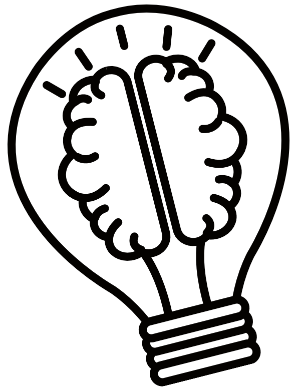
HACKTIV8

10-STEPS POEM SUMMARIZATION USING IBM GRANITE-3.3B-INSTRUCT

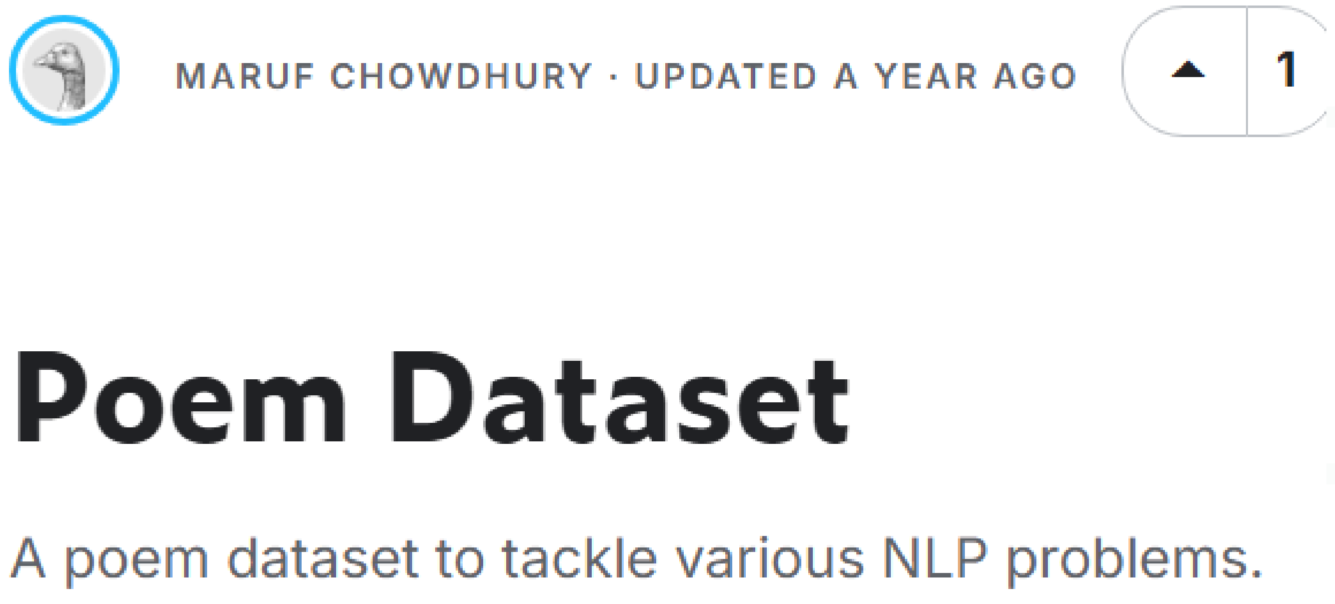
Leveraging Gen AI for Structured Poem Analysis



ADJI PUTRA NUGRAHA KUSUMA, 8/15/2025

RAW DATASET LINK

- Source: Kaggle Poem Dataset by Maruf Chowdhury
- Description: Contains Poems_Dataset.csv and poemDatasetWithSummary.csv with poem texts, titles, and summaries.
- Link: DOI: 10.34740/kaggle/dsv/7981215
- License: Apache 2.0



MARUF CHOWDHURY · UPDATED A YEAR AGO

▲ 1

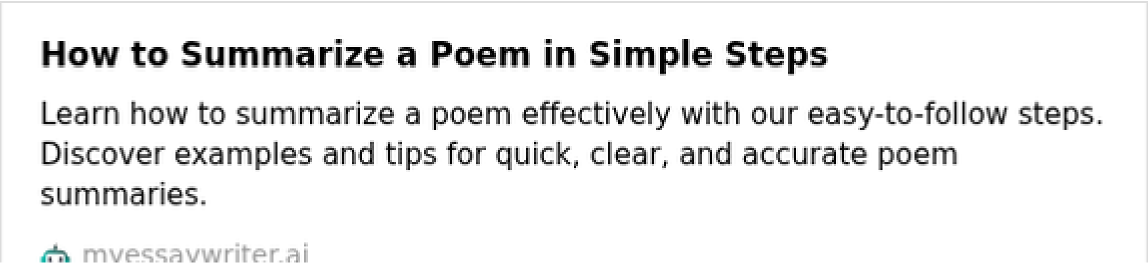
Poem Dataset

A poem dataset to tackle various NLP problems.

Project Overview

HACKTIV8

1. **Objective:** Use IBM Granite-3.3B-Instruct via Replicate and LangChain to summarize poems using a 10-step methodology inspired by MyEssayWriter.ai.
2. **Scope:** Summarize poems from the Kaggle Poem Dataset, focusing on structured outputs for two poems, including "The Road Not Taken."
3. **Purpose:** Demonstrate AI's ability to generate concise, thematic, and structured poem summaries.



How to Summarize a Poem in Simple Steps

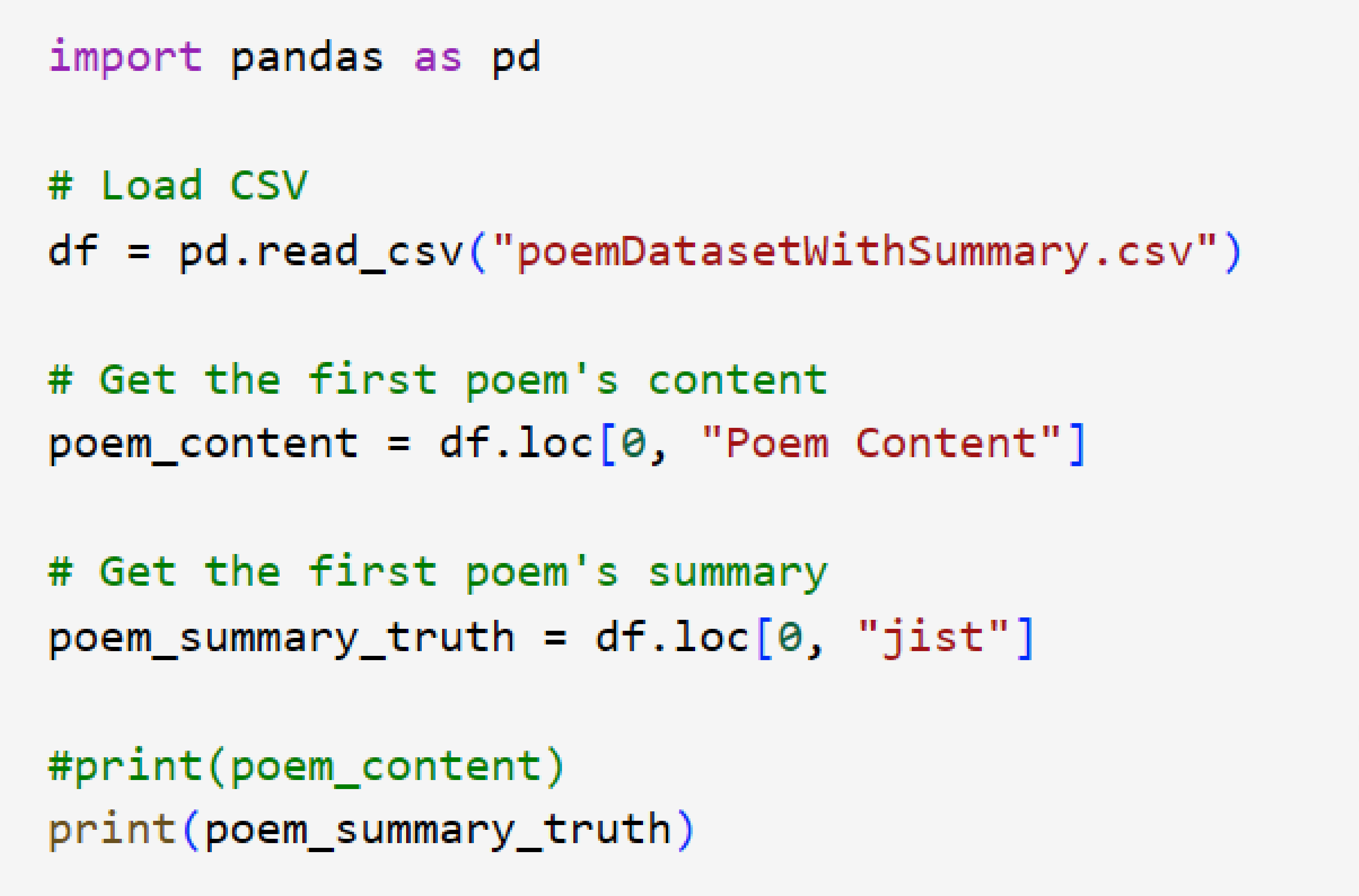
Learn how to summarize a poem effectively with our easy-to-follow steps. Discover examples and tips for quick, clear, and accurate poem summaries.

myessaywriter.ai

1. Write a One-Sentence Summary
2. Explain Each Stanza Separately
3. Highlight Key Words
4. Use a Metaphor
5. Incorporate Imagery
6. Pose a Reflective Question
7. Describe a Relevant Scenario
8. Include a Key Quote
9. Write a Comprehensive Summary
10. Reference a Specific Line or Paragraph

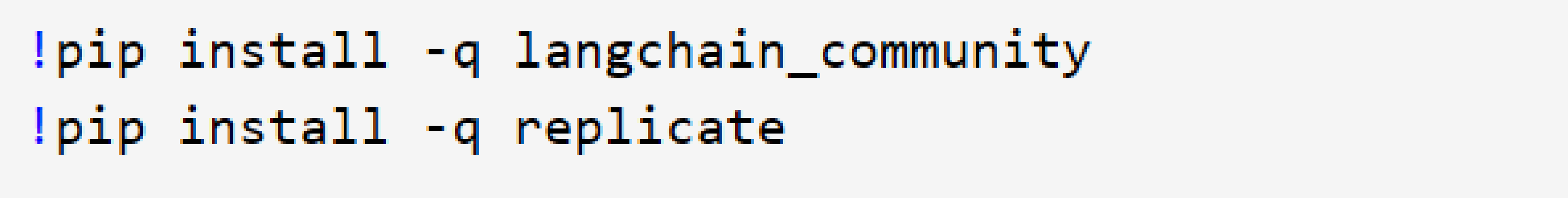
ANALYSIS PROCESS

Step 1: Load dataset (poemDatasetWithSummary.csv) and extract poem content and ground truth summaries

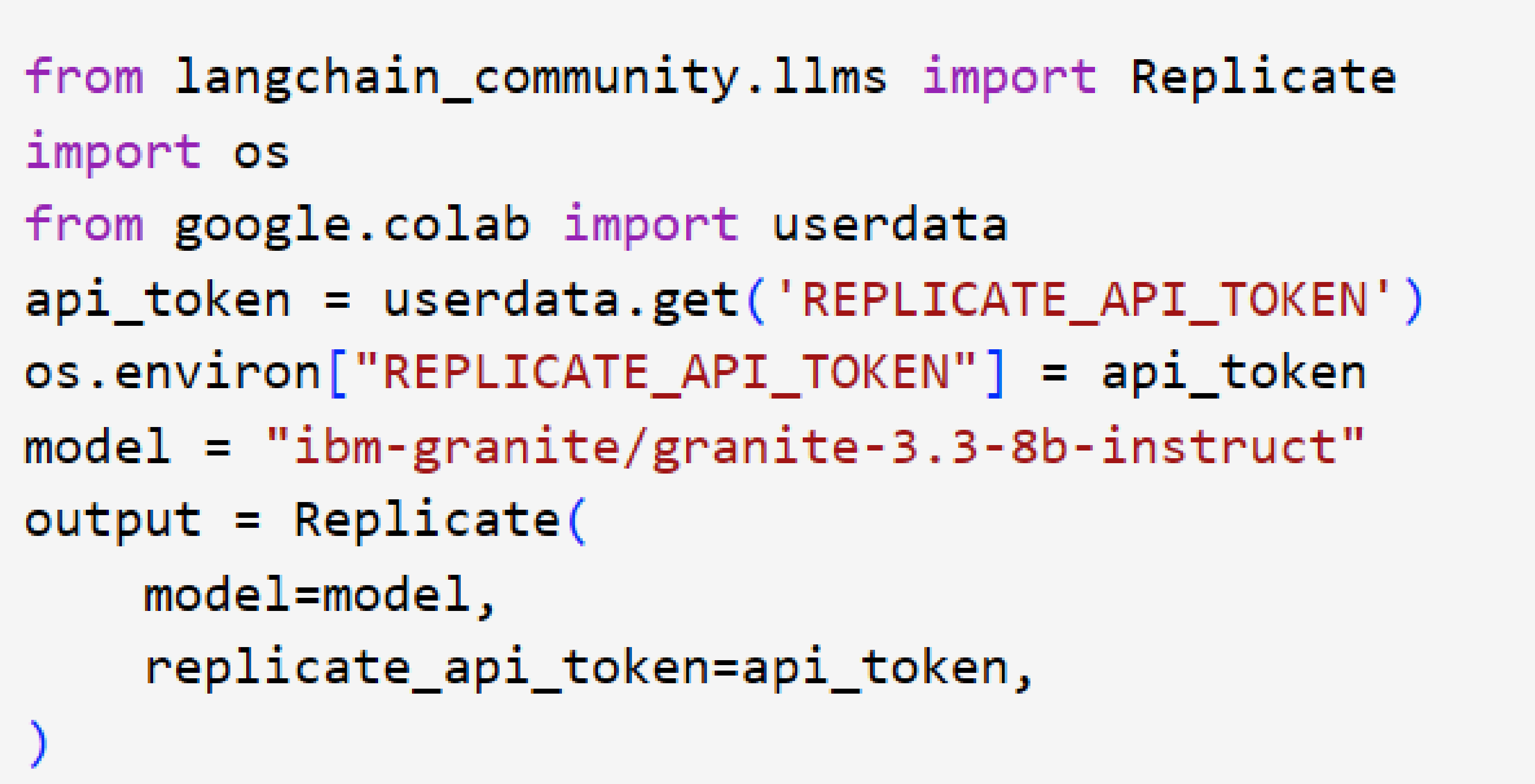


```python
import pandas as pd

# Load CSV
df = pd.read_csv("poemDatasetWithSummary.csv")

# Get the first poem's content
poem_content = df.loc[0, "Poem Content"]

# Get the first poem's summary
poem_summary_truth = df.loc[0, "jist"]

#print(poem_content)
print(poem_summary_truth)
```

Step 2: Use IBM Granite-3.3B-Instruct model via Replicate API



```python
!pip install -q langchain_community
!pip install -q replicate
```



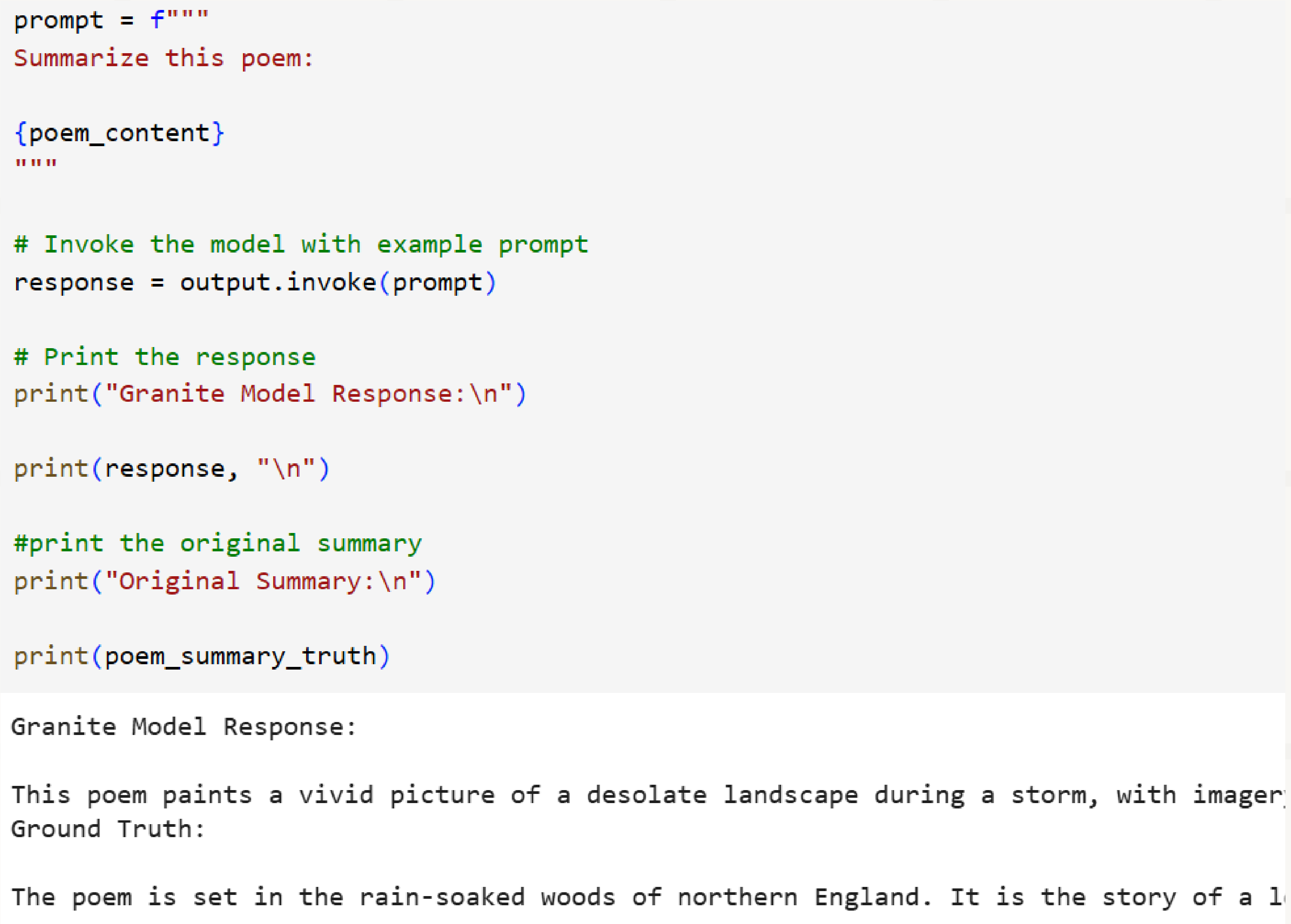
```python
from langchain_community.llms import Replicate
import os
from google.colab import userdata
api_token = userdata.get('REPLICATE_API_TOKEN')
os.environ["REPLICATE_API_TOKEN"] = api_token
model = "ibm-granite/granite-3.3-8b-instruct"
output = Replicate(
    model=model,
    replicate_api_token=api_token,
)
```

ANALYSIS PROCESS

Step 3: Apply progressive prompt refinements:

1. Basic summary (Cell 4).
2. Two-sentence summary (Cell 5).
3. Two-sentence summary with focus areas (setting, story) (Cell 6).
4. Structured format with 7-word sentences (Cell 7).
5. 10-step summarization methodology for two poems (Cells 8-11).

1.



```
prompt = f"""
Summarize this poem:

{poem_content}
"""

# Invoke the model with example prompt
response = output.invoke(prompt)

# Print the response
print("Granite Model Response:\n")

print(response, "\n")

#print the original summary
print("Original Summary:\n")

print(poem_summary_truth)
```

```
Granite Model Response:

This poem paints a vivid picture of a desolate landscape during a storm, with imager
Ground Truth:

The poem is set in the rain-soaked woods of northern England. It is the story of a l
```

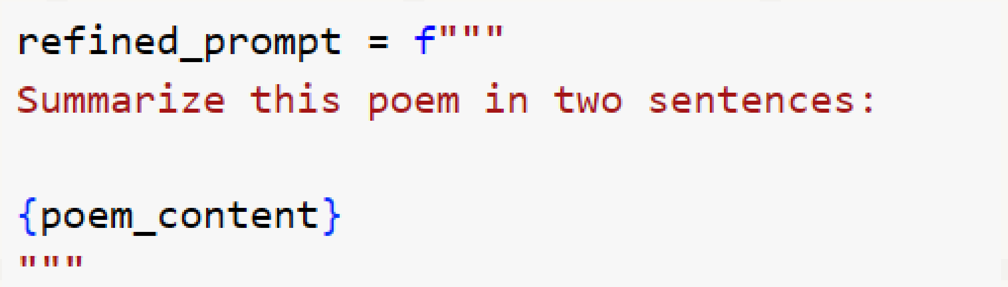
ANALYSIS PROCESS

HACKTIV8

Step 3: Apply progressive prompt refinements:

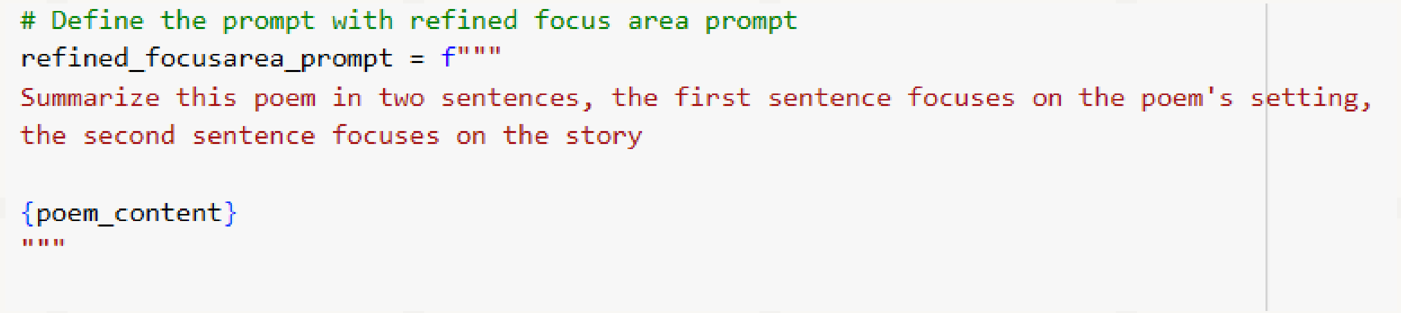
1. Basic summary (Cell 4).
2. Two-sentence summary (Cell 5).
3. Two-sentence summary with focus areas (setting, story) (Cell 6).
4. Structured format with 7-word sentences (Cell 7).
5. 10-step summarization methodology for two poems (Cells 8-11).

2.



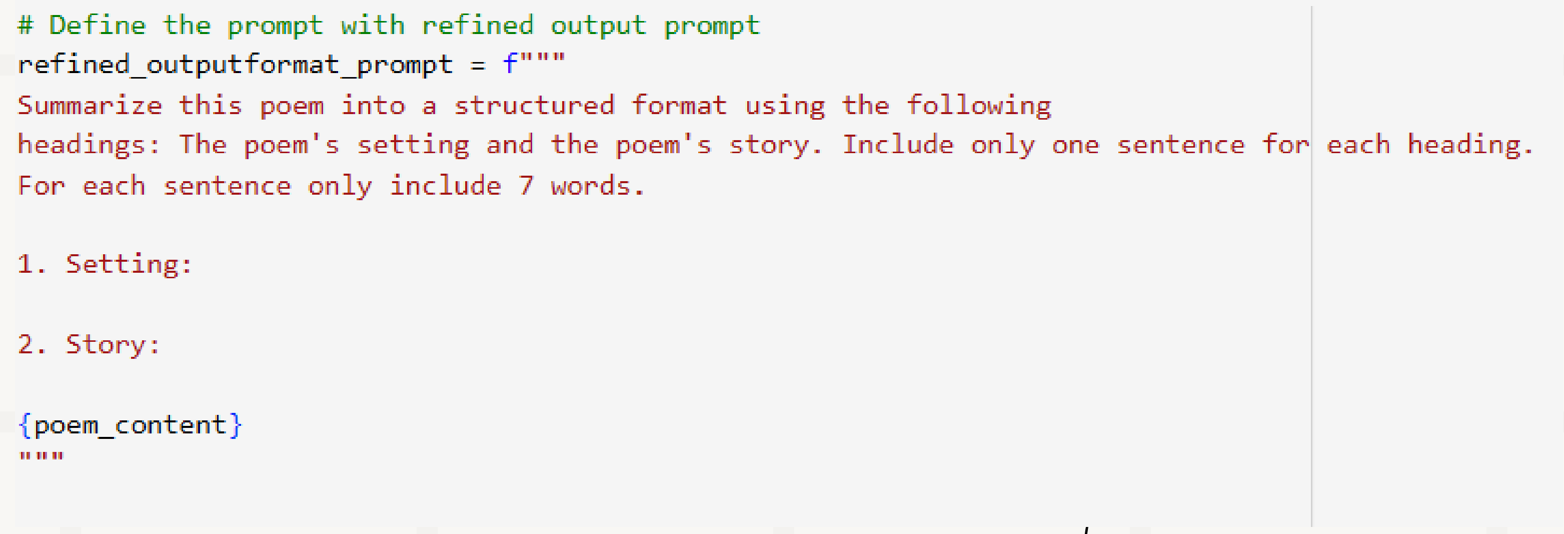
```
refined_prompt = f"""
Summarize this poem in two sentences:

{poem_content}
"""
```

3.



```
# Define the prompt with refined focus area prompt
refined_focusarea_prompt = f"""
Summarize this poem in two sentences, the first sentence focuses on the poem's setting,
the second sentence focuses on the story

{poem_content}
"""
```

4.



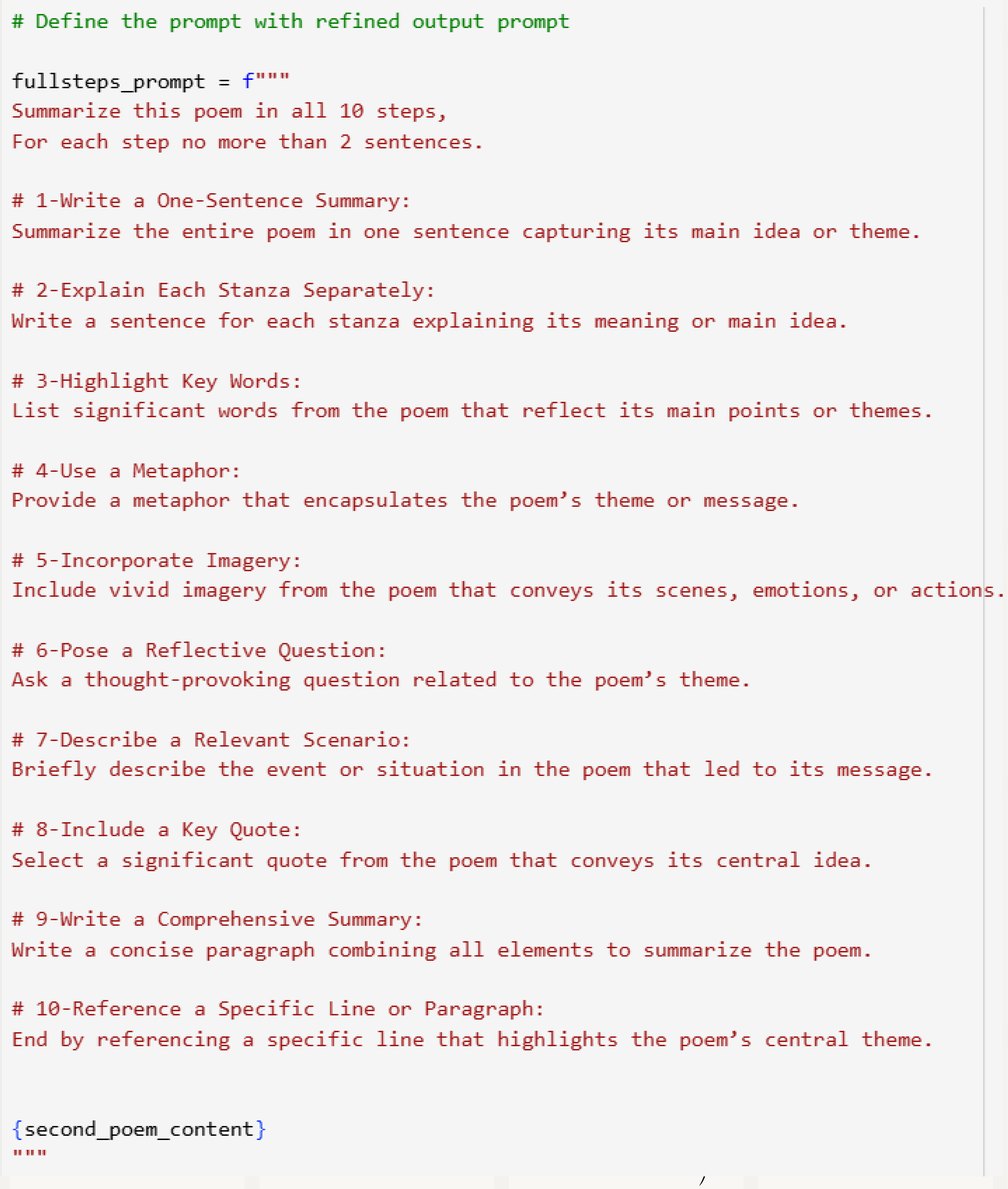
```
# Define the prompt with refined output prompt
refined_outputformat_prompt = f"""
Summarize this poem into a structured format using the following
headings: The poem's setting and the poem's story. Include only one sentence for each heading.
For each sentence only include 7 words.

1. Setting:

2. Story:

{poem_content}
"""
```

ANALYSIS PROCESS

Step 3: Apply progressive prompt refinements:

1. Basic summary (Cell 4).
2. Two-sentence summary (Cell 5).
3. Two-sentence summary with focus areas (setting, story) (Cell 6).
4. Structured format with 7-word sentences (Cell 7).
5. 10-step summarization methodology for two poems (Cells 8-11).

5.

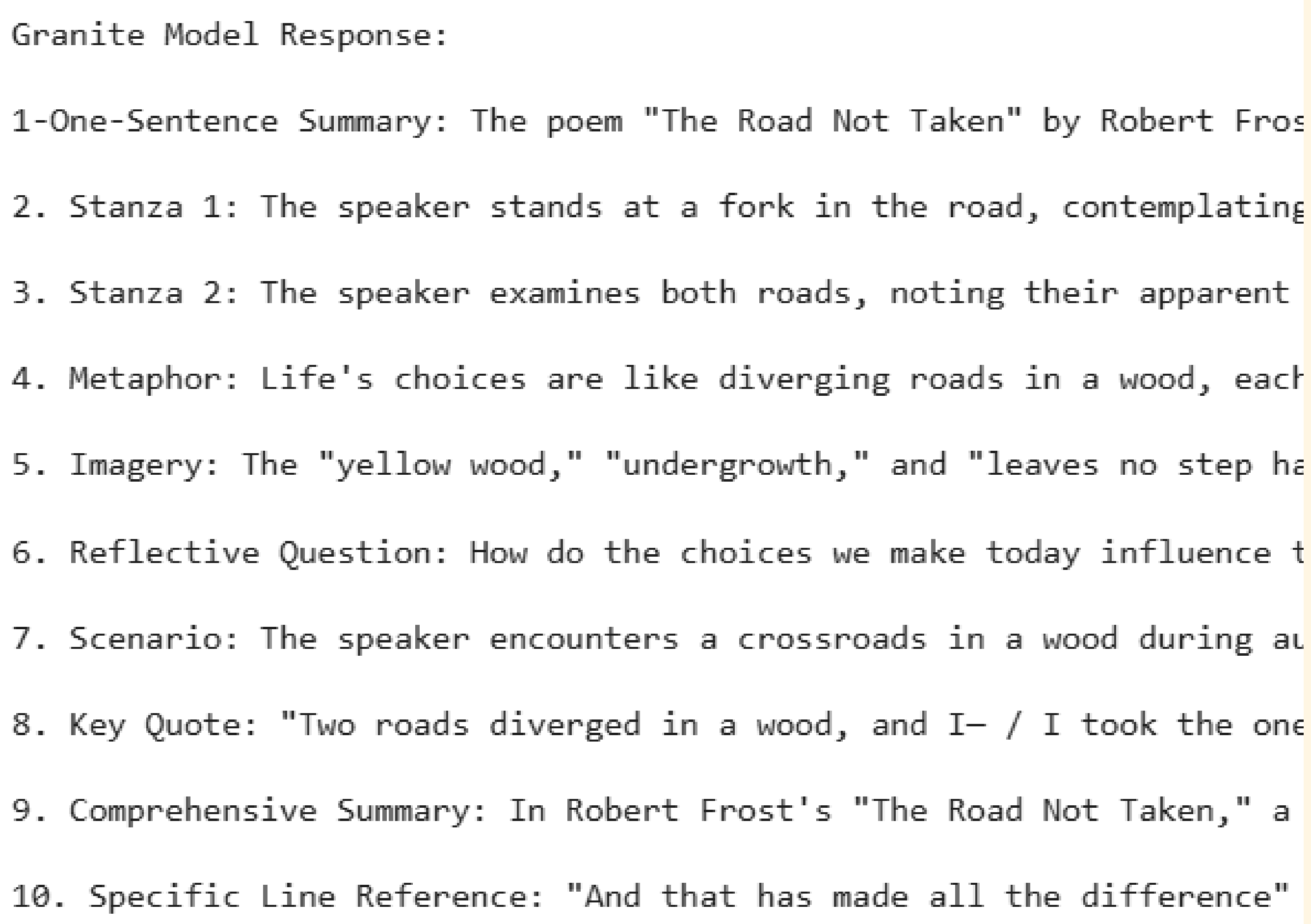


```
# Define the prompt with refined output prompt

fullsteps_prompt = f"""
Summarize this poem in all 10 steps,
For each step no more than 2 sentences.

# 1-Write a One-Sentence Summary:
Summarize the entire poem in one sentence capturing its main idea or theme.

# 2-Explain Each Stanza Separately:
Write a sentence for each stanza explaining its meaning or main idea.

# 3-Highlight Key Words:
List significant words from the poem that reflect its main points or themes.

# 4-Use a Metaphor:
Provide a metaphor that encapsulates the poem’s theme or message.

# 5-Incorporate Imagery:
Include vivid imagery from the poem that conveys its scenes, emotions, or actions.

# 6-Pose a Reflective Question:
Ask a thought-provoking question related to the poem’s theme.

# 7-Describe a Relevant Scenario:
Briefly describe the event or situation in the poem that led to its message.

# 8-Include a Key Quote:
Select a significant quote from the poem that conveys its central idea.

# 9-Write a Comprehensive Summary:
Write a concise paragraph combining all elements to summarize the poem.

# 10-Reference a Specific Line or Paragraph:
End by referencing a specific line that highlights the poem’s central theme.


{second_poem_content}
"""
```

INSIGHTS AND FINDINGS

Insights:

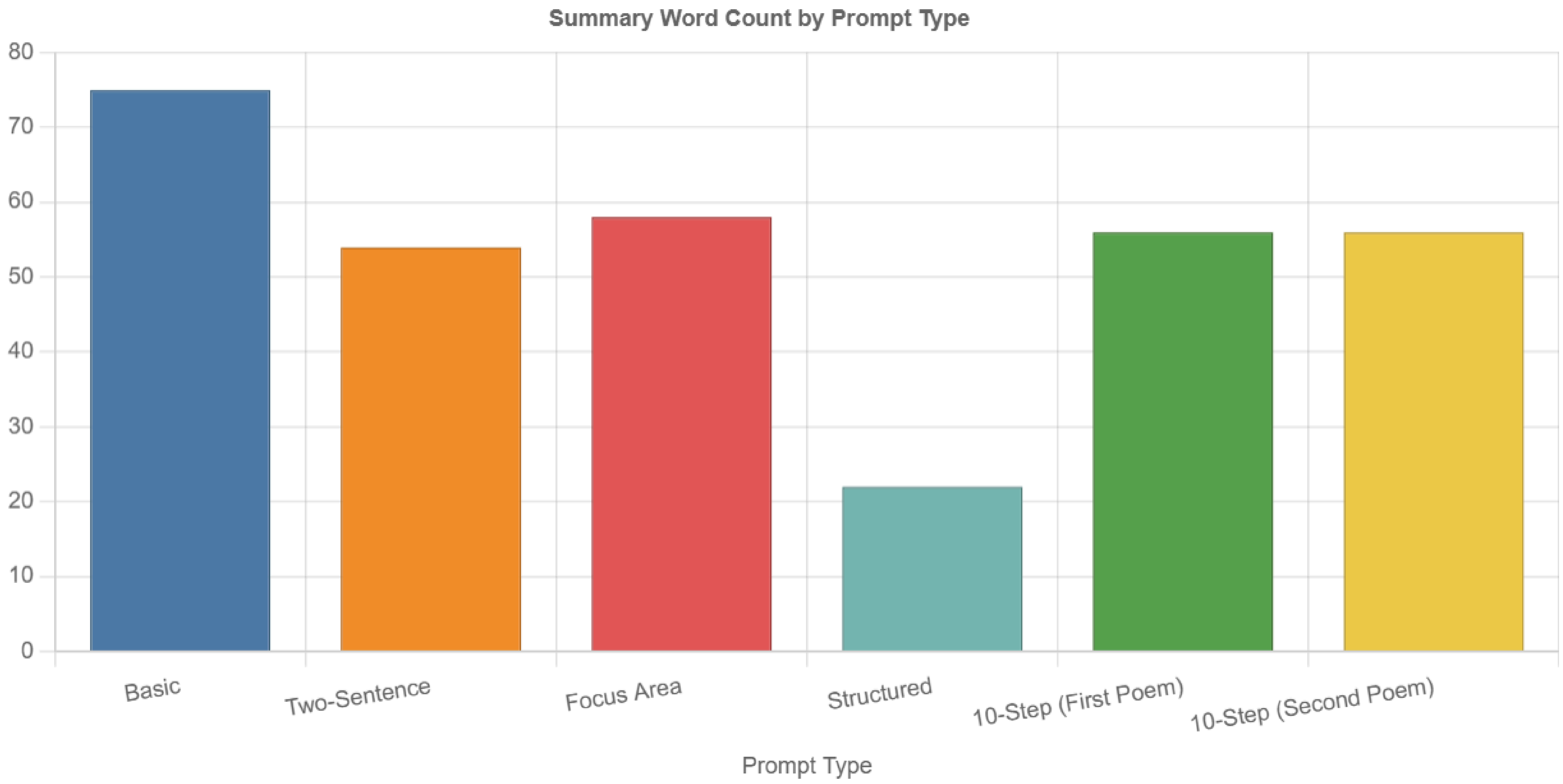
- The model accurately captures themes but elaborates beyond ground truth brevity for the first poem (Cells 4-7, 11).
- For "The Road Not Taken," the model aligns closely with the ground truth, emphasizing choice and its impact (Cell 10).
- The 10-step methodology provides deeper analysis (e.g., metaphors, imagery) but risks truncation with longer prompts (Cell 11, incomplete final line).
- Concise prompts with fewer context tokens improve output completeness



```
Granite Model Response:

1-One-Sentence Summary: The poem "The Road Not Taken" by Robert Fros

2. Stanza 1: The speaker stands at a fork in the road, contemplating

3. Stanza 2: The speaker examines both roads, noting their apparent

4. Metaphor: Life's choices are like diverging roads in a wood, each

5. Imagery: The "yellow wood," "undergrowth," and "leaves no step ha

6. Reflective Question: How do the choices we make today influence t

7. Scenario: The speaker encounters a crossroads in a wood during au

8. Key Quote: "Two roads diverged in a wood, and I— / I took the one

9. Comprehensive Summary: In Robert Frost's "The Road Not Taken," a

10. Specific Line Reference: "And that has made all the difference"
```

INSIGHTS AND FINDINGS

HACKTIV8



Insights & Findings

The IBM Granite-3.3B-Instruct model generates detailed poem summaries, with the 10-step methodology providing the deepest analysis but risking truncation. The bar chart shows that prompt complexity increases word count, with the Basic prompt (75 words) being the longest and the Structured prompt (22 words) the shortest. Concise prompts with fewer context tokens ensure complete outputs, especially for complex 10-step summaries.

Conclusions & Recommendations

Conclusion:

- The IBM Granite-3.3B-Instruct model effectively generates structured poem summaries, with the 10-step methodology offering the most detailed insights.
- Prompt refinement improves focus and brevity, but complex prompts risk truncation.

Recommendations:

- Use concise prompts with minimal context tokens for reliable outputs, especially for multi-step analyses.
- Explore further prompt optimization to balance depth and completeness.
- Apply the methodology to a larger subset of the dataset for broader validation.

AI Support Explanation

HACKTIV8

- **Model:** IBM Granite-3.3B-Instruct, accessed via Replicate API and LangChain (Cell 2).
- **Functionality:** Processes poem content to generate summaries in various formats, from single-sentence to 10-step analyses, capturing themes, imagery, and narratives.
- **Prompt Design:** Short, structured prompts (e.g., Cells 6, 7) ensure focused outputs; 10-step prompts (Cells 10, 11) provide comprehensive analysis but require token management to avoid truncation.
- **Strengths:** Handles complex literary analysis with vivid imagery and metaphorical interpretations, aligning well with ground truth summaries.

THANK YOU

HACKTIV8