting something the AI does not know or cannot handle accurately. For instance, you may want to limit your users from generating inappropriate content in a business AI application.

Improved user experience

Users avoid trial and error and still receive coherent, accurate, and relevant responses from AI tools. Prompt engineering makes it easy for users to obtain relevant results in the first prompt. It helps mitigate bias that may be present from existing human bias in the large language models' training data.

Further, it enhances the user-AI interaction so the AI understands the user's intention even with minimal input. For example, requests to summarize a legal document and a news article get different results adjusted for style and tone. This is true even if both users just tell the application, *"Summarize this document."*

Increased flexibility

Higher levels of abstraction improve AI models and allow organizations to create more flexible tools at scale. A prompt engineer can create prompts with domain-neutral instructions highlighting logical links and broad patterns. Organizations can rapidly reuse the prompts across the enterprise to expand their AI investments.

For example, to find opportunities for process optimization, the prompt engineer can create different prompts that train the AI model to find inefficiencies using broad signals rather than context-specific data. The prompts can then be used for diverse processes and business units.

TECHNICAL ASPECTS

The Technical aspects of the project has been listed below:

- Language Model: Gemini 1.5 Flash (Google Gemini)
- Programming Language: Python 3.13.0
- IDE (Integrated Development Environment): Google Colab

GEMINI 1.5 FLASH:

Gemini 1.5 Flash is a powerful and efficient AI model designed for high-volume, high-frequency tasks. It excels at multimodal reasoning, processing both text and visual information, and is capable of tasks like summarization, chat applications, image and video captioning, data extraction from documents and tables, and more.

PROPERTY	DESCRIPTION
Model code	models/gemini-1.5-flash
Supported data types	Inputs Audio, images, video, and text Output Text
Token limits	Input token limit 1,048,576 Output token limit 8,192
Audio/visual specs	Maximum number of images per prompt 3,600 Maximum video length 1 hour Maximum audio length Approximately 9.5 hours

Rate limits	Free: • 15 RPM • 1 million TPM • 1,500 RPD Pay-as-you-go: • 2,000 RPM • 4 million TPM
Capabilities	System instructions JSON mode JSON schema Adjustable safety settings Caching Tuning Function calling

Versions	• Latest: gemini-1.5-flash-latest
	• Latest stable: gemini-1.5-flash
	• Stable:
	  • gemini-1.5-flash-001
	  • gemini-1.5-flash-002

PYTHON 3.13.0:

Python 3.13.0: A Leap Forward

Python 3.13.0 is a significant release that introduces a range of innovative features and performance enhancements. This version aims to elevate the Python programming experience to new heights, making it more efficient, user-friendly, and powerful.

A Refined Interactive Interpreter

One of the key improvements in Python 3.13.0 is the enhanced interactive interpreter. It now supports multi-line editing, allowing developers to seamlessly edit and execute multiple lines of code within a single session. Additionally, the introduction of color support and colorized exception tracebacks significantly enhances code readability and debugging efficiency.

Unlocking Parallelism: Experimental Free-Threaded Mode

Python 3.13.0 introduces an experimental free-threaded mode, which relaxes the Global Interpreter Lock (GIL). This groundbreaking feature allows for the parallel execution of multiple threads, unlocking the potential for substantial performance gains in multi-threaded applications. While still in its experimental phase, this innovation promises to revolutionize the way Python handles concurrent tasks.

A Glimpse into the Future: Preliminary Just-In-Time (JIT) Compiler

Python 3.13.0 takes a significant step towards optimizing performance with the introduction of a preliminary Just-In-Time (JIT) compiler. This experimental feature compiles Python bytecode into machine code at runtime, potentially leading to significant speed improvements. While still in its early stages, the JIT compiler holds the promise of transforming Python into a high-performance language.

Enhanced Developer Experience

Python 3.13.0 prioritizes developer experience by providing clearer and more informative error messages. These improved tracebacks aid in swift identification and resolution of issues, streamlining the debugging process. Furthermore, the locals() function now exhibits well-defined semantics for mutating the returned mapping, empowering developers with greater control and flexibility.

GOOGLE COLAB:

Colab is a hosted Jupyter Notebook service that requires no setup to use and provides free of charge access to computing resources, including GPUs and TPUs. Colab is especially well suited to machine learning, data science, and education.

In order to provide access to students and under-resourced groups around the world, Colab prioritizes users who are actively programming in a notebook. Users in our free of charge tier commonly experience runtime terminations when attempting to bypass the notebook UI and using a web UI on a Colab managed runtime for content generation.

PROMPTING DEVELOPMENT PROCESS

CONTEXT DEVELOPMENT

Context is a crucial component of effective prompt engineering. It provides the AI model with the background information and specific requirements needed to generate the desired output. By incorporating relevant context into the prompt, we can significantly improve the quality and accuracy of the model's responses.

Key Techniques for Context Development:

1. **Specific Instructions:**
    - **Clear and Concise:** Clearly articulate the desired task or output.
    - **Action Verbs:** Use strong action verbs to guide the model's actions (e.g., "write," "summarize," "translate").
    - **Constraints:** Specify any limitations or constraints, such as word count, format, or style.
2. **Relevant Examples:**
    - **Illustrate the Desired Output:** Provide concrete examples to showcase the expected format or content.
    - **Highlight Key Features:** Emphasize the most important aspects of the examples.
    - **Adapt to the Model's Capabilities:** Tailor the examples to the model's strengths and limitations.
3. **World Knowledge and Factual Accuracy:**
    - **Incorporate Background Information:** Provide relevant historical, cultural, or scientific context.
    - **Verify Factual Claims:** Ensure the accuracy of any factual statements or assertions.
    - **Avoid Biases:** Strive for objectivity and neutrality in the provided context.
4. **Iterative Refinement:**
    - **Experiment with Different Prompts:** Test various prompt formulations to identify the most effective ones.
    - **Analyze Model Outputs:** Evaluate the quality of the generated responses and identify areas for improvement.
    - **Adjust Prompts Accordingly:** Refine the prompt based on the model's performance.

Real-World Applications of Context Development

- **Content Generation:** Creating articles, blog posts, scripts, or marketing copy.
- **Code Generation:** Generating code snippets or entire programs.
- **Translation:** Translating text from one language to another.
- **Summarization:** Condensing long texts into shorter summaries.
- **Creative Writing:** Generating poems, stories, or scripts.

By effectively developing and incorporating context into prompts, we can unlock the full potential of AI models and achieve exceptional results. The context for the Sentiment Analysis has been listed below.

CONTEXT USED IN PROGRAM:

""" Analyze the following product review and categorize it as positive, negative, or neutral. Here is your reviews.

Look like old product, volume decrease button not working properly if you hold it for 1-2 sec then its decreased the volume, cushions cover is clearly visible, ear Bar is also very tight..!! First' time' I returned any electronic product in the Amazon disappointing from Amazon and zeb.But headphone is okay. In this price range good sound quality.. ear cushions is also comfortable!!Good headphone you can go for it.!!Thank you

Quite good product from zebronics...button quality is slightly low and based is also low kind ...bt at this price battery durability is the best and overall sound quality...Clarity is good..so it is value for money...!

You have to build the answer based on the format given below:
Overall Response : If 80% of message is positive then tell me as Positive, or the percentage is equal to 50 to 55 inform me it is Neutral, or the percentage is between 0 to 49 mark it as Negative.
Expressed Sentiment : You will analyse the different types of emotions in text. So tell the emotion of the text is whether Happiness, Sadness, Anger, Fear, Surprise, Disgust,Love: affection, adoration, passion, compassion,Jealousy,Guilt,Pride,Hope,Empathy
Positive Percentage : Feed it in a fresh view that how positively the text is.
Negative Percentage : Feed it in a fresh view that how negatively the text is.
Neutral Percentage : Feed in a fresh view that how neutral the text is.

Review the comments one by one and show it. Lastly show the overall results with the combination of reviews.
Stretch it in a single line of each.
"""

UNRAVELING THE CONTEXT: A STEP-BY-STEP BREAKDOWN

PROMPT:

Analyze the following product review and categorize it as positive, negative, or neutral. ... Very important : You have to build the answer based on the format given below: ...

UNDERSTANDING THE TASK:

1. **Sentiment Analysis:** The primary task is to determine the overall sentiment of the product review. This involves identifying whether the reviewer's opinion is positive, negative, or neutral.
2. **Emotion Detection:** The prompt asks to identify specific emotions expressed in the text, such as happiness, sadness, anger, etc.
3. **Quantitative Analysis:** The prompt requires a numerical breakdown of the sentiment, expressed as percentages of positive, negative, and neutral sentiments.

EXTRACTING THE CONTEXT:

To effectively analyze the review, we need to consider the following contextual cues:

- **Explicit Sentiment Indicators:** Words like "good," "bad," "excellent," "poor," etc. directly convey the reviewer's opinion.
- **Implicit Sentiment Indicators:** Phrases like "not working properly," "disappointing," or "value for money" indirectly convey sentiment.
- **Emotion Words:** Words like "happy," "sad," "angry," etc. explicitly express emotions.
- **Sentence Structure and Tone:** The way sentences are constructed and the overall tone of the review can provide clues about the reviewer's feelings.

APPLYING THE CONTEXT TO THE ANALYSIS:

1. **Identify Sentiment Indicators:** We scan the review for explicit and implicit sentiment indicators.
2. **Detect Emotions:** We look for emotion words and analyze the overall tone of the review.
3. **Calculate Sentiment Percentages:** We assign a sentiment score to each sentence based on the identified indicators and emotions. We then calculate the overall positive, negative, and neutral percentages.
4. **Formulate the Response:** We structure the response according to the specified format, providing a clear and concise analysis of the review.

   By carefully considering these contextual factors, we can accurately assess the sentiment of the product review and provide a comprehensive analysis.

PROMPTS EXAMPLES:

By analyzing real-world product reviews, such as those for GTA 5, we can significantly enhance the capabilities of AI models in sentiment analysis. These reviews often contain a rich tapestry of emotions, expressed both explicitly through words and implicitly through emojis.

Key Insights from Real-World Reviews:

1. **Explicit Sentiment:** Reviews often contain explicit sentiment indicators like "great," "terrible," "love," or "hate."
2. **Implicit Sentiment:** Subtle cues like sarcasm, irony, or understatements can convey sentiment indirectly.
3. **Emoji-Enhanced Sentiment:** Emojis provide visual cues that can amplify or contradict the textual sentiment.
4. **Contextual Understanding:** The overall context of the review, including the product's nature and the reviewer's background, influences sentiment interpretation.

Training AI Models with Real-World Data:

By feeding AI models with a diverse dataset of real-world reviews, we can train them to:

- **Recognize Explicit Sentiment:** Identify keywords and phrases that directly express positive or negative opinions.
- **Detect Implicit Sentiment:** Understand the nuances of language and identify subtle cues that convey sentiment.
- **Interpret Emoji Sentiment:** Analyze emojis to enhance sentiment analysis accuracy.
- **Consider Contextual Factors:** Account for the broader context of the review to make more accurate assessments.

Benefits of Real-World Data:

Real-world data helps models learn from real-world nuances and complexities. Models trained on real-world data are better equipped to handle diverse sentiment expressions. AI models can become more resilient to noise, ambiguity, and sarcasm. By leveraging the power of real-world data, we can develop AI models that are more accurate, nuanced, and effective in understanding human sentiment.

Here is the Reviews feed in Amazon E-Commerce Site under the Product GTA 5 ENHANCED VERSION:

- "I got the key within 10 minute on my mail and the game is 100 percent genuine you can trust."
- "NOT THE GENUINE GAME, ITS ONLY A CRACKED VERSION."
- "It was quick, I got it in one day but it was sent through amazon via text message in the app. Great product, genuine product."
- "this osm game is playing it's a bless for me thanks for this game amazon. i really love it thank you so much"
- "I got only story mode. There is no code for online. And you cannot get refund also."

OUTPUT FORMAT:

The output was divided into individual and overall review analyses. Each review was assessed separately, and then the results were combined to provide an overall sentiment. This approach allows for a detailed understanding of both individual and collective opinions expressed in the reviews. The Expected and Acquired Output has been formatted by the following Alignment:

EXPECTED OUTPUT:

Overall Response: The aim is to categorize text-based sentiment as positive, neutral, or negative, based on the following thresholds:
Positive: If 80% or more of the sentiment is positive.
Neutral: If the sentiment is between 50% and 55% positive.
Negative: If the sentiment is 49% or less positive.

Expressed Sentiment: The task is to analyze text and identify the primary emotion expressed. This involves determining whether the text conveys happiness, sadness, anger, fear, surprise, disgust, love (including affection, adoration, passion, and compassion), jealousy, guilt, pride, hope, or empathy.

Positive Percentage: To quantify the positivity of the text, assigning a numerical value to its positive sentiment. This involves analyzing the text for positive language, positive emotions, and other indicators of positivity. The higher the numerical value, the more positive the text is deemed to be.

Negative Percentage: To quantify the negativity of the text, assigning a numerical value to its negative sentiment. This involves analyzing the text for negative language, negative emotions, and other indicators of negativity. The higher the numerical value, the more negative the text is deemed to be.

Neutral Percentage: To quantify the neutrality of the text, assigning a numerical value to its neutral sentiment.

THE ACQUIRED OUTPUT:

```
Here's
 an analysis of the product reviews, following your format:

**Review 1:**


* **Overall Response:** Negative
* **Expressed Sentiment:**
Anger, Disgust
* **Positive Percentage:** 20%
* **Negative Percentage:** 80%
* **Neutral Percentage:**
0%
```

```
**Review 2:**

* **Overall Response:** Neutral
* **Expressed Sentiment:**  Neutral
* **Positive Percentage:**
50%
* **Negative Percentage:** 30%
* **Neutral Percentage:** 20%
```

```
**Overall Results (Combined):**

* **Overall Response:**  Negative
* **Expressed Sentiment:** Anger, Disgust

* **Positive Percentage:** 35%
* **Negative Percentage:** 55%
* **Neutral Percentage:** 10%
```

PERFORMANCE DATA

The performance evaluation was conducted in a three-phased approach, each stage progressively delving deeper into specific parameters: data volume, input method, and prompt format.

Stage 1: Baseline Assessment

In the initial phase, a fundamental evaluation was conducted. Prompts, accompanied by relevant context and explicit instructions, were fed to the language model. The model's ability to process and comprehend these inputs, and subsequently generate accurate and relevant outputs, was assessed. This baseline evaluation established a benchmark against which subsequent stages could be compared.

“

```
prompt=""" Analyze the following product review and categorize it as
positive, negative, or neutral. Here Is your reviews...

Look like old product, volume decrease button not working properly if you
hold it for 1-2 sec then its decreased the volume, cushions cover is
clearly visible, ear Bar is also very tight..!! First' time' I returned
any electronic product in the Amazon disappointing from Amazon and zeb.But
headphone is okay. In this price range good sound quality.. ear cushions
is also comfortable!!Good headphone you can go for it.!!Thank you

Quite good product from zebronics...button quality is slightly low and
based is also low kind ...bt at this price battery durability is the best
and overall sound quality...Clarity is good..so it is value for money...!

Very important : You have to build the answer based on the.....
```

Stage 2: Scalability Testing

To gauge the model's scalability and potential, the number of prompts was significantly increased. This phase aimed to determine the

model's capacity to handle larger datasets without compromising accuracy or efficiency. The model's ability to maintain accurate and consistent performance, even when faced with a substantial increase in input, was a key focus of this stage. Additionally, the model's efficiency in processing and analyzing information was evaluated.

```
sentences = []

review_test1 = "Look like old product, volume decrease button not working
properly if you hold it for 1-2 sec then its decreased the volume,
cushions cover is clearly visible, ear Bar is also very tight..!! First'
time' I returned any electronic product in the Amazon disappointing from
Amazon and zeb.But headphone is okay. In this price range good sound
quality.. ear cushions is also comfortable!!Good headphone you can go for
it.!!Thank you
"

review_test2 = "Quite good product from zebronics...button quality is
slightly low and based is also low kind ...bt at this price battery
durability is the best and overall sound quality...Clarity is good..so it
is value for money...!
"

# Constructing the Prompt for the Model
context=f""" Analyze the following Reviews and categorize it as positive,
negative, or neutral.{sentences}

You have to build the answer based on the ......
```

Stage 3: Data Input Optimization

The final stage focused on optimizing the data input process. A dataset was directly read from a .txt file, stored in a structured format (array), and then fed to the model. This approach streamlined the data ingestion process, eliminating the need for manual intervention and reducing potential errors. By automating the data input process, the model could efficiently process large volumes of review data, leading to improved performance and scalability.

“

```
sentences = []

with open('Review Dataset.txt', 'r') as file:
    for line in file:
        line = line.strip()  # Remove leading/trailing whitespace
        if line:  # Check if line is not empty
            sentences.append(line)

# Constructing the Prompt for the Model
context=f""" Analyze the following Reviews and categorize it as positive,
negative, or neutral.{sentences}

You have to build the answer based on the .....
```

NOTE: There are 100+ Reviews has been used for the Model Training and the Datasets has divided into two separated files as .txt and .csv files respectively.

CONCLUSION

Sentiment analysis is a crucial technique for understanding public opinion and customer feedback. By effectively training AI models to accurately classify sentiment, businesses can make data-driven decisions and improve customer satisfaction.

To achieve accurate sentiment analysis, prompt engineering plays a pivotal role. By carefully crafting prompts, we can guide the AI model to interpret the nuances of language and identify the underlying sentiment. This involves designing prompts that are clear, concise, and specific, while also considering factors like context, sarcasm, and mixed emotions.

One effective approach is to use role-based prompts, where the model is asked to classify sentiment from the perspective of a specific role, such as a customer service representative. This helps the model to consider the underlying intent and emotions behind the text. Additionally, question-based prompts can be used to encourage the model to actively engage with the text and provide a reasoned response.

By iteratively testing and refining prompts, we can continuously improve the accuracy of sentiment analysis. This involves evaluating the model's performance on a diverse range of texts, identifying areas for improvement, and making adjustments to the prompts accordingly.