

THIAGO ADRIANO

POSTECH

IA PARA DEVS

TECH CHALLENGE

FASE 3

---

## TECH CHALLENGE

Tech Challenge é o projeto que engloba os conhecimentos obtidos em todas as disciplinas presentes na fase. Esta é uma atividade que, em princípio, deve ser desenvolvida em grupo. É importante atentar-se ao prazo de entrega, uma vez que essa atividade é obrigatória, valendo 90% da nota de todas as disciplinas da fase.

### O problema

No Tech Challenge desta fase, você precisa executar o fine-tuning de um foundation model (Llama, BERT, MINSTREL etc.), utilizando o dataset "The AmazonTitles-1.3MM". O modelo treinado deverá:

- Receber perguntas com um contexto obtido por meio de uma integração RAG (Retrieve-and-Generate), utilizando documentos relacionados aos produtos da Amazon.
- A partir do prompt formado pela pergunta do usuário e dos dados retornados do RAG, o modelo deverá gerar uma resposta baseada na pergunta do usuário e nos dados provenientes do RAG, incluindo as fontes.

### Fluxo de Trabalho Atualizado:

#### 1. Escolha do Dataset:

Descrição: O The AmazonTitles-1.3MM consiste em consultas textuais reais de usuários e títulos associados de produtos relevantes encontrados na Amazon, medidos por ações implícitas ou explícitas dos usuários.

#### 2. Preparação do Dataset:

Faça o download do dataset [AmazonTitles-1.3MM](https://drive.google.com/file/d/12zH4mL2RX8iSvH0VCNnd3QxO4DzuHWnK/view) - <https://drive.google.com/file/d/12zH4mL2RX8iSvH0VCNnd3QxO4DzuHWnK/view> e prepare os dados para o fine-tuning, garantindo que estejam organizados de maneira adequada para o treinamento do modelo. Limpe e pré-processe os dados conforme necessário para o modelo escolhido.

### **3. Execução do Fine-Tuning:**

Execute o fine-tuning do foundation model selecionado (por exemplo, BERT) utilizando o dataset preparado.

Documente o processo de fine-tuning, incluindo os parâmetros utilizados e qualquer ajuste específico realizado no modelo.

### **4. Configuração da Integração RAG:**

Configure uma integração RAG (Retrieve-and-Generate) para fornecer contexto ao modelo a partir dos documentos relacionados aos produtos da Amazon. Certifique-se de que a integração esteja funcionando corretamente para recuperar e fornecer dados contextuais ao modelo.

### **5. Geração de Respostas:**

Configure o modelo treinado para receber perguntas dos usuários. Quando uma pergunta for recebida, utilize a integração RAG para recuperar informações relevantes do dataset AmazonTitles-1.3MM. Combine a pergunta do usuário e os dados retornados do RAG para formar um prompt completo.

O modelo deverá gerar uma resposta baseada na pergunta do usuário e nos dados provenientes do RAG, incluindo as fontes fornecidas.

O que esperamos para o entregável?

- Documento detalhando o processo de seleção e preparação do dataset.
- Descrição do processo de fine-tuning do modelo, com detalhes dos parâmetros e ajustes utilizados. Código fonte do processo de fine-tuning e da configuração da integração RAG.
- Um vídeo demonstrando o modelo treinado gerando respostas a partir de perguntas do usuário e utilizando o contexto obtido por meio da integração RAG.

Se você ficou com alguma dúvida, não deixe de acessar o Discord para que alguém da equipe possa te ajudar.



POSTECH