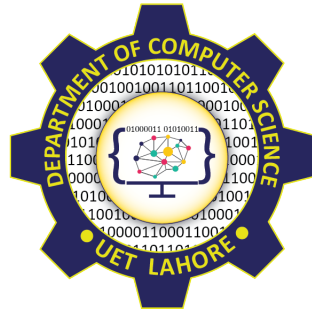


Gen AI Bootcamp 2024

Comparison Report: Zero-shot and Few-shot Chain-of-Thought (COT) Performance of Gemini and Mistral-7b-instruct-v0.1.Q2_K.gguf Models



Session 2022-2026

Submitted By

Shahbaz Ali 2022-CS-27

Submitted To

Mr. Muhammad Waseem

Department of Computer Science

**University of Engineering and Technology
Lahore, Pakistan**

Contents

1	Introduction	3
2	Zero-shot Chain-of-Thought (COT)	3
2.1	Gemini Output	3
2.2	Mistral-7b-instruct-v0.1.Q2_K.gguf Output	3
2.3	Comparison	4
3	Few-shot Chain-of-Thought (COT)	4
3.1	Gemini Output	5
3.2	Mistral-7b-instruct-v0.1.Q2_K.gguf Output	5
3.3	Comparison	6
4	Summary	6
5	Summary Table	7

1 Introduction

This report compares the performance of two AI models, Gemini and mistral-7b-instruct-v0.1.Q2_K.gguf, in solving arithmetic problems using Zero-shot and Few-shot Chain-of-Thought (COT) methodologies. The comparison is based on their responses to specific problem scenarios.

2 Zero-shot Chain-of-Thought (COT)

Input Question:

A store had 50 apples. They sold 20 apples and then received 30 more apples. How many apples do they have now?

2.1 Gemini Output

Gemini Output

Let's find out how many apples the store has now:
They started with 50 apples.
They sold 20, so they have $50 - 20 = 30$ apples left.
Then they received 30 more apples, so now they have $30 + 30 = 60$ apples.
Therefore, the store now has 60 apples.

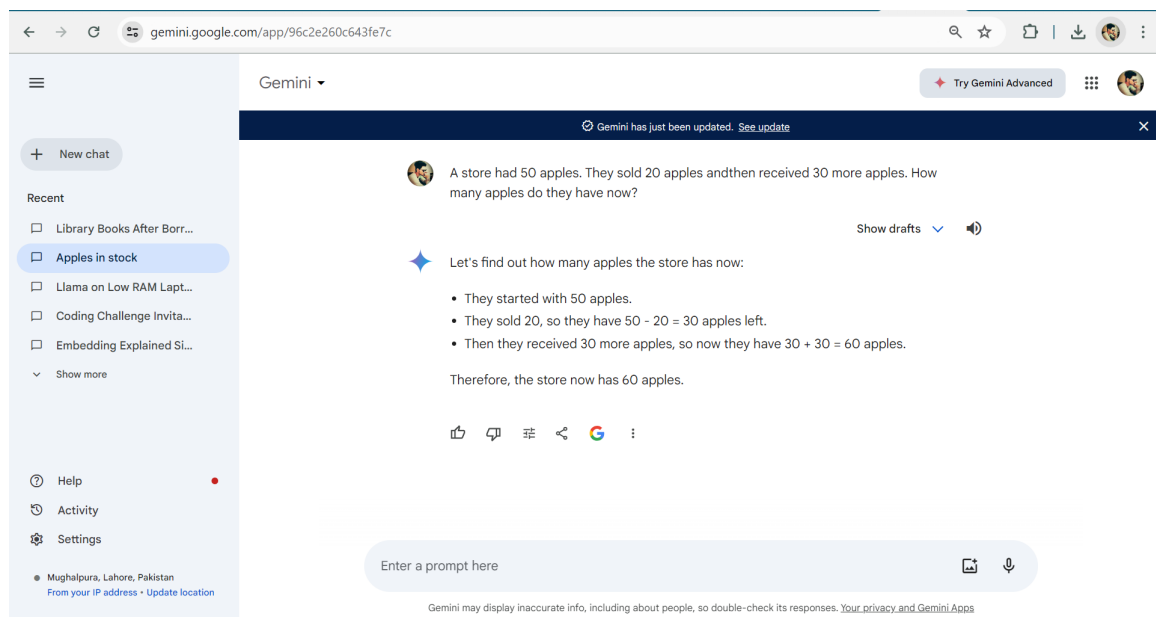


Figure 1: Gemini Prompt

2.2 Mistral-7b-instruct-v0.1.Q2_K.gguf Output

Mistral-7b-instruct-v0.1.Q2_K.gguf Output

The store now has 60 apples.

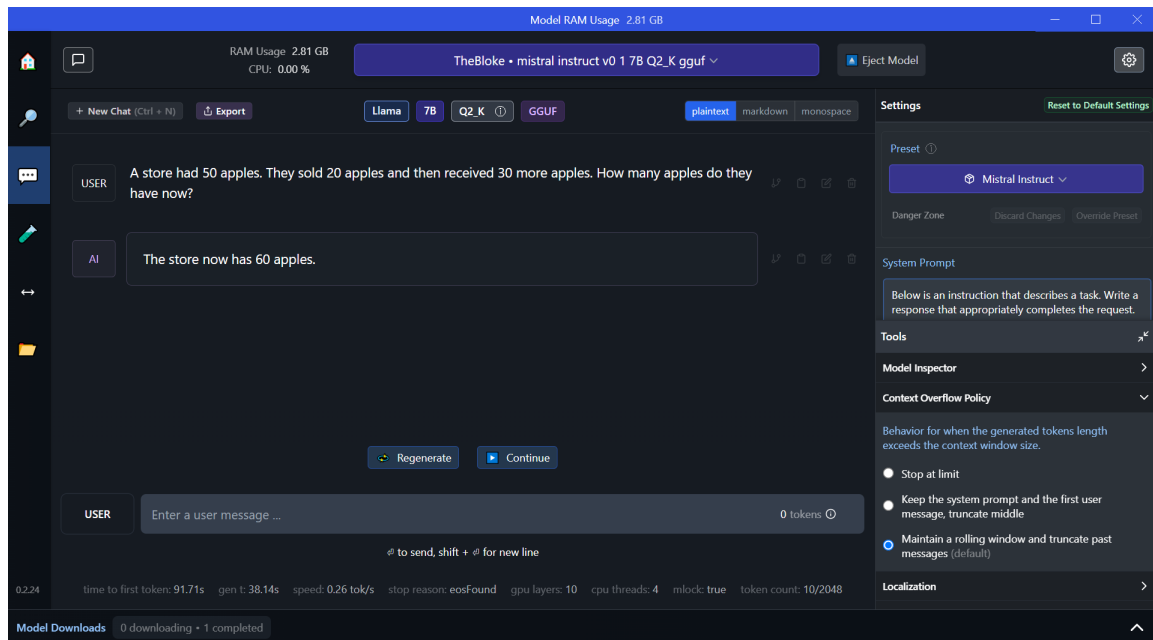


Figure 2: Mistral Prompt

2.3 Comparison

- **Detail and Clarity:** Gemini provides a detailed step-by-step explanation of how the final answer is derived, breaking down each step in the process. This approach is helpful for users who want to understand the logical progression of the solution.
- **Conciseness:** Mistral-7b-instruct-v0.1.Q2_K.gguf offers a concise answer directly. While it is accurate, it lacks the intermediate steps that demonstrate how the answer was calculated.

3 Few-shot Chain-of-Thought (COT)

Example Questions:

1. If a train travels 60 miles in 1 hour, how far will it travel in 3 hours?

A: Let's think this through. In 1 hour, the train travels 60 miles. In 3 hours, it will travel 60 miles/hour * 3 hours = 180 miles. So, the train will travel 180 miles in 3 hours.

2. There are 10 boxes with 5 apples each. How many apples are there in total?

A: We can break this down step by step. Each box has 5 apples. With 10 boxes, that's 10 * 5 = 50 apples in total.

New Question:

A library has 3 shelves with 20 books each. If 15 books are borrowed, how many books are left?

3.1 Gemini Output

Gemini Output

Following your examples, let's solve this problem step by step:

1. **Total Books:** The library has 3 shelves, and each shelf has 20 books. So, we need to find the total number of books on all shelves.
2. **Calculate Total:** We can multiply the number of shelves (3) by the number of books per shelf (20). $3 \text{ shelves} * 20 \text{ books/shelf} = 60 \text{ books}$.
3. **Borrowed Books:** We know 15 books are borrowed.
4. **Books Left:** To find the remaining books, we need to subtract the borrowed books from the total number of books. $\text{Total books (60)} - \text{Borrowed books (15)} = \text{Books left}$.

Therefore, there are $60 \text{ books} - 15 \text{ books} = 45 \text{ books}$ left in the library.

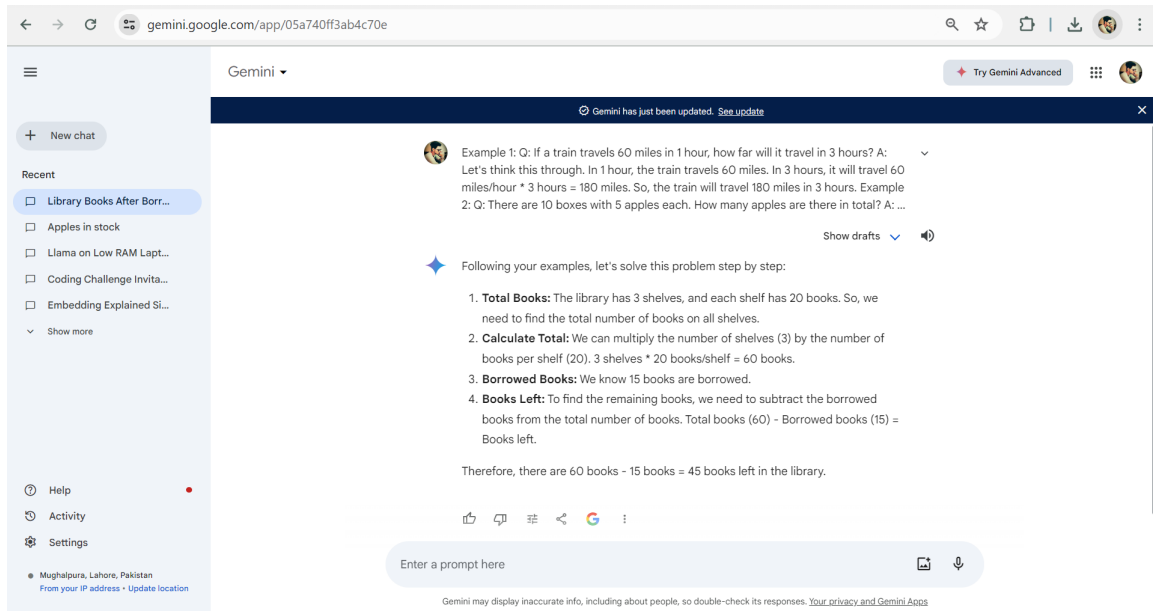


Figure 3: Gemini Output

3.2 Mistral-7b-instruct-v0.1.Q2_K.gguf Output

Mistral-7b-instruct-v0.1.Q2_K.gguf Output

First, let's calculate the total number of books that were initially on the shelves. We know there are 3 shelves and each shelf has 20 books, so:

Total books = Number of shelves * Books per shelf
 $= 3 * 20$
 $= 60 \text{ books}$

Next, we need to subtract the number of borrowed books from the initial total. We're told that 15 books were borrowed, so:

Remaining books = Total books - Borrowed books
 $= 60 - 15$
 $= 45 \text{ books}$

So, there are 45 books left in the library after 15 have been borrowed.

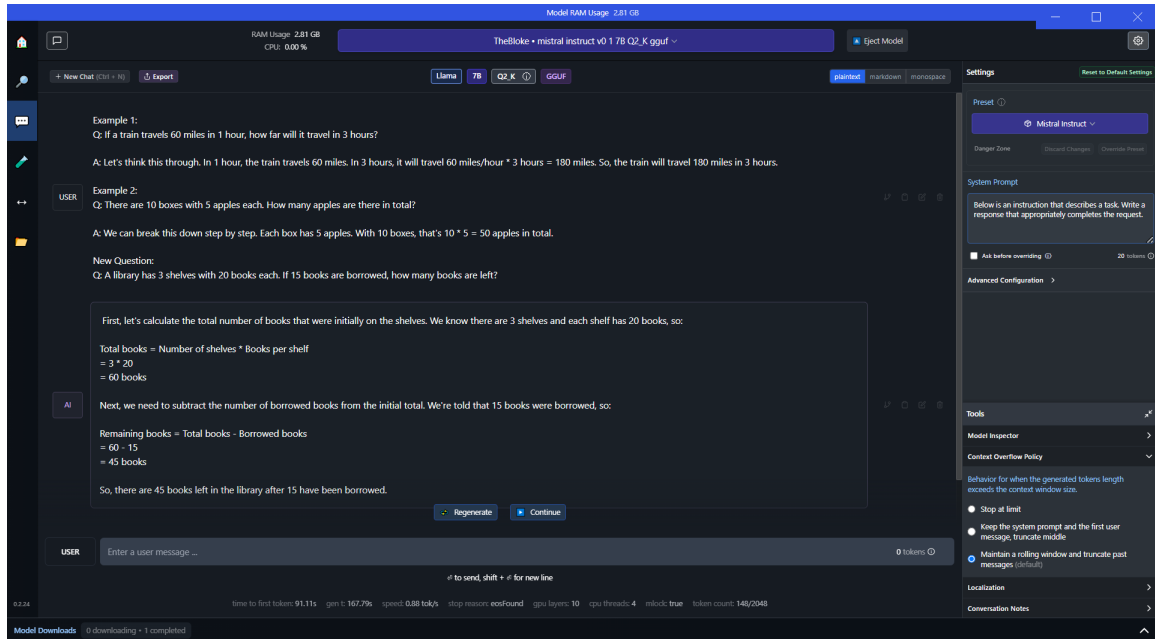


Figure 4: Mistral Output

3.3 Comparison

- **Detail and Clarity:** Both models provide detailed solutions, breaking down the problem into logical steps. This detailed explanation ensures that the user can follow the reasoning behind the final answer.
- **Step-by-Step Process:** Gemini and Mistral-7b-instruct-v0.1.Q2_K.gguf both follow a step-by-step approach, ensuring clarity in their problem-solving methods. Gemini adds structured bullet points to emphasize each step, while Mistral-7b-instruct-v0.1.Q2_K.gguf uses a narrative format.
- **Conclusion:** Both models arrive at the correct answer, but Gemini's structured format may be slightly easier to read for users who prefer clearly separated steps.

4 Summary

- **Zero-shot COT:** Gemini excels in providing detailed, step-by-step explanations, making it easier for users to understand the reasoning process. Mistral-7b-instruct-v0.1.Q2_K.gguf, while concise and accurate, does not provide the intermediate steps, which may be less educational for users needing the full reasoning process.
- **Few-shot COT:** Both models perform well, offering detailed and clear explanations. Gemini's use of bullet points for each step can enhance readability, while Mistral-7b-instruct-v0.1.Q2_K.gguf's narrative style remains clear and coherent.

Overall, Gemini is better suited for users who prefer detailed, step-by-step solutions, while Mistral-7b-instruct-v0.1.Q2_K.gguf is ideal for those who prefer concise answers with clear, logical explanations. Both models demonstrate strong capabilities in understanding and solving arithmetic problems through Chain-of-Thought reasoning.

5 Summary Table

Criteria	Gemini	Mistral-7b-instruct-v0.1.Q2_K.gguf
Detail and Clarity	High	Medium
Conciseness	Medium	High
Step-by-Step Process	Yes	Yes
User-Friendly	High	Medium

Table 1: Comparison of Gemini and Mistral-7b-instruct-v0.1.Q2_K.gguf Models