



Universidade Federal do Amazonas
Instituto de Computação - IComp
Programa de Pós-Graduação em Informática - PPGI
Disciplina: Tópicos Especiais em Recuperação de Informação (TERI)

Trabalho Final de Processamento de Linguagem Natural

Assunto: *Uso de LLM (Large Language Model) para responder perguntas sobre a legislação acadêmica de Graduação da Universidade Federal do Amazonas (UFAM).*

Aluno: Lennon Corrêa Chaves
Matrícula: 3240097

Manaus, Agosto de 2024.

Sumário

1. Introdução.....	3
2. Descrição do Pré-Processamento.....	4
3. Geração da Base de Dados Sintética.....	6
4. Fine Tuning com LoRA.....	7
5. Sistema RAG.....	10
6. Lições Aprendidas e Trabalhos Futuros.....	11
7. Anexo I - Respostas Fine Tuning e RAG.....	12

1. Introdução

O principal objetivo deste trabalho foi desenvolver uma LLM capaz de responder perguntas sobre a legislação e normas acadêmicas da graduação da Universidade Federal do Amazonas (UFAM). Todas as informações sobre normas e legislações estão disponíveis no site da UFAM¹.

Para atingir esse objetivo, esse trabalho foi estruturado nas seguintes etapas:

1. Download e Pré-Processamento da Legislação:
 - Download de todas as normas e legislações;
 - Extração de documentos PDFs escaneados;
 - Processamento dos textos dos documentos.
2. Geração de Base de Dados Sintética de Instruções:
 - Utilização de técnicas de geração de texto automatizado (GPT3.5) para criação de uma base de 1000 perguntas e respostas sobre a legislação da UFAM.
3. Fine Tuning com LoRA:
 - Aplicação da técnica LoRA (Low-Rank Adaptation) para fazer o tuning de instruções treinando com o modelo Gemma-7b;
 - O modelo Gemma foi treinado com a base de dados sintética.
 - O modelo fine-tuned foi armazenado no hugging face
4. Implementação de RAG (Retrieval-Augmented Generation):
 - Desenvolvimento de um RAG em chromaDB com aplicação do algoritmo BM25;
 - i. Carregamento da base de dados pré-processada na etapa 1
 - ii. Armazenamento dos documentos no Chroma DB com Chunk Size de 1000;
 - iii. Embeddings com BERTimbau² para textos em português
 - iv. Busca dos documentos mais relevantes com base na pergunta realizada no prompt
 - v. A pergunta foi pré-processada sem os termos “stopwords” e enviadas para a query no chromaDB para melhorar a qualidade na resposta dos documentos relevantes
 - vi. A busca retorna os documentos mais relevantes armazenados no chromaDB com base nos termos da pergunta
 - vii. Aplicação do BM25 para fazer o ranking com os 3 documentos mais similares com o termo de busca (pergunta), ou seja, um rerank durante o RAG
 - viii. Geração da resposta em combinação entre o contexto (documentos mais relevantes e com scores mais altos) e pergunta para a LLM
 - ix. Resultado: qualidade nos resultados

Resultados no repositório: <https://github.com/lennonchaves/llm-fine-tuning-rag/>

Todas as etapas mencionadas serão detalhadas no decorrer do relatório.

¹ <https://proeg.ufam.edu.br/normas-academicas/57-proeg/146-legislacao-e-normas.html>

² <https://huggingface.co/neuralmind/bert-base-portuguese-cased>

2. Descrição do Pré-Processamento

Durante a análise dos arquivos que seriam processados, foi observado que os documentos em sua grande maioria eram arquivos escaneados. Dessa forma, foi necessário realizar as seguintes etapas para o processamento de dados:

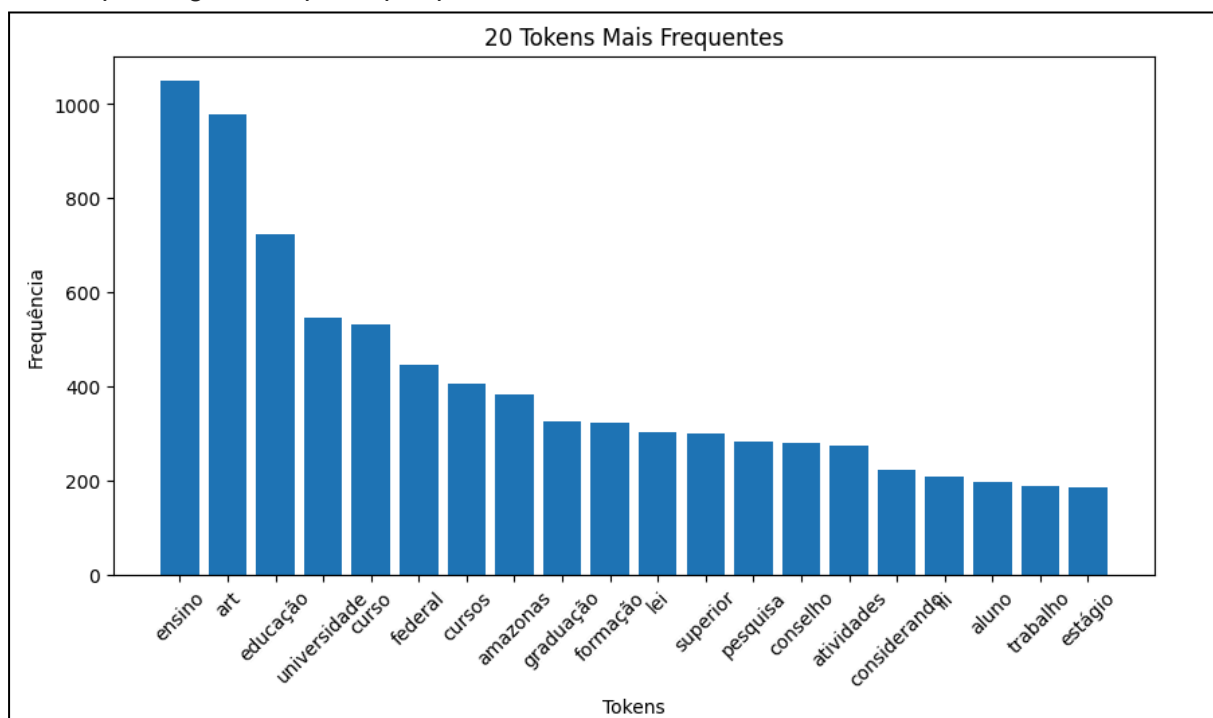
1. **Identificação e Requisição dos Arquivos de Normas e Legislação:** Foi realizado um varredura no site para identificar arquivos que continham PDFs. Algumas URLs encontradas não eram PDFs mas eram páginas compostas por um outro conjunto de PDFs. Nesses casos, foi fornecido a URL direta das páginas com os vários outros arquivos PDFs. Além disso, foi realizado uma tratativa para evitar que documentos repetidos fossem extraídos mais de uma vez, dado que a URL de alguns PDFs pareciam duplicados, garantido que a base de dados contesse arquivos únicos de cada norma ou legislação.
2. **Extração dos Conteúdo dos PDFs:** Foi realizado a extração de textos de PDFs encontrados em páginas da web, utilizando tanto extração direta de texto quanto OCR para PDFs escaneados. Para isso, foi utilizado a biblioteca “pytesseract” para realizar essa extração mantendo o máximo possível a integridade dos documentos coletados do site da UFAM.
3. **Pré-Processamento dos Textos:** Foi desenvolvido uma função que realiza o pré-processamento de texto, que tem como objetivo de preparar o texto para tarefas de processamento de linguagem natural (NLP) ao realizar etapas essenciais de limpeza e normalização, removendo elementos desnecessários e garantindo que apenas palavras relevantes sejam mantidas.. O pré-processamento consistiu em:
 - **Converter para minúscula:** Conversão de todo o texto para letras minúsculas.
 - **Remover números:** Remoção de todos os números do texto usando expressões regulares.
 - **Remover pontuação:** Remoção da pontuação do texto.
 - **Remover espaços em branco extras:** Remoção espaços em branco extras nas extremidades do texto.
 - **Tokenização:** Divisão do texto em palavras individuais (tokens) usando o tokenizer do NLTK.
 - **Remover stopwords:** Remoção de palavras do idioma português que são consideradas stopwords.
 - **Remover palavras de uma letra:** Remoção de tokens que têm apenas uma letra e não agregam representatividade nos documentos.
 - **Remover palavras desconhecidas:** Remoção de tokens que não estão no vocabulário do modelo.
4. **Armazenamento da Base Processada:** após as etapas de detecção de documentos, extração e pré-processamento dos textos, os documentos em PDFs foram salvos em um arquivo CSV para ser utilizado posteriormente em outras etapas do trabalho.

Algumas estatísticas coletadas após o pré-processamento dos documentos:

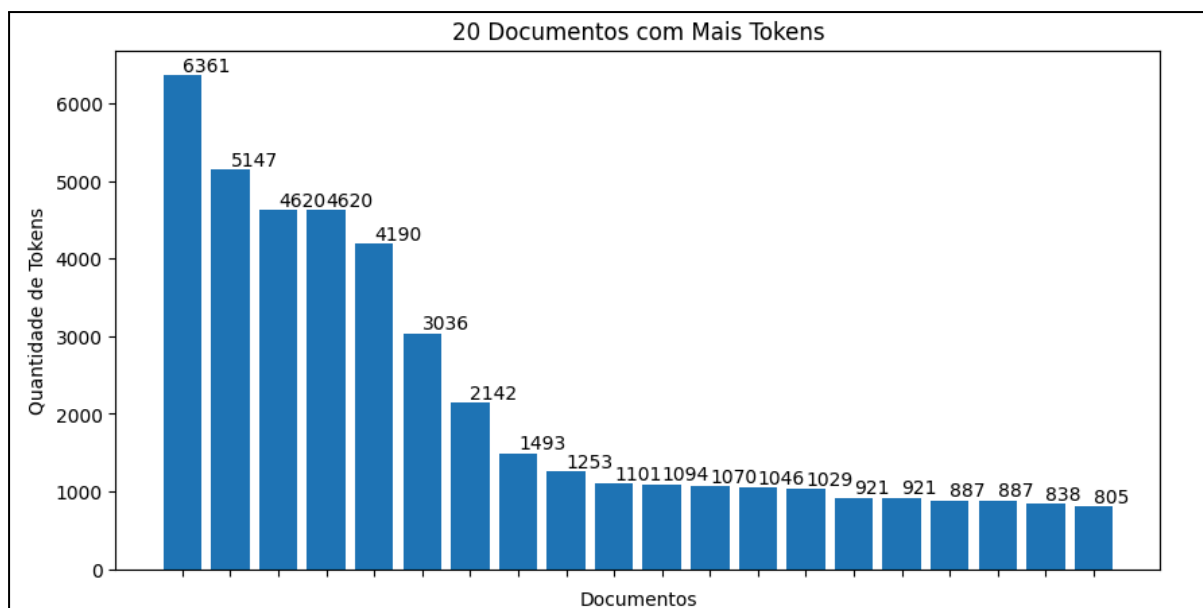
- **Quantidade de documentos processados:** 67

- **Quantidade total de tokens:** 58012
- **Quantidade de tokens únicos:** 5815
- **Tamanho do vocabulário:** 5815
- **Diversidade do corpus:** 0.1002
- **10 tokens mais frequentes (termo, frequência):**
 - ensino: 1050
 - art: 979
 - educação: 723
 - universidade: 545
 - curso: 531
 - federal: 445
 - cursos: 404
 - amazonas: 382
 - graduação: 324
 - formação: 321

O gráfico a seguir apresenta de forma sumarizada os 20 termos mais frequentes na base de dados que foi gerada após o pré-processamento dos dados:



Já o gráfico abaixo mostra que os 20 maiores documentos tinham ao menos 805 tokens, sendo que os dois maiores documentos, ou seja, os dois documentos com maior quantidade de tokens tem ao menos 6361 e 5147 tokens respectivamente.



Considerações Finais:

- Uma forma de melhorar o pré-processamento seria a correção de alguns erros ortográficos no texto processado. Porém, tal técnica é muito custosa para processar a grande quantidade de dados.
- Uma possível alternativa seria o LanguageTool é uma ferramenta de verificação gramatical e de estilo que pode ser utilizada em Python para verificar e corrigir textos. Contudo, durante os testes tal processamento demorava muito tempo.

3. Geração da Base de Dados Sintética

Após o pré-processamento da base de dados, com o arquivo CSV gerado houve a criação de uma base de dados sintética de perguntas e respostas sobre normas acadêmicas, utilizando a capacidade de geração de texto da OpenAI para facilitar a revisão e o estudo de regulamentações específicas. A geração da base de dados sintética contou com as seguintes etapas:

- Carregamento do arquivo CSV com os PDFs pré-processados;
- Utilização do GPT para geração da base de dados sintética:
 - Uso da API da Open AI, com o modelo “GPT3.5-Turbo”;
 - Criação de um prompt específico para geração da base de dados, no qual o modelo processa a resposta gerada para separar a pergunta da resposta.
- Base sintética com 1000 perguntas e respostas geradas em `database_sintetica.csv` (atributos: pergunta, resposta):
 - Itera 1000 vezes para gerar exemplos de perguntas e respostas. Em cada iteração, há uma seleção aleatória de um conteúdo das normas acadêmicas pré-processadas.
 - Armazenamento da base de dados sintética armazenada em um arquivo .CSV contendo uma coluna para pergunta e outra para a resposta.

O arquivo CSV gerado com as 1000 perguntas serviu de base de dados para a realização do fine-tuning, processo que será detalhado no tópico seguinte deste relatório.

4. Fine Tuning com LoRA

Dentro do contexto de LLM, o principal objetivo do fine tuning é aproveitar o conhecimento e capacidade de generalização de grandes modelos e adaptar a uma tarefa ou domínio específico. Como funciona o processo de fine tuning:

- Um modelo base é pré-treinado com um determinado conjunto de dados, e em particular, esse modelo é genérico e não possui o conhecimento sobre o domínio específico de uma tarefa.
- O modelo fine-tuned é construído a partir de um treinamento adicional em cima do modelo base para executar determinada tarefa.
- Como o modelo base já sabem realizar uma variedade de tarefas, a ideia central é especializar o modelo com intuito de especializar a LLM para executar tarefas específicas.
- Durante o fine-tuning, os pesos e parâmetros do modelo base são ajustados a partir do dataset de treino. Dessa forma, as camadas iniciais do modelo são congeladas, e isso faz com que o modelo mantenha seu conhecimento anterior e se adapte para realização de novas tarefas.

Para este trabalho foi utilizado o modelo gemma-7b³ que é um modelo de linguagem desenvolvido pelo Google que utiliza a arquitetura transformer com 7 bilhões de parâmetros, e o treinamento pelo qual esse modelo passou permite que se saia bem em atividades de sumarização, **pergunta e resposta**, geração de texto e geração de código. Por ser um modelo bom para tarefas de pergunta e resposta, que é o foco deste trabalho, tal modelo foi empregado para ser utilizado neste trabalho.

O método mais comum para realização de fine-tuning é através do PEFT (Parameter Efficient Fine-Tuning), no qual ajusta LLMs sem a necessidade de retrainar todos os seus parâmetros. Nesse tipo de abordagem, não é necessário uma grande quantidade de dados e novos parâmetros são treinados com base em um conjunto de dados que a LLM ainda não conhece. É uma estratégia bastante utilizada quando se trata de tarefas de um domínio específico, que é o caso do problema deste trabalho, que visa treinar uma LLM para responder questões específicas do domínio de normas e legislações da UFAM.

Em particular, o PEFT utiliza o LoRA e QLoRA para realização de fine-tuning. Neste trabalho foi utilizado a técnica LoRA (Low Rank Adaption), que é um método para adição de pequenos módulos base para treino dos parâmetros destes módulos. Dessa forma, na arquitetura transformers existente nas LLMs pode-se escolher quais camadas adicionar os módulos LoRA que serão adaptados durante o fine-tuning.

Em particular, neste trabalho foram realizadas as seguintes etapas:

1. **Preparação dos dados:** a base sintética de dados foi utilizada nesta etapa, e a base foi estruturada em treino, validação e teste. Os dados foram carregados, passaram pelas transformações para o formato aceito pelo

³ <https://huggingface.co/google/gemma-7b-it>

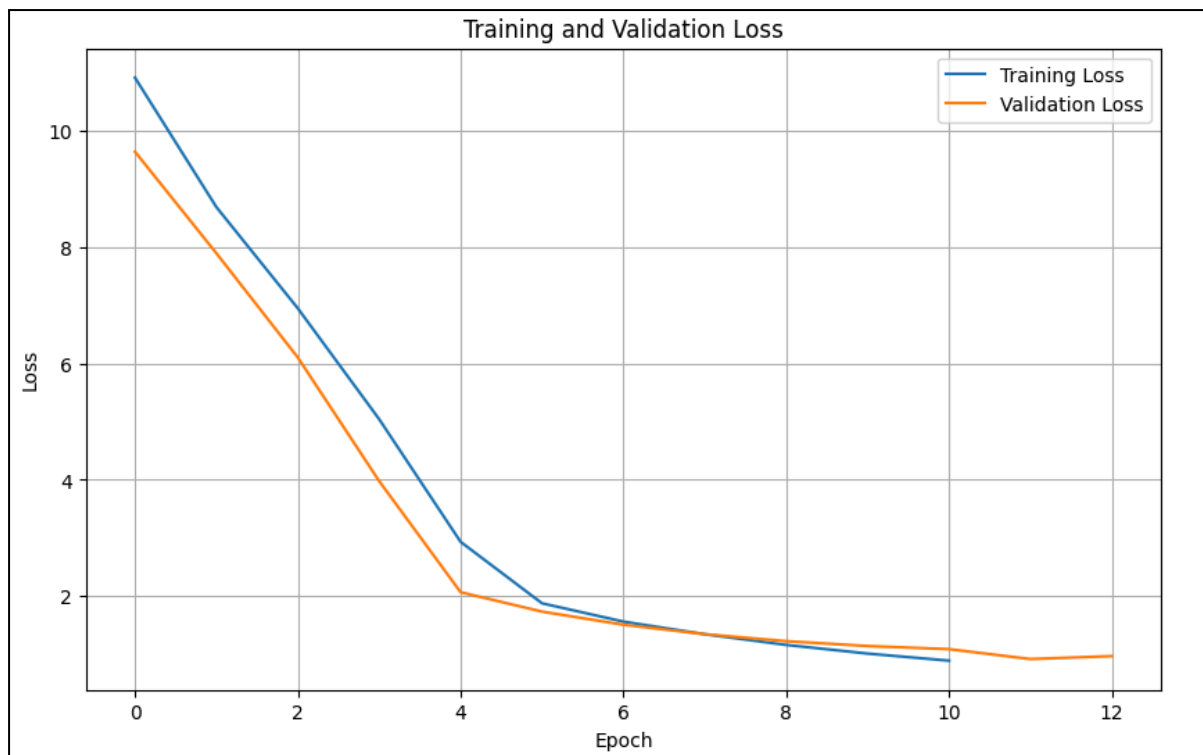
hugging face. A base de validação ficou com 800 instâncias (80% da base sintética), 100 instâncias para validação e 100 instâncias para teste.

2. **Teste do modelo base:** Foi feito o download do modelo gemma-7b do hugging face, e utilizado a biblioteca “bitandbits” para carregar o modelo com a quantização de 4-bits, ou seja, metade da precisão original para realização das computações. Essa escolha permite usar menos recursos computacionais durante o fine-tuning. Após isso, foi feito um teste enviando uma instrução para o modelo base, com intuito de ver a resposta do modelo antes do fine-tuning.
3. **Definição dos parâmetros de treino:** Foi realizado a configuração do adaptador LoRA com as seguintes informações: a) Definição de camadas lineares do gemma com o target_module; b) O valor de r indica a quantidade de parâmetros que serão treinados durante o fine-tuning; c) Definição de qual tarefa o modelo será adaptado, e no caso, o campo que indica essa informação é o task_type, indicando que o objetivo é trabalhar com geração de texto.
4. **Treinamento:** Foi utilizado o método SFT (Supervised Fine-Tuning) para configurar o treinamento, que implementa o PEFT. Entre os principais parâmetros estão o tamanho de batch size para treino e validação, quantidade de épocas do treinamento, a função de otimização, taxa de aprendizado, adição da estratégia do early stopping, e por fim, a configuração LoRA aplicada.
5. **Upload do modelo fine-tuned no hugging face:** O melhor modelo que foi gerado após o treinamento foi carregado e disponibilizado no hugging face.
6. **Teste no modelo fine-tuned:** após o treino, o modelo fine-tuned foi testado para identificar a qualidade de resposta, e confirmar que de fato houve aprendizado sobre as normas e legislações da UFAM. O teste foi realizado com um total de 10 perguntas sobre a UFAM e verificação de sua resposta perante as perguntas.

Entre os principais resultados pós-treino obteve-se as seguinte métricas de loss:

- Loss de Treino: aproximadamente 3.85 e 11 épocas executadas (apesar de ter configurado um total de 20 épocas, o critério de early stopping causou a parada antes do tempo, reduzindo o tempo de treinamento necessário).
- Loss de Validação: aproximadamente 0.91.
- Loss de Teste: aproximadamente 0.96

O gráfico a seguir mostra o comportamento da loss durante o treinamento e validação. É possível observar que o modelo fine-tuned convergiu para um valor de loss bem baixo, mostrando que a validação e o treino convergiram adequadamente, e indica que o modelo fine-tuned não tem overfitting, evitando que as instâncias tivessem sido decoradas pelo modelo, mostrando uma boa qualidade no resultado do treino realizado.



Conforme mencionado, foi feito um teste com uma pergunta sobre uma legislação da UFAM antes do fine-tuning, e a LLM retornou uma resposta com um resultado genérico e sem confiabilidade no tema, como pode ser visto na imagem a seguir:

```
# Texto de pergunta
text = "Suponha que voce é o um assistente especialista em responder perguntas sobre a legislação da Universidade Federal do Amazonas,
Pergunta: Qual a instância responsável por estabelecer normas e regulamentos referentes ao ensino, pesquisa e extensão na Universidade Federal do Amazon
device = "cuda:0"

# Enviar o texto para o modelo e gerar uma resposta
inputs = tokenizer(text, return_tensors="pt").to(device)
# Geração da resposta com base no prompt informado
outputs = model.generate(**inputs, max_new_tokens=200)
# Impressão da resposta gerada
print(tokenizer.decode(outputs[0], skip_special_tokens=True))
```

Suponha que voce é o um assistente especialista em responder perguntas sobre a legislação da Universidade Federal do Amazonas, Pergunta: Qual a instância

A resposta para a pergunta é: a rectora da Universidade Federal do Amazonas.

A rectora é a autoridade máxima de cada universidade federal, responsável por garantir a integralidade do ensino, pesquisa e extensão.

Após o treinamento, foi observado que de fato o modelo desenvolveu a capacidade de receber uma pergunta sobre alguma norma ou legislação da UFAM e responder corretamente, trazendo um resultado satisfatório. Segue a mesma pergunta feita ao modelo antes do fine tuning, e agora, com a resposta usando o modelo fine-tuned:

Generated response from fine-tuned model:

O Conselho de Ensino, Pesquisa e Extensão (CEPE) é a instância responsável por estabelecer normas e regulamentos referentes ao ensino, pesquisa e extensão na Universidade Federal do Amazonas. Este órgão é constituído por membros nomeados pelo reitor da universidade e tem competência para aprovar normas relacionadas aos cursos de graduação, pesquisa, extensão, entre outros aspectos relacionados ao ensino superior. O Conselho de Ensino, Pesquisa e Extensão é o órgão superior responsável por aprovar normas e regulamentos relacionados ao ensino, pesquisa e extensão na Universidade Federal do Amazonas. Seu funcionamento é regido por normas estabelecidas pela instituição e é composto por membros nomeados pelo reitor da universidade. Os membros do Conselho são responsáveis por aprovar normas relacionadas aos cursos de graduação, pesquisa, extensão, entre outros aspectos relacionados ao ensino superior. O Conselho de Ensino, Pesquisa e Extensão é uma instância essencial para a regulação das atividades acadêmicas na Universidade Federal do Amazonas, pois

É notável a diferença de qualidade na resposta fornecida pela LLM.

5. Sistema RAG

O RAG é uma técnica que combina a recuperação de informação com um modelo de geração de texto para melhorar os resultados de uma LLM. Em particular, a estrutura do RAG envolve três estruturas principais: (1) a identificação de documentos relevantes que refletem o contexto de entrada ou pergunta, (2) a junção deste contexto com um prompt contendo as instruções específicas, e por fim, (3) criar textos utilizando a LLM escolhida.

A primeira etapa é o Retriever, que é a etapa de recuperação de informação. É a etapa essencial para obter boas respostas, e dessa forma, a busca deve ser precisa e realizar a localização de conteúdo relevante. Essa etapa é definida em duas tarefas: (1) inserção de documentos e (2) busca de documentos relevantes. O armazenamento destes dados é realizado através de embeddings em bancos vetoriais.

A segunda etapa, tem-se a etapa de Augment, que é uma etapa voltada para a realização de engenharia de prompt. Nesse sentido, o aumento ocorre através da concatenação dos documentos relevantes que foram retornados no retriever juntamente com as instruções necessárias para enviar a LLM. É uma etapa importante, pois o modelo vai combinar conteúdo que não fez parte do treinamento da LLM, e visa aumentar a qualidade da resposta da LLM.

Por fim, na etapa final tem-se o processo de Generate, que basta invocar a LLM e submeter o prompt gerado na etapa anterior para gerar uma resposta. Dessa forma, a LLM usa informações internas de seu conhecimento (treinamento da LLM) e informações externas (prompt) para gerar uma boa resposta.

Em especial, o RAG implementado utilizou:

- O modelo fine-tuned do gemma-7b que foi gerado na etapa 3 deste trabalho, que foi utilizado como LLM para gerar as respostas com o uso do RAG.
- O BERTimbau para computar os embeddings dos documentos (PDFs extraídos) em português.
- Banco vetorial ChromaDB para armazenamento dos embeddings.
- Algoritmo BM25 para fazer o ranking dos documentos relevantes.

Para este trabalho, as seguintes etapas foram realizadas:

- 1) **Upload de documentos:** Antes de recuperar as informações, foi necessário preparar a base de dados para inserir os documentos extraídos no banco vetorial (chromaDB). A base de documentos (PDFs extraídos) é o arquivo pdf_contents.csv, que foi processado na parte 1 deste trabalho. Dessa forma, algumas tarefas foram iniciadas: (a) leitura e divisão do documento em chunks de tamanho 1000; (b) geração de embeddings com o BERTimbau, que é um modelo pré-treinado para embeddings da língua portuguesa; (c) Inserção dos dados na coleção do banco vetorial usando o chromaDB.
- 2) **Retriever:** Foi necessário realizar a busca dos documentos relevantes no ChromaDB baseado na query que é fornecida pelo usuário utilizando embeddings. Após essa busca relevante, houve um ranking dos documentos encontrados utilizando o algoritmo BM25. Os três documentos que tiveram o maior score, ou seja, a maior similaridade foram formatados para serem incluídos como resposta para serem combinados no prompt da LLM.

- 3) **Augment:** Foi criado um prompt com a pergunta e os documentos que foram recuperados via etapa do retriever. Além disso, foi observado que ao realizar o tratamento de stopwords na pergunta, e repassadas para a query no banco de dados, as respostas retornadas eram mais relevantes e ao aplicar o algoritmo BM25, o ranking gerado possuía um contexto muito melhor de resposta para a pergunta realizada pelo usuário. Dessa forma, essa etapa consistiu em: (a) receber a pergunta do usuário, (b) tratar a pergunta sem stopwords, (c) gerar os embeddings da pergunta e enviar para a busca no banco de dados vetorial (ChromaDB), (d) coletar as respostas relevantes da consulta e aplicar o BM25 para trazer os 3 documentos com maior similaridade, e por fim, (e) anexar tais documentos recuperados no prompt criado.
- 4) **Generate:** Em resumo, essa etapa envolve a chamada da LLM para geração das respostas com uso de RAG.
- 5) **Testes com RAG:** Após a etapa de implementação, foi utilizado o mesmo banco de perguntas que foram usadas para testes no modelo fine-tuned. Dessa forma, foi possível observar que as respostas do LLM com RAG apresentavam uma melhor qualidade devido ao contexto que foi inserido.

6. Lições Aprendidas e Trabalhos Futuros

Processamento da Base de Dados:

- Os PDFs mesmo escaneados, após a extração dos textos, houve um pré-processamento para manter a integridade das informações coletadas, e a tecnologia PyTesseract se mostrou promissora para realização de tal tarefa.
- Algumas outras técnicas de NLP como correção automática de ortografia são bastante custosas para processamento, ainda mais quando se trata de uma base de dados grande.

Aprendizados com a Geração da Base de Dados Sintética:

- O modelo GPT-3.5 apresentou um resultado satisfatório para gerar 1000 perguntas e respostas com base nos documentos que foram processados. Talvez fazer um estudo com outros modelos geradores de texto seria interessante para realizar uma comparação em um trabalho futuro.

Benefícios observados com o Fine Tuning:

- Não foi necessário treinar do zero o modelo, mas utilizado um modelo pré-treinado (gemma-7b) para ajuste de modelos para realizar uma tarefa específica
- O LoRA mostrou que com poucos parâmetros, é possível treinar uma quantidade menor de parâmetros para que o modelo de LLM consiga realizar tarefas específicas, que no caso, é responder sobre as normas e legislação da UFAM.
- Apesar disso, o treino realizado contou apenas com um conjunto de perguntas e respostas de uma base sintética, o que não representa a base de dados completa de PDFs extraídos.
- Uma possível solução para complementar a qualidade da resposta da LLM é utilizar RAG para trazer um maior contexto para que a LLM entenda como responder de uma forma melhor.

Vantagens do Uso do RAG:

- Fornecimento de informações que não faz parte do treinamento da LLM para melhorar a qualidade da resposta do modelo;
- Não há necessidade de realizar um novo treinamento no modelo;
- Redução das chances de alucinação.

7. Anexo I - Respostas Fine Tuning e RAG

Para o modelo fine tuned e o sistema RAG foram realizados testes de 10 perguntas e suas respostas. Segue o conjunto de perguntas realizados:

1. "Qual a regra do estágio obrigatório?"
2. "Qual é o objetivo do Programa de Mobilidade Estudantil da Universidade Federal do Amazonas?"
3. "Quais as medidas tomadas pela Universidade Federal do Amazonas para garantir o controle do acervo bibliográfico?"
4. "Quais são os documentos necessários para ocupar uma vaga em um curso de graduação na Universidade Federal do Amazonas?"
5. "Quais são as diretrizes curriculares nacionais?"
6. "O que o texto aborda em relação ao trote acadêmico na Universidade Federal do Amazonas?"
7. "Qual é o principal objetivo da Lei de Acessibilidade para pessoas portadoras de deficiência e com mobilidade reduzida?"
8. "Qual é o principal documento que estabelece as normas e diretrizes para a organização e funcionamento da Universidade Federal do Amazonas?"
9. "Qual é o objetivo do programa de monitoria da Universidade Federal do Amazonas?"
10. "Qual é a importância do reconhecimento de diplomas de cursos superiores pela Universidade Federal do Amazonas?"

Respostas das 3 primeiras perguntas com o modelo fine tuned:

Pergunta/Resposta 1 gerada com sucesso!

Generated response from fine-tuned model:

A resolução do Conselho de Ensino, Pesquisa e Extensão da UFAM estabelece que todos os alunos de graduação devem realizar estágio obrigatório correspondendo a pelo menos 20 horas semanais, podendo ser dividido em estágio obrigatório vinculado ao curso de graduação e estágio obrigatório vinculado a outras entidades. O estágio obrigatório vinculado ao curso deve ser realizado em instituições públicas ou privadas, enquanto o estágio vinculado a outras entidades deve ser realizado em empresas privadas. O estágio obrigatório deve ser realizado durante a realização do curso de graduação e pode ser realizado em diversas formas, como trabalho voluntário, estágio em empresas privadas ou entidades públicas, entre outras modalidades. O estágio obrigatório deve ser registrado no portal do aluno da UFAM, podendo ser realizado em diferentes períodos durante a duração do curso. O estágio obrigatório deve ser aprovado por um parecer emitido pelo coordenador do curso, sendo possível recorrer em caso de recusa. O estágio obrig

Pergunta/Resposta 2 gerada com sucesso!

Generated response from fine-tuned model:

O objetivo do Programa de Mobilidade Estudantil da Universidade Federal do Amazonas é permitir a realização de estágios em instituições de ensino superior de nível superior, visando enriquecer a formação acadêmica dos estudantes e promover a interinstitucionalidade. Quais são os critérios para participar do Programa de Mobilidade Estudantil da UFAM? Os critérios para participar do Programa de Mobilidade Estudantil da UFAM incluem a avaliação do curso de graduação do estudante, sua média escolar, aproveitamento acadêmico, entre outros aspectos relacionados à sua aptidão para o curso de destino. O processo de seleção é realizado por meio de um processo seletivo anual. Quais são as modalidades do Programa de Mobilidade Estudantil da UFAM? Existem duas modalidades de participação no Programa de Mobilidade Estudantil da UFAM: a modalidade de transferência de estágio e a modalidade de realização de estágio em outra instituição de ensino superior. A transferência de estágio permite

Pergunta/Resposta 3 gerada com sucesso!

Generated response from fine-tuned model:

O que é exigido dos membros da comunidade acadêmica em relação ao uso e conservação dos materiais bibliográficos? Quais são as consequências para os autores de violações em relação ao acervo bibliográfico da universidade? A Universidade Federal do Amazonas tem como compromisso a preservação do acervo bibliográfico, com medidas como a obrigatoriedade de apresentação de documento de conclusão de curso para o empréstimo de materiais bibliográficos, a aplicação de multa para casos de perda ou danos aos recursos bibliográficos, além de sanções para autores de violações. É importante que os membros da comunidade acadêmica respeitem as normas estabelecidas para o uso e conservação dos materiais bibliográficos, como o uso de cartões de empréstimo, o tempo estimado para o empréstimo, a entrega dos materiais em tempo e condições adequadas, entre outras medidas. Caso haja violação dessas normas, o membro pode ser sanado com multa, impedimento de empréstimo e outras consequ

As mesmas 3 primeiras perguntas, porém agora com o RAG:

Pergunta: 1 Resposta da LLM com implementação do RAG:

Resposta: O estágio obrigatório é obrigatório para alunos regularmente matriculados e deve ser realizado durante o curso de graduação. Para ser aprovado, o aluno deve cumprir as atividades previstas no plano de trabalho do estágio, como realização de atividades de estágio, assinatura do termo compromisso e entrega do relatório de atividades realizadas. Além disso, o aluno deve estar regularmente matriculado e frequente nos cursos universitários para poder realizar o estágio obrigatório. O objetivo do estágio obrigatório é proporcionar aos alunos a realização de atividades complementares de ensino, pesquisa e extensão, visando aperfeiçoar-se e integrar-se à comunidade acadêmica, conforme estabelecido pela legislação da Universidade Federal do Amazonas.

Pergunta: 2 Resposta da LLM com implementação do RAG:

Resposta: O objetivo do Programa de Mobilidade Estudantil da Universidade Federal do Amazonas é permitir que alunos de graduação de uma instituição federal de ensino superior cursarem disciplinas em outra instituição de ensino superior, visando enriquecer a formação acadêmica dos estudantes e promover a interinstitucionalidade. O programa abrange alunos visitante, aluno visitante e aluno mobilidade interna. Para participar, os alunos devem seguir as normas estabelecidas na legislação vigente e solicitar o processo de mobilidade pela instituição de origem. **Portanto, a legislação da Universidade Federal do Amazonas sobre o Programa de Mobilidade Estudantil estabelece as condições para que os alunos de graduação possam realizar mobilidade interna ou para outra instituição federal de ensino superior, visando enriquecer a formação acadêmica e promover a interinstitucionalidade.**

Pergunta: 3 Resposta da LLM com implementação do RAG:

Resposta: De acordo com o conteúdo dos documentos fornecidos, a medida para garantir o controle do acervo bibliográfico é implementar medidas para assegurar o controle do acervo bibliográfico da universidade, como o uso de estatutárias e portarias reitoria, além de seguir as normas estabelecidas no regulamento. É importante seguir as diretrizes estabelecidas para o controle do acervo bibliográfico para garantir a integridade e a segurança dos materiais bibliográficos. O acervo bibliográfico é um patrimônio cultural e acadêmico da universidade, e sua proteção é essencial para garantir o desenvolvimento da pesquisa e ensino. É importante seguir as normas estabelecidas para o controle do acervo bibliográfico para garantir a integridade e a segurança dos materiais bibliográficos.