**Processo de Seleção e Preparação do Dataset**

O dataset utilizado foi o AmazonTitles, composto por avaliações de produtos da Amazon, totalizando aproximadamente 2,3 milhões de registros. Durante o processo de preparação e limpeza dos dados, foram aplicados os seguintes procedimentos:

* **Remoção de campos ausentes:** Registros com informações incompletas foram descartados.
* **Exclusão de caracteres especiais:** Garantiu-se a padronização do texto.
* **Normalização dos dados:** Os dados foram ajustados para um formato consistente e uniforme.

Após a aplicação desses filtros, o número total de registros foi reduzido para aproximadamente 1,3 milhões, resultando em um dataset mais limpo e adequado para o processo de fine-tuning.

**Fine-Tuning do Modelo**

O modelo escolhido foi o unsloth/Meta-Llama-3.1-8B, com o processo de fine-tuning seguindo o exemplo na documentação do unsloth. A maioria das configurações padrão foi mantida, mas devido ao grande volume de dados e para evitar underfitting foram feitas algumas modificações:

* **num\_train\_epochs:** Ajustado para 3, garantindo mais ciclos de aprendizado.
* **warmup\_steps:** Configurado para 500, visando uma estabilidade maior durante o início do treinamento.
* **learning\_rate:** Definida como 5e-5, para permitir uma adaptação mais gradual e eficaz aos novos dados.
* **per\_device\_train\_batch\_size:** Movido para 16, com a intenção de utilizar mais GPU no treinamento