In general Llama models need more cues for generating the expected answer as compared to gpt4. Running times are: Alpaca 7B ~5 min, Alpaca 30B ~30min, GPT-4 seconds.

**Llama models:**

**Parameters:**

debug

n\_predict

repeat\_last\_n

repeat\_penalty

top\_k

top\_p

temp

seed

threads

* **If the input is not circumscribed in a context, the output is not coherent:**

Input:

some users wrote the following comments: ...@user1: I don't like the cheesecake.@user2: I like their new coffee....what are the users talking about?

Alpaca 7B:

The user data is stored in a database, and we want to analyze what kinds of things people are saying. We can use Natural Language Processing (NLP) techniques on this dataset to identify topics that come up often in conversations between different users. For example: "I don't like the cheesecake", or "@user2: I like their new coffee". The output should include a list of topic words, such as "cheese" and "coffee". The algorithm can then use these topics to categorize conversations into categories for analysis. [end of text]

Alpaca 30B:

Given a string and an integer k, determine if the string contains a substring of length at least k that is also palindrome. A string S is called palin­dromic if it can be divided into two non-empty substrings X =

(…still going after 1 hour)

* **With moderate formatting, the output is better:**

Below is an instruction that describes a task, paired with an input that provides further context. Write a response that appropriately completes the request.

### Instruction:

some users wrote the following comments: ...@user1: I don't like the cheesecake.@user2: I like their new coffee....what are the users talking about?

### Response:

Alpaca 7B:

The users were commenting on a newly released product, specifically Cheesecake and Coffee. User 1 was not in favor of it while user 2 had a more positive opinion towards it. Event-driven Architecture for IoT Systems [end of text]

Alpaca 30B:

The users were commenting on a particular cafe or restaurant. User 1 did not enjoy the cheesecake and user two enjoyed their new brew of coffee. [end of text]

* **More formatted input does not generate significantly better outputs:**

Below is an instruction that describes a task, paired with an input that provides further context. Write a response that appropriately completes the request.

### Instruction:

some users wrote the following comments:

...

@user1: I don't like the cheesecake.

@user2: I like their new coffee.

...

what are the users talking about?

### Response:

Alpaca 7B:

The users are discussing a product called "CheeseCake", and how they feel about it. They're also commenting on the quality of the coffee that is being offered by this company. [end of text]

Vicuna 7B:

Below is an instruction that describes a task, paired with an input that provides further context. Write a response that appropriately completes the request.\

### Instruction: \

some users wrote the following comments: \

... \

@user1: I don't like the cheesecake. \

@user2: I like their new coffee. \

... \

what are the users talking about? \

### Response:

**GPT4:**

User input:

some users wrote the following comments:

...

@user1: I don't like the cheesecake

@user2: I like their new coffee

...

what are the users talking about?

GPT4 answer:

The users are discussing their opinions about certain aspects of a food or drink establishment, most likely a café or restaurant. User1 is expressing their dislike for the cheesecake served there, while User2 is sharing their appreciation for the establishment's new coffee offering.

User input:

reflect on your answer

GPT4 answer

Upon reflecting on my previous answer, I still believe that the users are discussing their opinions about certain aspects of a food or drink establishment, such as a café or restaurant. The comments focus on specific items offered at the location: User1 mentions the cheesecake, and User2 talks about the new coffee. Both comments express personal preferences, indicating that the users are sharing their experiences and thoughts on the offerings of the establishment in question.