

Testes de Temperatura

1. Introdução

Este documento tem como objetivo avaliar o impacto da variável de temperatura na qualidade e coerência das respostas do chatbot. Foram testadas várias configurações de temperatura para encontrar o valor que oferece um equilíbrio ideal entre coerência e criatividade nas respostas.

2. Metodologia

Perguntas padrão:

“Qual é o melhor notebook?”

“Me fale sobre um livro bem avaliado”

“Qual produto as pessoas não recomendam?”

“Indique produtos para um atleta.”

“Fale sobre produtos que apresentam defeitos.”

Configurações de temperatura: Os valores de temperatura que foram testados variam entre 0.1 e 1.0.

Critérios de avaliação: Coerência, Relevância, nível de detalhamento e variedade de respostas.

3. Análise dos testes

Intervalo de temperatura de 0.1 a 0.4

Nessas temperaturas, a IA adota uma abordagem cautelosa, destacando a subjetividade das perguntas. Em uma pergunta como “Qual é o melhor notebook?”, a IA ressalta que a resposta depende das necessidades do usuário, e então descreve opções com base em critérios como uso, desempenho e preço. Esse cuidado em contextualizar antes de recomendar resulta em respostas que ajudam o usuário a refletir sobre as opções.

Liv: A pergunta “Qual é o melhor notebook?” é complexa e a resposta depende do seu uso e necessidades. A análise dos reviews permite ter uma ideia geral, mas não há um “melhor” absoluto.

O “notebook Asus Vivobook X510UABR540T Intel Core i5” é bem avaliado por sua rapidez, tela de tamanho ótimo e boa duração da bateria, ideal para estudo e trabalho, apesar de não atingir as 8 horas prometidas. Já o “notebook Lenovo Ideapad 320 Intel® Core i77500U” se destaca pelo ótimo custo-benefício, ideal para tarefas menos exigentes e com boa qualidade.

Por outro lado, notebooks como o “Acer Aspire 57506 BR858 BR” e o “HP Pavilion G41000” também são bem avaliados em relação ao teclado, mas não há informações suficientes para compará-los aos modelos anteriores.

Vale lembrar que alguns produtos, como a “bateria para notebook Acer Aspire 5810TZ 6 células”, tiveram reviews negativas, com relatos de defeitos e devoluções.

Para escolher o melhor notebook para você, é essencial definir suas necessidades e prioridades. Se você precisa de um notebook para uso intenso, como jogos ou edição de vídeo, notebooks com processadores mais potentes e placas de vídeo dedicadas são recomendados. Se você precisa de um notebook para uso casual, como navegação na Internet e tarefas básicas, notebooks mais simples e com preços mais baixos podem ser suficientes.

Lembre-se de pesquisar outros reviews e comparar preços antes de fazer sua escolha.

Similaridade nas respostas: Em perguntas como “Me fale sobre um livro bem avaliado”, a IA recomenda os mesmos livros em todos os testes, indicando uma consistência nas recomendações.

User: Me fale sobre um livro bem avaliado

Liv: O livro “Assim Acaba” da Colleen Hoover é muito bem avaliado pelos leitores. A entrega é rápida e o produto chega em ótimas condições, além de ser lindamente embalado. A história é considerada excepcional, com muitos destacando que é o melhor livro que já leram. A escrita da autora é elogiada, e o livro é descrito como envolvente, despertando a vontade de continuar a leitura até o final. O enredo é emocionante e, para muitos, vale a pena cada página. Se você gosta de romances envolventes e emocionantes, “Assim Acaba” é uma ótima pedida.

Esclarecimento de contexto: A IA busca mais informações ou oferece respostas diversificadas, promovendo uma experiência útil para perguntas amplas.

Qualidade das recomendações: Em questões como “Indique produtos para um atleta”, algumas respostas fugiram do contexto esportivo, recomendando itens como protetores de canto, kits de piscina e até bandanas temáticas. Essa escolha evidencia uma limitação na compreensão do contexto esperado.

****Bandana Naruto:**** A bandana é um acessório divertido para os fãs do anime, mas é importante verificar se as fotos da loja correspondem ao produto que será enviado. Apesar da boa qualidade do produto, alguns compradores reclamaram de receberem um modelo diferente do que esperavam.

****Kit reparo para piscina Nautika:**** Este kit é ideal para manter a piscina em ordem e evitar transtornos. O produto é bem avaliado pela qualidade e a entrega é rápida. É uma ótima opção para quem busca praticidade e eficiência na manutenção da piscina.

****Ativador de cachos:**** Se você está em busca de um presente para uma mulher que cuida dos cabelos, este ativador de cachos pode ser uma boa opção. Os usuários elogiam o produto, dizendo que ele é excelente para definir e hidratar os cachos. A entrega é rápida e o produto é recomendado por muitos.

****Recomendações:****

É fundamental pesquisar outros reviews e comparar preços antes de comprar qualquer produto. Vale lembrar que a experiência de cada pessoa com o produto pode variar, por isso é importante ler diversas opiniões para ter uma ideia mais completa.

****Dica extra:****

Para encontrar produtos específicos para atletas, como suplementos alimentares, roupas esportivas e

Intervalo de temperatura de 0.5 a 0.7

Em temperaturas médias, a IA começa a simplificar a abordagem. As respostas ainda são informativas, mas há uma menor tendência a qualificar a subjetividade das perguntas. Para "Qual é o melhor notebook?", a IA fornece uma lista de opções com algumas características, sem detalhar tanto as nuances de escolha pessoal.

User: Qual é o melhor notebook?

Liv: A pergunta "Qual é o melhor notebook?" é muito subjetiva, pois depende do que você procura em um notebook. Para te ajudar, vamos analisar as avaliações dos produtos:

- **Acer Aspire 57506 BR858 BR:**** O teclado tem ótimo acabamento e qualidade, mas há dúvidas se é original.
- **Acer 4251 5736z 4551 5551 5251 5741 e1431 e1471 e1521 e1531 e1571 AS10D31:**** A bateria está funcionando, mas ainda é cedo para avaliar sua durabilidade.
- **LG C400 A410:**** O teclado atendeu às expectativas, cumprindo o prometido na descrição.
- **Acer 19v 342a LA1148:**** A fonte não é original, esquenta um pouco, mas funciona.
- **Acer PA165080 PA165086 19v 342a:**** A fonte demorou a chegar e não funcionou.
- **HP Pavilion G41000:**** O teclado está funcionando perfeitamente até o momento.
- **ASUS X450L C:**** O teclado tem ótima qualidade, idêntica ao original.
- **Acer Aspire 5810TZ:**** A bateria apresentou defeito e foi devolvida.

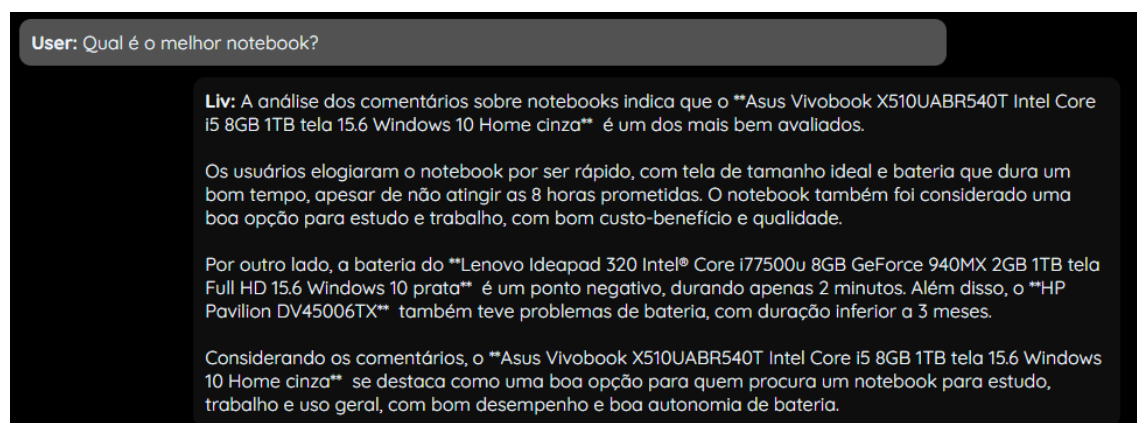
Diferença na abordagem: A IA ainda recomenda os mesmos produtos em diversas respostas, mas com menos ênfase em explicar por que uma escolha pode ser subjetiva.

Menor consistência contextual: A IA mantém a relevância, mas começa a perder o foco nas nuances que perguntas amplas exigem.

Contexto moderado: A IA ainda responde com consistência, mas perguntas mais complexas podem não receber uma análise tão aprofundada quanto nas temperaturas mais baixas.

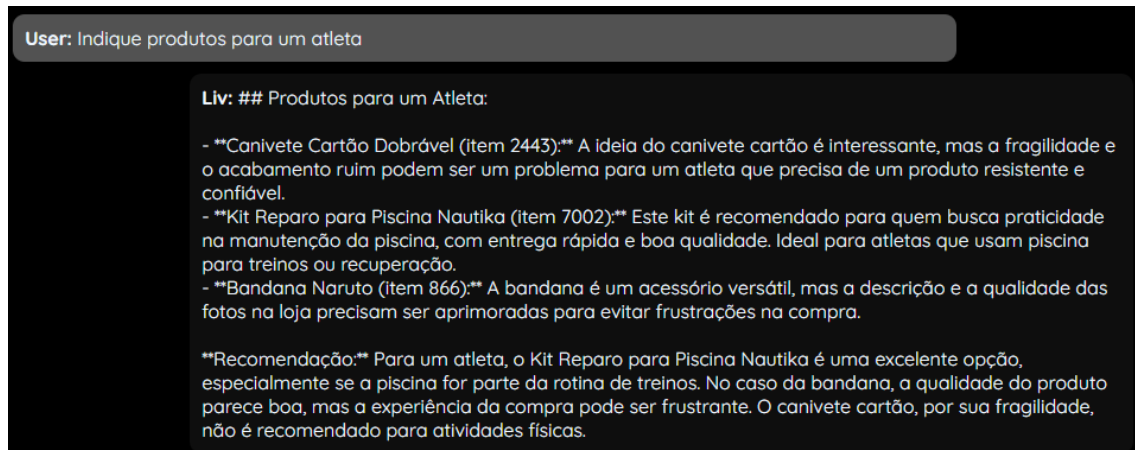
Intervalo de temperatura de 0.8 e acima

Em temperaturas mais altas, a IA busca oferecer respostas diretas e objetivas, o que resulta em respostas que ignoram nuances e subjetividades. Para perguntas amplas como “Qual é o melhor notebook?”, a IA recomenda modelos específicos sem contextualizar as necessidades do usuário. A objetividade excessiva pode levar a respostas que soam certeiras, mas não respondem plenamente à questão.



Simplificação excessiva: A IA, em vez de considerar nuances, fornece respostas diretas, o que nem sempre é ideal para recomendações complexas.

Falta de alinhamento com o contexto: Em perguntas como “Indique produtos para um atleta”, a IA sugere itens que pouco têm a ver com o esporte, como bandanas e kits de reparo para piscina, prejudicando a relevância.



Consistência inferior: A IA tende a recomendar produtos aleatórios ou menos pertinentes, o que pode frustrar as expectativas do usuário.

4. Conclusão

As respostas mostram que temperaturas mais baixas (0.1 a 0.4) são mais eficazes para perguntas que exigem análise e contextualização. Elas preservam nuances e fornecem recomendações úteis e personalizadas. Temperaturas médias (0.5 a 0.7) ainda são eficazes, mas podem não ter o mesmo cuidado ao destacar a subjetividade. Por outro lado, temperaturas mais altas (0.8 e acima) simplificam as respostas e, embora possam ser úteis em perguntas muito diretas, não se ajustam bem a perguntas amplas, falhando em capturar o contexto desejado.