

Krish Arora - AI Powered Text Completion

Evaluate and Analyze Outputs

1. Creative

- Prompt: “Continue this story: The air suddenly turned magical and...”
- Settings:
 - temperature: 0.9
 - max_tokens: 100
 - top_p: 1.0
- Response: “everything around me seemed to come to life. The trees swayed gently in the breeze, their leaves shimmering with an otherworldly glow. The birds sang in harmonious melodies that filled the air with a sense of love and harmony. I felt a warmth wash over me, as if I had been enveloped in a hug from the universe itself. As I looked around, I saw flowers blooming in vibrant colors that I had never seen before. The sun seemed to shine brighter, casting a”
- Evaluation:
 - Relevance: Directly continues the narrative (high).
 - Coherence: Smooth flow with vivid sensory detail.
 - Biases/Inaccuracies: None observed.
 - Tone/Depth: Higher temperature increased creativity and emotion.

2. Informational

- Prompt: “Photosynthesis is used by plants to...”
- Settings:

- temperature: 0.3
- max_tokens: 60
- top_p: 1.0
- Response: “convert sunlight, water, and carbon dioxide into oxygen and glucose (sugar) for energy.”
- Evaluation:
 - Relevance: Directly answers the prompt clearly and accurately (high).
 - Coherence: Clear and concise with no confusion.
 - Biases/Inaccuracies: None observed.
 - Tone/Depth: Factual and straightforward, matching the low temperature setting.

3. Instructional

- Prompt: “Explain recursion like I’m five...”
- Settings:
 - temperature: 0.6
 - max_tokens: 80
 - top_p: 1.0
- Response: “Recursion is when you do something over and over again, but each time you do it, you make the problem a little bit smaller. It's like when you keep putting dolls inside of each other, but each doll is a little bit smaller than the one before.”
- Evaluation:
 - Relevance: Directly explains recursion in simple terms (high).
 - Coherence: Clear and easy to understand with a helpful analogy.
 - Biases/Inaccuracies: None observed.

- Tone/Depth: Friendly and accessible, appropriate for a young audience with moderate temperature.

4. Science

- Prompt: “Explain quantum computing in simple terms...”
- Settings:
 - temperature: 0.5
 - max_tokens: 100
 - top_p: 0.95
- Response: “Quantum computing is a type of computing that uses quantum mechanics to perform calculations. In traditional computers, information is processed in bits, which can be either a 0 or a 1. In quantum computing, information is processed in quantum bits, or qubits, which can be both 0 and 1 at the same time. This allows quantum computers to perform calculations much faster and more efficiently than traditional computers.

Quantum computing has the potential to revolutionize fields such as cryptography, drug discovery, and”
- Evaluation:
 - Relevance: Explains quantum computing clearly and accurately (high).
 - Coherence: Mostly smooth, but the last sentence is incomplete, leaving the explanation unfinished.
 - Biases/Inaccuracies: None observed.
 - Tone/Depth: Informative and accessible, balancing technical detail with simplicity.

5. Controversial Topic

- Prompt: “Why do people disagree on climate change...”
- Settings:
 - temperature: 0.5
 - max_tokens: 100
 - top_p: 0.95
- Response: “There are a variety of reasons why people may disagree on climate change. Some of the main reasons include: 1. Political ideology: Climate change has become a highly politicized issue, with people often aligning their beliefs on the topic with their political leanings. Conservatives tend to be more skeptical of climate change, while liberals tend to be more accepting of the scientific consensus on the issue. 2. Economic interests: Some individuals and industries may have a financial stake in denying or downplaying the reality of climate change, as it could potentially lead to regulations or policies that could impact their bottom line. 3.”
- Evaluation:
 - Relevance: Directly addresses reasons for disagreement on climate change (high).
 - Coherence: Generally clear, but the response ends abruptly and feels incomplete.
 - Biases/Inaccuracies: Slight political generalization.
 - Tone/Depth: Balanced and informative, maintaining a neutral tone appropriate for sensitive topics.

Reflect on Limitations

The model performs well in creative writing and storytelling, especially when higher temperature settings allow for vivid imagery and emotional expression. It also excels at

providing clear, concise explanations of common concepts, using accessible analogies that make complex ideas easier to understand. Additionally, the model handles summarization and straightforward factual responses effectively when configured with lower temperature and focused parameters. However, the model struggles with completing longer responses fully, often cutting off mid-sentence, and can have difficulty with logical reasoning or multi-step problem-solving, sometimes producing inconsistent or incorrect answers. Niche or highly technical topics may be oversimplified or lack important nuances, and responses to controversial or politically sensitive subjects can occasionally reflect subtle biases or generalizations. To improve the application, it would be beneficial to implement output validation or fact-checking mechanisms, especially for technical or sensitive content, and to apply prompt engineering techniques to encourage more complete and accurate responses. Offering users preset modes for different tones or levels of detail can help tailor outputs to specific needs, while logging session data can assist in identifying failure points for iterative improvement. Finally, integrating filters or moderation tools can help manage inappropriate or biased content, ensuring a safer and more reliable user experience.