

Prompt 1

Prompt: “Once upon a time, in a land far far away, lived a young princess who loved her kingdom very much.”

Response: Once upon a time, in a land far far away, lived a young princess who loved her kingdom very much. The world that she loved was different from her kingdom. When the princess died, her name was the Queen. This was the beginning of her reign. In the end, this was a world she saw to be destroyed. Even now, in this land far away, lived the Princess who had brought peace and prosperity to the kingdom. The Princess was the princess who lived to see her day on. Today, she lives alone. The Princess lived. Today, she had to leave this land and travel to the other world. And she had to live with her family to get there. And she had to face the danger of the other world with a certain face. The princess did this. The Princess lived. And now, she is home. In the next world... the future. She lived. She was born. After a short time, a boy named Goh, whose name was Gokhama, left her to become a princess. But that was the beginning of the end. Gokhama was called Gokhama of Yotsu. She was a woman who

Settings: max_tokens=100, temperature=1.0, top_p=0.9

The reaction is first relevant and captivating due to its appropriate fairytale tone. However, as the text gets repetitious and includes disjointed or contradictory plot details (e.g., "the princess died" followed immediately by her restoring peace to the realm), the coherence rapidly deteriorates. Although the identities and events are illogical at the end, there are no detrimental biases or factual errors. Though it also made the narrative less controlled, the circumstances, particularly the high temperature (1.0), obviously fostered imaginative storytelling. This demonstrates how temperature may foster creativity at the expense of reason and order.

Prompt 2

Prompt: Hi John, I'd like to arrange a meeting for Tuesday

Response: Hi John, I'd like to arrange a meeting for Tuesday between you and the other members of our council to discuss how we can proceed in future. I understand you're quite busy, and you have to be in the midst of meetings, but you're an absolute pleasure to have the time to talk to me and talk to the other council members. I'd also like to make the point that you know about our proposal to make a referendum on our new state bill, and that's going to include all the people who voted for you as you've voted on it. I think the time to act is now. If we can't pass it in the next week, we'll be left

with only the details of what we're going to do about it. But in the meantime, here's a quick summary. We're going to vote on a referendum that says that, under no circumstances should we allow any more children to go to school, whether in their private school or at a public school in Ontario. We're going to ban all private school admissions at all public schools. We're going to ban any school board from discriminating in any way on students based on their gender identity. We're going to ban any new, full-time job openings, from the schools which have refused to allow more

Settings: max_tokens=50, temperature=1.0, top_p=0.9

At first, the response remains pertinent and appears to be the start of a professional email, but it soon veers off course and becomes political commentary about public policy and education that has nothing to do with the prompt. This change weakens the response's coherence and causes confusion. Even while there aren't any overt factual errors, the sudden political turn could unintentionally create bias, particularly when using strong language like "banning school admissions." This unexpected tone might have been caused by the usage of `top_p=0.9` and `default temperature=1.0`, indicating that lower temperature or limited sampling is more suitable for structured communication tasks.

Prompt 3

Prompt: The average heart pumps 2000 gallons of blood everyday

[illegible]

Settings: max_tokens=100, temperature=0.5, top_p=0.9

The response rapidly loses consistency as it starts repeating language concerning the city's health department and doctors, even though the first sentence is factually correct and on subject. This recurrence reduces the usefulness of the response and makes the user experience annoying. Although there aren't any blatant errors, the value is limited by the inability to stay on topic. Degeneration was not stopped by the lower temperature setting (0.5), which was meant to promote stability and factual consistency. This could have been caused by model limits rather than parameter tuning. It demonstrates that the shortcomings of smaller models in managing factual continuity cannot be entirely made up for, even in conservative conditions.

Prompt 4

Prompt: "Hey Lisa how are you?"

Response: "Hey Lisa how are you?" Lisa looked up towards the sky with wide golden eyes and said, "There's some really, real monsters out there out there, right Lisa? Why would I give an explanation?" Lisa said: "Look, it can't be anything personal. It's not like we'll ever know what it is. It's all about the monster. What are you doing, trying to make sense out of everything at once? Like maybe you're really a magician and want to see what it shows you. Or maybe you want to see it like you're part of a team working to see if there's anything you can do with a weapon that's an old fashioned sword. Or perhaps you see it with you in your past." I'm not gonna be able to see where they're going, so let's see how many other things might lie between us." After I turned around, Lisa said, "When are you going to go back to class? If in a future time they find out you're the last of your kind, they can stop working for a while and you, too, can quit. They could put on some form of an armor-piercing collar about your neck. Or they could make you come with some sort of

Settings: max_tokens=100, temperature=1.0, top_p=1.2

After the initial greeting, the tone of this response drastically changes, devolving into a fantasy-like conversation involving monsters and magical weapons. Although the writing is creative, it mostly lacks coherence in respect to the challenge, which called for a conversational or informal answer. Although the content is not offensive or biased, it is useless for practical activities due to its lack of topical relevance. It's interesting to note that a top_p number larger than 1 (invalid, but probably regarded as max) and a high temperature (1.0) produced incredibly imaginative but disorganized results. This demonstrates how aggressive sampling can skew answers to straightforward questions.

Prompt 5

Prompt: Explain photosynthesis like I'm five.

Response: Explain photosynthesis like I'm five. You'll also have a lot more information about some of the best things to do in your day. We're gonna be a little late this weekend.

Settings: max_tokens=200, temperature=1.0, top_p=0.9

The comment makes no effort to explain photosynthesis and is wholly off-topic. Rather, it dives right into vague, useless remarks about "doing things this weekend," rendering it completely irrelevant. Since it doesn't even touch on the subject, there aren't any factual inaccuracies. Although the reason for the model's failure is unknown, it's probable that the prompt's tokenization confused the generating process. Even with the right settings (temperature=1.0, top_p=0.9), the model's failure points to a problem with task comprehension or decoding rather than a problem with the sampling setup. This emphasizes the necessity of reruns or fallback handling in the application.

Reflections

Prompts that support open-ended creative generation, such narrative or expressive discourse, are the most effective for the approach. For instance, even though coherence eventually faded, the model's response to the fairytale-style suggestion about the little princess had a rich narrative tone and creative structure. In situations where logical consistency is not strictly necessary, these examples demonstrate the model's ability to produce fluid, stylistically suitable language. Its capacity to generate intricate and vibrant fictional content is further demonstrated by the bizarre response to "Hey Lisa, how are you?" This can be useful in creative writing circumstances.

But it's obvious that the model has trouble with cues that call for structured reasoning, practical experience, or specialized task comprehension. It completely failed to stay on topic in the "photosynthesis" assignment, providing an answer that was not only superficial but completely unrelated. This implies that there may be trouble understanding particular teaching cues. Additionally, it struggled to stay coherent when producing longer factual or procedural responses, such those in the business communication question or the heart-related fact prompt, which both drifted into repetitious or off-topic content. With default parameters and no further safeguards, these generative models still struggle with logical reasoning, especially when factual accuracy or consistency is crucial.

A number of tactics could be used to enhance the application. When responses deviate too much from the original inquiry, it may be possible to identify this by filtering outputs using relevance score or keyword checks. Reliability would also be increased by including a straightforward retry option in case outputs are too brief, irrelevant, or empty (as in the case of photosynthesis). Controlling parameters, such as setting a temperature limit for factual prompts or providing preset modes (such as "creative," "formal," or "explanatory"), could help users expect better outcomes. Most significantly, adding levels of fact-checking or summary validation—possibly with the aid of an external model—would lessen hallucinations and give users more reliable content, particularly for business or educational purposes.