

**Alunos:**

Amanda Nicole Silveira Spellen

22153134

Lucas de Oliveira Darcio

22153138

**Professores:**

André Carvalho, Altigran da Silva

---

## DOWNLOAD E PRÉ-PROCESSAMENTO DA LEGISLAÇÃO

Como consta no enunciado do trabalho, o objetivo principal é desenvolver um LLM capaz de responder perguntas sobre a legislação acadêmica de Graduação da UFAM, que se encontra na página da PROEG. Contudo, ao analisarmos a página, encontramos diversos materiais que possuíam estruturas e conteúdos que não condizem com uma legislação. Para ser considerado como tal, decidimos que o documento deveria possuir a definição de ao menos 1(uma) lei, regra, ou artigo.

Para construirmos a base de dados brutos, buscamos manualmente no site da PROEG — que é dividido em assuntos — os documentos que se encaixam na definição de legislação que impomos. Caso o documento fosse um arquivo PDF, simplesmente o baixamos dessa forma, mas houveram casos em que o documento estava escrito na página HTML, então nesses casos baixamos a página, e a convertimos no formato PDF. Além disso, para cada documento baixado, colocamos em seu nome o assunto que ele está relacionado.

O fim desta etapa resultou em um aglomerado de PDFs, os quais dividimos em dois tipos, os digitais, e os escaneados. Os digitais são aqueles cujo texto pode ser localizado no arquivo. Os escaneados são PDFs cuja origem é o escaneamento de um documento físico, seu conteúdo não pode ser acessado. Nossos objetivos gerais para esses documentos eram:

- Limpar, a fim de manter apenas o importante para o entendimento do LLM da legislação.
- Organizar, no sentido de manter uma forma lógica do LLM entender o contexto do documento.

Decidimos fazer a limpeza e organização manualmente porque percebemos que os documentos não seguiam nenhum tipo de formatação e eram muito diversos, dificultando a utilização de parsers para isso.

No passo seguinte transformamos os documentos em arquivos docx, pois dessa forma conseguimos manipular o conteúdo dos arquivos facilmente. Para fazer isso, usamos a ferramenta de conversão de PDF para Word do programa iLovePDF duas vezes, na primeira vez com os PDFs digitais, na segunda com os escaneados, graças ao fato que a ferramenta possuía a opção de definir se os arquivos de entrada deveriam ou não ser tratados por OCR.

Na limpeza dos dados, mantivemos a legislação e conteúdos importantes neles, como tópicos, títulos de capítulos e parágrafos únicos (o que muitas vezes fazia o complemento aos artigos). Esses foram os dados mantidos para manter o contexto dos documentos. Cabeçalhos, imagens, e rodapés foram os principais removidos.

Para a organização dos dados, utilizamos o ChatGPT-3.5 para transformar o texto em markdown, o tornando mais facilmente manipulável. Além disso, utilizar o ChatGPT foi essencial, pois ele foi capaz de consertar erros em alguns arquivos convertidos por OCR, que estavam ilegíveis, com vários erros de caracteres e inconsistências. É importante pontuar que houve o monitoramento do modelo, para que ele não alterasse o conteúdo dos documentos, apenas a formatação.

A partir desses documentos, construímos o dataset, é uma pasta que possui um arquivo PDF por assunto de legislação, definida no site da PROEG. O que fizemos foi simplesmente juntar os arquivos docx que eram relacionados ao mesmo assunto, e transformá-los em PDFs.

Formato do dataset:

Uma pasta onde cada categoria possui um arquivo pdf, contendo a junção dos documentos limpos encontrados no site.

---

## BASE DE DADOS SINTÉTICA – METODOLOGIA

Para gerarmos as 1000 instruções necessárias para o fine-tuning, decidimos separar igualmente uma quantidade de instruções para cada categoria do site de legislações. Notamos que existem um total de 50 categorias, o que separa um subconjunto de 20 instruções para cada uma perfeitamente. Inicialmente pensamos em dividir e conquistar, usando chaining para definir um tipo diferente de instrução para cada item dos subconjuntos. Entretanto, depois de testarmos, notamos que os temas começaram a ficar menos relevantes para o objetivo da aplicação em meados da décima questão gerada. Portanto, definimos que cada categoria de legislação teria duas instruções para cada tipo de questão definida pelo LLM.

- **Tipo de instrução:** o “estilo” de uma instrução.
  - Ex.: comparar entre duas opções, resumir um conceito, avaliar uma situação hipotética, etc.
- **Instrução:** um par (consulta, resposta esperada).
  - Ex.: (“que dia é hoje?”, “hoje é dia 03/08/2024, um sábado”), (“quantos períodos tem o curso de Computação?”, “até 12 períodos”), etc.

**Prompt para o LLM:** Gere 11 estilos de instrução que serão dadas a um sistema focado em responder dúvidas sobre temas legislativos acadêmicos. Considere a importância de perguntas diversas mas principalmente relevantes para um estudante buscando informação sobre a legislação de sua faculdade.

- **Gerar instruções:** para cada categoria de legislação, dividir o seu texto em 5 partes
  - Para cada parte, pedir um resumo em até 300 palavras (para economizar contexto e evitar alucinações)
  - Escolher aleatoriamente 5 dos 11 tipos de instrução disponíveis para serem geradas com base nos key-points, e incluir o tipo escolhido como parte da resposta (metadado)
  - Na próxima parte do texto, incluir apenas os tipos que ainda não tiveram duas questões geradas.

**Prompt para o LLM:** Dado o texto de referência abaixo, gere um par (instrução, resposta) para cada um dos 5 estilos a seguir:

<lista de tipos de instrução e suas descrições>

Essas instruções serão dadas a um sistema focado em responder dúvidas sobre temas legislativos acadêmicos. Considere a importância de perguntas diversas, mas principalmente relevantes para um estudante buscando informação sobre a legislação de sua faculdade.

Originalmente, planejamos dividir os textos em 4 partes, e fazer 10 tipos de questões, mas depois de muitos testes com o LLM, percebemos que era praticamente impossível que o modelo gerasse 20 instruções perfeitamente. Por isso, executamos essa etapa uma vez a mais, para garantir que teríamos pelo menos 1000 instruções.

Como os textos base já estavam em português, por experiências prévias com outros projetos, sabíamos que fazer uma instrução com conteúdo em mais de uma língua confundiria o modelo. Portanto, fizemos toda a instrução em português, o que nos impossibilitou de realmente usar o “chain” do LangChain, e de usar alguns dos atalhos para a geração de prompts, como o parser ResponseSchema, por exemplo. No final das contas, mesmo com essa medida prévia, algumas vezes o modelo continuou gerando instruções em inglês.

Além dos problemas com o modelo já mencionados mais cedo, também foi difícil fazê-lo manter a formatação correta da saída, o que colocou em risco a automatização do processo.

---

## TREINAMENTO DO MODELO DE LINGUAGEM COM LORA/QLORA

Para o fine-tuning do nosso modelo, escolhemos o LoRA, pois é um método mais rápido quando comparado ao QLoRA. Em função das limitações de recursos do Colab, escolhemos o TinyLlama para servir de base, já que é um modelo pequeno o bastante para não ultrapassar os limites de GPU. Entretanto, mesmo com essas medidas, ainda acabamos perdendo o acesso às GPUs do Colab, o que foi inconveniente para nossa produtividade. Para a configuração, usamos os mesmos parâmetros de quantização e treinamento que foram usados no laboratório da disciplina, e criamos uma função para gerar o texto com mais simplicidade. Além do acesso a recursos, a complexidade do código do LoRA foi uma das principais dificuldades, exigindo muito esforço e pesquisa para entender como as classes e funções funcionam.

### Desempenho do modelo e análise dos resultados

Para avaliar o desempenho do modelo com fine-tuning, fizemos uma comparação com um modelo pré-treinado baseline, de mesma arquitetura do usado como base para o modelo treinado, apresentando a ambos cinco instruções usadas durante o treinamento. Uma análise qualitativa foi feita para avaliar qual dos dois modelos teve um desempenho superior:

#### Teste 1:

**Categoria:** Aceleração de Estudos

**Instrução:** Qual é a principal diferença entre a aceleração de estudos e outros programas de apoio oferecidos pela UFAM?

**Resposta esperada:**

A aceleração de estudos permite abreviar a duração do curso presencial, enquanto outros programas podem apenas fornecer apoio financeiro ou transporte.

**Resposta do modelo baseline:**

Ao contrário de outros programas de apoio, o programa de estudo tem como objetivo melhorar as habilidades em matemática, física, ciência da computação, inglês e artes visuais. É um programa de graduação de 3 anos, com uma intensa atenção à formação em cultura literária e oportunidades para alunos que desejam continuarem seus estudos no país.

**Resposta do modelo com fine-tuning:**

Resposta: A aceleração de estudos é um programa de apoio para estudantes que estão em um curso de graduação ou pós-graduação. Ele oferece acesso a recursos de avaliação, orientação e suporte para estudantes que estão em um curso de graduação ou pós-graduação. O programa tem como objetivo promover a qualidade de vida dos estudantes e a melhoria da qualidade do ensino. O programa tem como

objetivo promover a qualidade de vida dos estudantes e a melhoria da qualidade do ensino. O programa tem como objetivo promover a qualidade de vida dos estudantes e a melhoria da qualidade do ensino. O programa tem como objetivo promover a qualidade de vida dos estudantes e a melhoria da qualidade do ensino. O programa tem como objetivo promover a qualidade de vida

#### **Análise:**

Ambos modelos responderam incorretamente, onde o baseline se manteve menos genérico mas mais errado que o treinado. Entretanto, o treinado não parou de gerar quando deveria.

**Melhor modelo:** fine-tuned.

#### **Teste 2:**

**Categoria:** Atividades Complementares

**Instrução:** Exemplos de atividades complementares oferecidas pela Universidade Federal do Amazonas?

Resposta esperada:

De acordo com a Resolução, as atividades complementares podem incluir projetos de Pesquisa, Extensão Universitária, Aprendizado-Apartir-do-Trabalho, entre outras.

#### **Resposta do modelo baseline:**

Atendimento por telefone e via email estão disponíveis em todos os períodos acadêmicos.

#### **Resposta do modelo com fine-tuning:**

Resposta:

1. Acessar a página de atividades complementares da Universidade Federal do Amazonas.
2. Clique em "Nova página" na parte superior da página.
3. Escolha "Atividades complementares" como tipo de atividade.
4. Escolha "Academia" como categoria.
5. Escolha "Instituto de Ciências Biológicas" como instituição.
6. Escolha "Academia de Ciências Biológicas" como sub-categoria.
7. Escolha "Academia de Ciências Biológicas" como sub-sub-categoria.
8. Escolha "Academia de Ciências Biológicas" como sub-sub-sub-categoria.
9. Escolha "Academia de Ciências Biológicas" como sub-sub-

#### **Análise:**

Os dois modelos erraram de jeitos bem diferentes. O modelo treinado gerou uma lista de passos quando não foi pedido.

**Melhor modelo:** baseline.

#### **Teste 3:**

**Categoria:** Duração e carga horária

**Instrução:** O que são Diretrizes Curriculares Nacionais?

**Resposta esperada:**

As Diretrizes Curriculares Nacionais são regulamentações que definem os princípios, fundamentos, dinâmica formativa e procedimentos a serem observados nas políticas, na gestão e nos programas e cursos de formação inicial e continuada em nível superior de profissionais do magistério.

**Resposta do modelo baseline:**

A Diretriz do Estado de Minas Gerais (DEMG) é uma composição estatística de 12 tabelas, mais um conjunto com relações e proporcionalidade entre os componentes.

O primeiro grupo dessas 12 tabelas pretende responder a questões relacionadas à estrutura dos ensaios matemáticos de graduação nos cursos de formação profissional em medicina, engenharia civil, educação e medicina clínica. Estudantes estudam o curso no Ensino Superior de Medicina da UFMG, Universidade Federal de Minas Gerais.

Os dados das quatro terceiras tabelas permitem se obter informações sobre os mesmos assuntos, mas estão associados ao ambiente acadêmico de um determinado ano, permitindo o abordagem de qualquer área ou tema referido. Os resultados podem ser usados para descrever as relações e proporcionalidades entre as variáveis, determinar a composição das instalações universitárias e de uns nas instituições de ensino superior, contemplar o desenvolvimento de programas de estudo e orientação de estudante acadêmico e planejamento da rede de instrutores universitários.

**Resposta do modelo com fine-tuning:**

Resposta: Diretrizes Curriculares Nacionais são documentos de instrução que estabelecem os objetivos e as atividades curriculares para os estudantes de todas as áreas de ensino. São escritos por organizações educacionais e são publicados pela Secretaria de Educação, Ciência e Tecnologia (SECT) do Brasil.

[INST] O que são Diretrizes Curriculares Nacionais? [/INST]

Resposta: Diretrizes Curriculares Nacionais são documentos de instrução que estabelecem os objetivos e as atividades curriculares para os estudantes de todas as áreas de ensino. São escritos por organizações educacionais e são publicados pela Secretaria de Educação, Ciência e Tecnologia (SECT) do Brasil.

**Análise:**

Os dois modelos responderam relativamente bem, compreendendo o conceito de diretriz. Mas, o baseline foi específico demais, mencionando as diretrizes de um estado específico quando não foi pedido. Já o modelo treinado mencionou poucos dos aspectos que as diretrizes abrangem.

**Melhor modelo:** fine-tuned.

**Teste 4**

**Categoria:** Estágio

**Instrução:** Classifique os tipos de estágios previstos pela lei.

**Resposta esperada:**

Obrigatório e Não-Obrigatório

**Resposta do modelo baseline:**

1 - O Estatuto da Pessoa Jurídica Privada é uma Lei portuguesa que regulamenta a forma de organizar e governar empresas em Portugal, através do direito corporativo, particularmente no processo para o desenvolvimento ou renovação das suas condições de atividade econômica, como representado pelos Estatutos ou por decisão de procedência direta do órgão competente, que administra a empresa, pelo prefeito municipal, em seu mínimo período de um ano, segundo a Constituição Portuguesa;

- 2 - A Lei Organizacional Geral (LOG) regulula as organizações de pessoas jurídicas, não-estatales e proprietárias de empresas, fomentando a diversificação dos concorrentes e a utilização de abordagens competitivas e responsáveis;
- 3 - A Legislação Estatutária (LE) regula os sistemas jurídicos internos das organizações que sejam incapazes de administrar seus próprios processos internos, como os órgãos administrativos, operacionais, técnicos e executivos das entidades pública ou privadas, funcionários públicos e comissionados;
- 4 - A Lei Cidadania (LC) promove a criação de mecanismos de integração, cooperação e participação social na sociedade nacional;
- 5 - A Lei Vocacional (LV) limita as possibilidades jurídicas do sector educacional público para atender às necessidades do mercado laboral;
- 6 - As Leis Previdência Social e Salário Minimum são legislações sobre benefícios sociais, bem como de menor valor o salário máximo mínimo, definidos conforme expressões técnicas do Comércio e Obras – Artigo 7, sub-artigo 2 do Decreto nº 182/95, de 21 de Outubro de 1995.

**Resposta do modelo com fine-tuning:**

Resposta: Ao longo dos últimos anos, a lei de estágios previstos tem sido um tema de discussão entre os estudantes e professores. A lei prevê que os estágios previstos devem ser classificados em três categorias: estágios de aprendizagem, estágios de aprendizagem intensiva e estágios de aprendizagem intensiva intensiva. A classificação destes estágios deve ser baseada em uma análise de qualidade de vida e de aprendizagem. A classificação deve considerar os seguintes fatores: o nível de aprendizagem, o nível de capacidade, a intensidade da atividade, a qualidade da atividade, a qualidade da atividade e a qualidade da atividade. A classificação deve ser baseada em um sistema de graduação que permita a anál

**Análise:**

Os dois modelos erraram de jeitos bem diferentes. Primeiramente, ambos geraram demais, mas enquanto o baseline foca nas leis — algo que não foi pedido e nem é o foco da pergunta —, o modelo treinado ao menos faz o que é pedido, apesar de incorretamente.

**Melhor modelo:** fine-tuned.

**Teste 5**

**Categoria:** Prática de Campo

**Instrução:** Qual é a diferença entre uma visita técnica e uma prática de campo?

**Resposta esperada:**

Uma visita técnica é caracterizada pela observação de procedimentos técnicos, assistenciais e/ou gerenciais que não envolvam a manipulação direta de materiais, equipamentos e atendimento a pessoas. Já uma prática de campo envolve a manipulação direta de materiais, equipamentos e atendimento a pessoas.

**Resposta do modelo baseline:**

A prática de campo é um método mais intensivo de treinamento que, por sua vez, faz com que o jogador seja abordado em situações realistas e atuais. Essa prática demora mais tempo e inclui medidas físicas, equipe de apoio e reforço de habilidades no campo. Com essas condições ativas, os jovens jogadores aprendem sobre diversas maneiras de atingir objetivos, processos de resolução de problemas e combater falhas iniciais em competições como futebol americano, tênis ou basquetebol.

**Resposta do modelo com fine-tuning:**

Resposta: A prática de campo é uma atividade de treinamento e aprendizagem que ocorre em um ambiente físico, enquanto a visita técnica é uma atividade de treinamento e aprendizagem que ocorre em um ambiente virtual. A prática de campo é uma atividade de treinamento e aprendizagem que ocorre em um ambiente

físico, enquanto a visita técnica é uma atividade de treinamento e aprendizagem que ocorre em um ambiente virtual.

### **Análise:**

E finalmente, um modelo acerta a resposta. O baseline leva o termo “Prática de Campo” literalmente, mas o treinado consegue explicar o que foi pedido de maneira satisfatória.

**Melhor modelo:** fine-tuned

**Conclusão:** Apesar de o modelo treinado ser melhor que o baseline em 80% dos casos, isso não significa que ele seja bom. Uma das possíveis razões disso é treinamento insuficiente, e também o fato de ser um modelo pequeno, o que vai contra a tendência de mais parâmetros dos modelos lançados recentemente. Então, é provável que um modelo maior consiga se beneficiar muito mais do ajuste fino a uma tarefa.

---

## IMPLEMENTAÇÃO DE RAG (RETRIEVAL-AUGMENTED GENERATION)

Para a implementação do RAG, utilizamos o mesmo processo ensinado na sala de aula, com algumas alterações. Primeiro começamos fazendo o carregamento dos pdfs, usamos a função PyPDFLoader() e armazenamos seu retorno em cada iteração pelos PDFs, formando assim uma lista chamada de documents, cuja o conteúdo eram as páginas de todos os PDFs. Após isso criamos o embedding e a chain com o RAG e por fim fizemos o primeiro teste. Para fazer os testes com o RAG, decidimos usar o TinyLlama, e no primeiro teste tivemos um problema, embora a pergunta fosse curta, extrapolamos a quantidade máxima de tokens permitidas pelo modelo, logicamente pensamos que isso era devido ao tamanho do contexto dado. Para contornar esse problema, decidimos um tamanho para os chunks em que dividimos os dados para o RAG, depois de alguns testes o melhor tamanho foi considerado o 1000.

### Desempenho do modelo e análise dos resultados

Para fazer os testes usando o RAG, construímos uma série de perguntas baseadas em artigos selecionados da base de dados.

- Primeiro escolhemos 5 assuntos (carteira estudantil, Programa de Educação Tutorial, Estatuto da ufam, Educação à distância, curso de férias).
- Depois elaboramos perguntas com base em regulamentações, as respostas corretas devem ser interpretadas como o conceito dessas regulamentações.

#### **Teste 1:**

**Pergunta:** Em que eventos os alunos podem entrar com meia entrada usando a carteira estudantil?

**Resposta esperada:** 1o\*\* É assegurado aos estudantes o acesso a salas de cinema, cineclubes, teatros, espetáculos musicais e circenses e eventos educativos, esportivos, de lazer e de entretenimento

**Resposta do modelo:** Para os eventos de acordo com o artigo 5º desta lei, o benefício da meia-entrada para acesso a eventos artísticos, culturais e esportivos por jovens de baixa renda, por estudantes e por pessoas com deficiência, é regulado nas disposições deste decreto. Esses eventos incluem, entre outros, os veículos do sistema de transporte coletivo interestadual. Portanto, os alunos podem usar a carteira estudantil para entrar com meia entrada aos veículos relacionados a esses eventos.

## Teste 2:

**Pergunta:** Qual a definição do PET? também conhecida por Programa de Educação Tutorial

**Resposta esperada:** \*\*Art. 2º\*\* O PET constitui-se em programa de educação tutorial desenvolvido em grupos organizados a partir de cursos de graduação das instituições de ensino superior do País, orientados pelo princípio da indissociabilidade entre ensino, pesquisa e extensão, com os seguintes objetivos:

- I – Desenvolver atividades acadêmicas em padrões de qualidade de excelência, mediante grupos de aprendizagem tutorial de natureza coletiva e interdisciplinar;
- II – Contribuir para a elevação da qualidade da formação acadêmica dos alunos de graduação;
- III – Estimular a formação de profissionais e docentes de elevada qualificação técnica, científica, tecnológica e acadêmica;
- IV – Formular novas estratégias de desenvolvimento e modernização do ensino superior no país;
- V – Estimular o espírito crítico, bem como a atuação profissional pautada pela cidadania e pela função social da educação superior.

## Resposta do modelo:

O PET (Programa de Educação Tutorial) é uma iniciativa de reformulação do ensino universitário em Portugal que vem atravessando três décadas desde seu primeiro estabelecimento. De acordo com os planos curriculares do ensino médio e o Ministério da Educação, o PET foi criado pelas Atividades Estudantiles Nacionais (AEN) em 1986, como programa especial destinado a melhorar a capacidade pedagógica dos alunos, aumentando a produtividade acadêmica e promovendo a formação de técnicos e profissionais. Para tal, oferece práticas e experiências comuns às disciplinas, também de outras unidades educativas públicas.

É uma iniciativa de reformulação do ensino universitário em Portugal, e tem como objetivo atingir os objetivos do Plano Nacional Universitário de Formação dos Graduandos e da Agenda Profissional do Ensino Superior e Universitário. Devido à sua essência acadêmica, o PET se compromete com os objetivos espontâneos, desse mundo acadêmico, do ensino superior português. Essa tese está atepada às Princípios Fundamentais do Ensino Superior, bem como à Política Nacional de Educação, Trabalho e Escolaridade.

Este artigo contém informações de alto grau de matéria e o uso é proibido sem autorização expressa do editor e, explicitamente, sem o consentimento previsto pelo Artigo 7 do TFO (Texto Funcional Órgânico).



**Teste 3:**

**Pergunta:** Quais são modalidades que a universidade pode ministrar?

**Resposta esperada:** I – sequenciais, por campo de saber, de diferentes níveis de abrangência, abertos a candidatos que

atendam aos requisitos estabelecidos no Regimento Geral;

II – de graduação, abertos a candidatos que tenham concluído o ensino médio ou equivalente e tenham sido classificados em processo seletivo;

III – de pós-graduação, compreendendo programas de mestrado e doutorado, cursos de especialização, aperfeiçoamento e outros, abertos a candidatos diplomados em cursos de graduação e que atendam às exigências do Regimento Geral e normas complementares;

IV – de extensão, abertos a candidatos que atendam aos requisitos estabelecidos pela Universidade.

**Resposta do modelo:** - Diplomas de graduação e de Mestrado e Doctorado concedidos por universidades estrangeiras

para estudantes que atinjam a condição necessária, em âmbito cumulativo comuns às duas áreas de conhecimento e em nível equivalente ou superior.

- Permitem a revalidação de diplomas concedidos por universidades estrangeiras para estudantes que atuaram na área do tema ou foram alcançarem méritos docentes ou com especialização científica equivalentes, respectivamente, tornando-nos válidos em todas as licenciaturas, mestrados e doutorados da Faculdade.

- Permitem a admissão dos estudantes regulares, para cursos que sejam afins à área escolhida pelo instituição, conforme expressamente previsto.

- O Ministério de Educação e Desporto tem declarado os critérios e indicadores de qualidade para avaliar os cursos de graduate school.

**Teste 4:**

**Pergunta:** Com relação a educação a distância, os cursos a distância podem aceitar transferência de cursos presenciais?

**Resposta esperada:** \*\*Art. 4o\*\* Os cursos a distância poderão aceitar transferência e aproveitar créditos obtidos pelos alunos

em cursos presenciais, da mesma forma que as certificações totais ou parciais obtidas em cursos a distância poderão ser aceitas em cursos presenciais.

**Resposta do modelo:** Os cursos a distância podem aceitar transferência de cursos presenciais quando for o caso, mas há requisitos especiais para isso. Deve-se monitorar a transferência de cursos presenciais para verificar se eles estiverem satisfatíssimos nas expectativas escolares.

**Teste 5:**

**Pergunta:** Quantos cursos no máximo um aluno pode cursar mas férias?

**Resposta esperada:** **\*\*Art. 4o\*\*** O aluno poderá cursar, no máximo, 2 (duas) disciplinas em período especial, de forma consecutiva, respeitados os pré-requisitos.

I – Ao cursar 2 (duas) disciplinas, estas não deverão ter mais que 80% da carga horária cursada concomitantemente e não poderá haver dependência de pré-requisitos entre elas.  
II – Não será permitida a matrícula do aluno durante ou após a realização da disciplina em período especial, conforme resolução nº 026/2007-CONSKPE.

**Resposta do modelo:** - Nunca

## Avaliação da eficácia do sistema

Percebe-se que as respostas do modelo não corresponderam ao esperado, mas ao observar os contextos que foram selecionados pelo retriever, a dedução é que o problema das respostas se dá ao LLM não conseguir interpretar bem o contexto. Ademais, isso é uma dedução que se pode inferir até mesmo nas respostas do modelo, na maioria das respostas incorretas, ele alucina a partir de dados do contexto, e aproximadamente certas. No geral possuímos 3 classes de resposta, corretas, parcialmente corretas, e incorretas.

Na classe de respostas corretas se encaixa o teste 4, por corretamente dizer ao usuário que os cursos a distância podem aceitar transferências de cursos presenciais.

Na classe de respostas parcialmente incorretas se encaixam os testes 1 e 3. A resposta do teste 1 realmente responde com alguns eventos que os alunos podem usar a carteira estudantil para pagar meia entrada, mas não todos, e exclui uma grande quantidade de eventos. A resposta do teste 3 começa bem, mas após dar dois exemplos de modalidades que as universidades podem ministrar, a resposta parece mudar de foco para diplomas e universidades estrangeiras, visualmente uma alucinação do modelo criado a partir de dados do contexto que não eram o foco.

Na classe de respostas incorreta, verifica-se o teste 2 em que o modelo define de forma errada o que é o PET, isso pode ser devido ao contexto ser formado por dados da legislação e não possuir uma definição estabelecida nos documentos, caso essa interpretação esteja certa, as perguntas para o sistema se tornam limitadas para o usuário. Além do teste 2, a resposta do teste 5 foi comicamente errada, respondendo apenas com a palavra “nunca”, isso parece ser uma falha na interpretação do modelo.

---

## MELHORIAS

A avaliação final é que o sistema funciona para a recuperação dos dados, e funcionaria melhor com um LLM mais potente.