FUNDAÇÃO GETULIO VARGAS ESCOLA DE MATEMÁTICA APLICADA CURSO DE GRADUAÇÃO EM CIÊNCIA DE DADOS

BERNARDO VARGAS DA MOTTA ALVES

Chatbot Jurídico para respostas de perguntas relacionadas a decisões de recursos que chegaram até o Supremo Tribunal Federal.

Introdução	3
Objetivo Geral	4
Funcionalidades da Plataforma	4
Impacto e Potencial da Plataforma	5
Componentes do Sistema	6
1. Backend	6
2. Frontend	7
3. Estilo e Layout	8
Fluxo do Usuário	9
Contexto Adicional	13
Explicação	13
Testes	13
Diferenças entre as respostas	16
Possíveis Melhorias	18
Conclusão	19

Introdução

A integração entre diferentes áreas do conhecimento tem gerado soluções inovadoras e impactantes em diversas disciplinas. Este projeto foi desenvolvido por um estudante de Ciência de Dados, mas que possui formação prévia em Direito e visa ilustrar como a combinação de expertise jurídica e competências tecnológicas pode ser aplicada na criação de ferramentas inteligentes para resolver problemas complexos.

A motivação para este trabalho surge do crescente volume de informações jurídicas geradas diariamente, como decisões judiciais, leis e regulamentos, que frequentemente tornam o trabalho de advogados, juízes e pesquisadores mais desafiador. Assim como fiz em meu TCC para a graduação em Direito, buscar interseção entre as duas áreas pareceu ser uma forma inovadora e desafiadora de utilizar as duas graduações.

A Ciência de Dados, com suas técnicas avançadas de manipulação, análise e modelagem de dados, oferece uma oportunidade única para enfrentar essas barreiras. Aliada ao conhecimento jurídico, possibilita o desenvolvimento de sistemas capazes de interpretar decisões judiciais, extrair informações relevantes e oferecer respostas contextuais, otimizando o trabalho dos profissionais do Direito.

Este TCC tem como objetivo desenvolver uma plataforma inteligente que permita aos usuários pesquisar decisões judiciais, analisar seus textos e interagir com um chatbot jurídico capaz de responder perguntas específicas e fornecer informações baseadas em leis e artigos mencionados nas decisões. O projeto demonstra como a união entre Direito e Ciência de Dados pode não apenas modernizar a prática jurídica, mas também contribuir para o acesso à justiça e a eficiência do sistema legal.

A iniciativa reflete a busca por uma ferramenta que pode, com seu desenvolvimento presente e futuro, ajudar usuários com diferentes níveis de familiaridade e conhecimento com o Direito, podendo ser utilizada tanto por leigos quanto por profissionais.

Objetivo Geral

O objetivo principal deste projeto é desenvolver uma plataforma inteligente capaz de integrar e explorar dados jurídicos por meio de técnicas de Ciência de Dados e Inteligência Artificial. Mais especificamente, a proposta é criar um sistema que permita a análise automatizada de decisões judiciais, especialmente aquelas que chegaram ao Supremo Tribunal Federal (STF), utilizando um chatbot jurídico avançado, projetado para auxiliar na compreensão e contextualização dessas decisões, respondendo perguntas específicas e fornecendo informações detalhadas sobre leis e artigos mencionados nos textos.

Funcionalidades da Plataforma

O sistema desenvolvido visa oferecer as seguintes funcionalidades principais:

1. Busca e Exibição de Decisões:

Permitir ao usuário pesquisar decisões do STF por diversos critérios, como número do processo, relator ou número identificador, facilitando o acesso às informações de maneira rápida e organizada.

2. Interpretação Automática das Decisões:

O sistema utiliza LLMs para analisar o texto completo das decisões, identificando leis, artigos e outros elementos relevantes.

3. Resposta a Perguntas Jurídicas:

O chatbot jurídico é capaz de compreender questões formuladas pelos usuários, oferecendo respostas baseadas no texto da decisão e em legislações mencionadas, proporcionando contextualização e insights para profissionais e estudantes de Direito.

4. Extração e Mapeamento de Leis:

Automatizar a extração de artigos e leis citados nas decisões, formatando-os de maneira padronizada e enriquecendo as respostas com informações complementares retiradas de bases de dados externas, como arquivos CSV contendo o texto das leis.

5. Suporte ao Estudo e Análise Jurídica:

A plataforma não apenas responde às perguntas do usuário, mas também fornece uma listagem das leis relevantes para o caso, auxiliando estudantes e profissionais na elaboração de pareceres e análises mais profundas, permitindo uma rápida identificação dos documentos legais avaliados na decisão para pesquisa posterior. A

passagem dessas leis como contexto ao bot também permite que perguntas sobre essas leis sejam feitas.

Impacto e Potencial da Plataforma

Ao utilizar decisões judiciais do STF como objeto de análise, o projeto atinge dois objetivos centrais:

1. Contribuir para o Acesso à Justiça:

Tornar mais acessível e compreensível a interpretação de decisões de alto impacto, reduzindo as barreiras de entendimento para estudantes, pesquisadores e cidadãos comuns interessados em temas constitucionais.

2. Inovar na Prática Jurídica:

Provar o potencial da Ciência de Dados para transformar a forma como a informação jurídica é organizada, analisada e utilizada. Ao explorar decisões tão ricas em conteúdo como as do STF, a plataforma serve como um modelo para futuras aplicações em tribunais inferiores e em outras áreas do Direito.

Componentes do Sistema

1. Backend

O arquivo app.py implementa a lógica principal do servidor.

Principais funcionalidades:

o Configuração do Flask e API da OpenAI:

Inicializa a aplicação Flask e define a chave de API para a comunicação com os serviços da OpenAI.

o Rota /

Exibe a página inicial da aplicação, carregando o arquivo index.html.

O Rota /search

Recebe uma requisição POST com os parâmetros de busca (query e search by) e consulta o arquivo Excel (updated data.xlsx).

Busca por decisões judiciais baseadas nos critérios:

- ID da decisão.
- Relator da decisão.
- Número do Processo.

Retorna os resultados como um JSON contendo as decisões encontradas.

Rota /get_txt

Oferece para download o texto completo da decisão selecionada em formato .txt.

O Rota /chat

Processa as mensagens do usuário enviadas via chat.

- Extrai as leis mencionadas no texto da decisão usando a função map_laws (modelo GPT4o faz a busca na decisão).
- Pesquisa os textos das leis e artigos correspondentes em um arquivo CSV (articles.csv) usando a função buscar artigo no csv.
- Gera uma resposta com base na pergunta do usuário e no contexto das leis utilizando a API da OpenAI (modelo GPT4o).

• Funções Auxiliares:

- generate_txt: Salva o texto da decisão em um arquivo .txt.
- map_laws: Extrai leis e artigos mencionados no texto da decisão utilizando um prompt especializado na API da OpenAI.
- buscar_artigo_no_csv: Busca detalhes de leis e artigos em um arquivo
 CSV de referência.
- generate_response: Gera respostas para perguntas jurídicas com base no texto da decisão e no contexto de leis relevantes.

2. Frontend

A interface do usuário é desenvolvida em HTML e utiliza JavaScript para interações dinâmicas. (arquivo index.html)

Componentes:

Seção de Busca:

Permite buscar decisões judiciais com filtros por ID, Relator ou Número do Processo.

Lista de Decisões:

Exibe uma lista de decisões encontradas com base nos filtros de busca.

o Texto da Decisão:

Mostra o texto completo da decisão selecionada.

o Chatbot Jurídico:

Permite que o usuário envie perguntas baseadas na decisão exibida.

As respostas são geradas dinamicamente com base no texto da decisão e no contexto das leis mencionadas.

o Contexto Jurídico:

Exibe as leis e artigos utilizados como contexto para as respostas do chatbot.

3. Estilo e Layout

- o CSS estiliza a página com um design moderno e responsivo.
- o Divisão em três seções principais: busca, chatbot e contexto jurídico.

Fluxo do Usuário

1. Instalação de bibliotecas:

No terminal, instalar as seguintes bibliotecas: pandas, flask, openai, os, re, ast, csv. Para instalar, rodar:

pip install -r requirements.txt

2. Rodar o código:

No terminal, rodar:

python app.py

3. Busca:

O usuário insere um critério de busca e recebe uma lista de decisões correspondentes.

4. Seleção de Decisão:

Ao selecionar uma decisão, o texto completo é exibido na interface.

5. Interação com o Chatbot:

O usuário envia perguntas relacionadas ao texto da decisão.

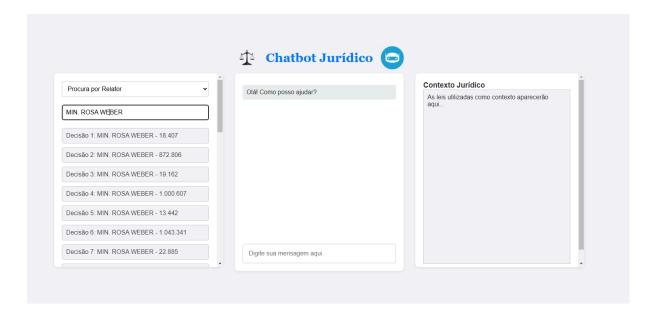
 O chatbot analisa o texto e menciona leis relevantes para fornecer respostas contextuais.

6. Exibição das leis citadas como Contexto Jurídico:

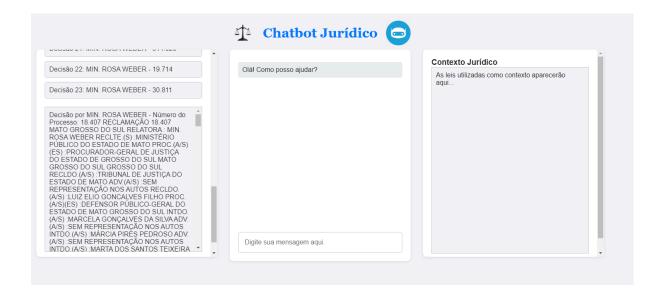
As leis mencionadas no texto da decisão e usadas como base para as respostas aparecem em uma área dedicada para contexto jurídico.

Caso Concreto:

Selecionando busca por relator e escolhendo a Min. Rosa Weber:



Escolhendo a decisão 1 - 18.407, seu texto na íntegra já aparece na parte debaixo



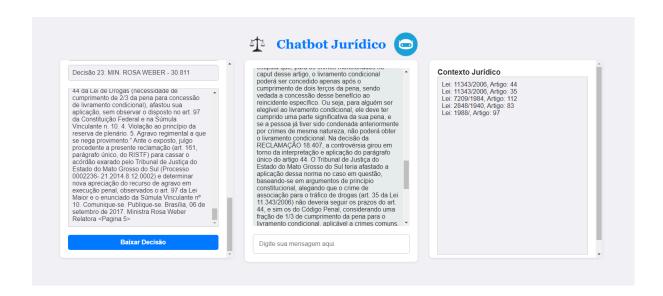
Sendo assim, já é permitido fazer perguntas ao bot:











Contexto Adicional

Explicação

Como dito anteriormente, foi dado ao modelo um contexto jurídico adicional, que permite a

listagem de leis que ocorre na plataforma, mas também permite a passagem de conhecimento

adicional à LLM, visando uma melhoria no prompt desse modelo e assim melhores respostas

às perguntas dos usuários.

Utilizando uma outra LLM, o texto da decisão é passado e esse modelo faz a leitura e

identificação das leis citadas nesse documento. Sendo assim, um código python acessa uma

base de dados local (articles.csv) e passa para o modelo que responde perguntas o texto

integral dessas leis, caso existentes na base local. Dessa forma, objetiva-se enriquecer o

modelo que responde perguntas, oferecendo mais informações jurídicas relevantes através do

seu prompt.

Testes

Como forma de testar a principal adição na aplicação desde o sua concepção (adição das leis

como contexto para o modelo), foram feitas algumas perguntas para duas aplicações

diferentes: uma recebendo esse contexto e outra não recebendo.

Pergunta: Quais os principais pontos da decisão e os artigos e leis relevantes.

Resposta com contexto:

"No caso em questão, o Tribunal de Justiça do Estado de Mato Grosso do Sul afastou a

aplicação do parágrafo único do art. 44 da Lei de Drogas (Lei 11.343/2006) que estabelece

que a concessão do livramento condicional no caso de crimes relacionados ao tráfico de

drogas só pode ser concedido após o cumprimento de dois terços da pena. O Tribunal

justificou essa decisão com base no princípio constitucional da isonomia, alegando que o

crime de associação para o tráfico de drogas (art. 35 da Lei 11.343/06) não é considerado um

crime hediondo e, portanto, a progressão de regime e o livramento condicional deveriam

seguir os lapsos temporais previstos no art. 112 da Lei de Execução Penal (Lei 7.210/1984) e

no art. 83 do Código Penal. O Ministério Público do Estado de Mato Grosso do Sul, por sua

13

vez, ajuizou reclamação alegando que o ato do Tribunal de Justiça havia violado o enunciado da Súmula Vinculante nº 10 do Supremo Tribunal Federal, que estabelece que é necessário o plenário para afastar a incidência de uma norma alegando sua inconstitucionalidade. A Ministra Rosa Weber, do Supremo Tribunal Federal, acatou a reclamação argumentando que a decisão do Tribunal de Justiça do Estado de Mato Grosso do Sul, ao afastar a aplicação de uma norma legal baseando-se em princípios constitucionais, violou a cláusula de reserva de plenário prevista no art. 97 da Constituição Federal, que exige a manifestação do plenário do tribunal para a declaração de inconstitucionalidade de uma norma. A decisão foi embasada nos seguintes dispositivos legais: - Artigos 102, I, I, e 103-A, § 3º, da Constituição Federal; - Parágrafo único do art. 44 da Lei 11.343/2006 (Lei de Drogas); - Art. 97 da Constituição Federal (princípio da reserva de plenário); - Súmula Vinculante nº 10 do Supremo Tribunal Federal; - Art. 161, parágrafo único, do Regimento Interno do Supremo Tribunal Federal."

Resposta sem contexto:

"No caso em questão, o Ministério Público do Estado do Mato Grosso do Sul ajuizou uma reclamação constitucional alegando que o Tribunal de Justiça da mesma unidade federativa desrespeitou o enunciado da Súmula Vinculante nº 10, ao proferir uma decisão que afastou a incidência do parágrafo único do art. 44 da Lei 11.343/06 (Lei de Drogas) em um caso de condenação pelo crime de associação para o tráfico de drogas. O argumento central do Ministério Público está na ideia de que, ao tomar essa decisão, o Tribunal de Justiça estaria desrespeitando o princípio constitucional da reserva de plenário - presente no artigo 97 da Constituição Federal - que determina que a inconstitucionalidade de uma lei ou ato normativo do poder público só pode ser declarada pelo voto da maioria absoluta dos membros de um tribunal. Em sua decisão, a Ministra Rosa Weber observou que, mesmo que o Tribunal de Justiça não tenha declarado expressamente a inconstitucionalidade do disposto no parágrafo único do art. 44 da Lei de Drogas, ao afastar sua aplicação sem a observância da cláusula de reserva de plenário, estaria infringindo essa norma constitucional. Portanto, a Ministra julgou procedente a reclamação do Ministério Público, cassando o acórdão proferido pelo Tribunal de Justiça e determinando uma nova apreciação do recurso de agravo em execução penal,

dessa vez observando corretamente o art. 97 da Constituição Federal e o enunciado da Súmula Vinculante nº 10"

Pergunta: Cite os artigos da Lei 13343 utilizados na decisão e seu texto na íntegra.

Resposta com contexto:

"Na decisão da RECLAMAÇÃO 18.407 do Estado do Mato Grosso do Sul, foram citados os seguintes artigos da Lei 11.343/06 (Lei de Drogas): 1. *Artigo 35: "Art. 35. Associarem-se duas ou mais pessoas para o fim de praticar, reiteradamente ou não, qualquer dos crimes previstos nos arts. 33, caput e § 1°, e 34 desta Lei: Pena - reclusão, de 3 (três) a 10 (dez) anos, e pagamento de 700 (setecentos) a 1.200 (mil e duzentos) dias-multa. Parágrafo único. Nas mesmas penas do caput deste artigo incorre quem se associa para a prática reiterada do crime definido no art. 36 de" 2. **Artigo 44, Parágrafo Único*: "Art. 44. Os crimes previstos nos arts. 33, caput e § 1°, e 34 a 37 desta Lei são inafiançáveis e insuscetíveis de sursis, graça, indulto, anistia e liberdade provisória, vedada a conversão de suas penas em restritivas de direitos. Parágrafo único. Nos crimes previstos no caput deste artigo, dar-se-á o livramento condicional após o cumprimento de dois terços da pena, vedada sua concessão ao reincidente" Esses artigos foram discutidos no contexto da reclamação, que tratou da aplicabilidade do parágrafo único do artigo 44 em relação ao livramento condicional para o crime de associação para o tráfico, previsto no artigo 35. O entendimento do Supremo Tribunal Federal na decisão foi de que a decisão da Corte Estadual ao afastar a aplicação deste parágrafo único sem a observância da cláusula de reserva de plenário violou o enunciado da Súmula Vinculante nº 10."

Resposta sem contexto:

"Na decisão mencionada, os artigos da Lei nº 11.343, de 2006 (comumente referida como Lei de Drogas), que são citados e utilizados, são o artigo 44 e 35. O texto do artigo 44 é: Art. 44. Nos crimes previstos no caput deste artigo, dar-se-á o livramento condicional após o cumprimento de dois terços da pena, vedada sua concessão ao reincidente específico. O artigo mencionado trata sobre as condições para o livramento condicional em casos relacionados a crimes previstos no seu escopo, especificamente no que se refere à quantidade da pena que deve ser cumprida para que tal benefício seja concedido, bem como a

impossibilidade de concessão ao reincidente específico. O artigo 35 da Lei 11.343/2006 estabelece o crime de associação para o tráfico de drogas, que ocorre quando duas ou mais pessoas se associam para a prática reiterada deste tipo de ilícito. No contexto da decisão da Reclamação 18.407 do Mato Grosso do Sul, essa questão foi mencionada em relação ao cumprimento da pena e ao livramento condicional. A decisão discutida refere-se à aplicação ou não do parágrafo único do art. 44 da mesma lei, que estipula requisitos específicos para concessão de livramento condicional para crimes do caput, ou seja, após o cumprimento de dois terços da pena, vedado ao reincidente específico. Neste caso analisado, houve uma controvérsia sobre a necessidade de aplicar tal regra de forma diferenciada em relação ao crime descrito no art. 35. O Supremo Tribunal Federal determinou que o Tribunal de Justiça do Mato Grosso do Sul tivesse que reavaliar essa questão, respeitando a cláusula de reserva de plenário e evitando qualquer indevido afastamento da aplicação de normas legais sem o devido procedimento de declaração de inconstitucionalidade."

Diferenças entre as respostas

Embora as respostas de ambos modelos sejam boas para entendimento do contexto da decisão e da aplicabilidade dos artigos mencionados, podemos perceber uma maior completude nas respostas do modelo com contexto, apresentando os textos na íntegra dos artigos que são passados a ele (apenas citando as leis e artigos daqueles que não possui o texto integral), enquanto o modelo sem contexto responde à pergunta de forma satisfatória, mas apenas cita partes do artigo que vem descritas na decisão em si, sem contexto adicional.

Na primeira pergunta percebemos uma listagem melhor das leis citadas e uma explicação mais completa do caso em questão (ao relacionar possível incidência do artigo 44 da Lei 13.343 ao artigo 35 da mesma Lei, por exemplo). Em relação à segunda questão, que visa justamente ver a diferença entre respostas com e sem contexto, percebe-se claramente que o modelo com contexto responde o texto integral das leis de forma precisa, enquanto o outro modelo (sem contexto) responde listando apenas o parágrafo único art. 44 (ignorando o caput) e parafraseia o artigo 35, não cita seu texto na íntegra.

Como dito anteriormente, ambas os modelos fornecem respostas aceitáveis para entender o caso e a aplicabilidade das leis, mas a adição de um contexto serve não só para uma melhoria na riqueza das respostas, oferecendo detalhes jurídicos muitas vezes relevantes, como

também como uma alternativa de diminuir a amplitude das respostas do bot. Nem sempre nas decisões as leis e artigos vão ter seus textos escritos, às vezes vão só ser citadas e tendo suas aplicações explicadas, então passando o texto desses artigos para o bot, buscamos eliminar algumas variáveis.

Utilizando *prompt engineering* visamos padronizar as respostas do modelo para as perguntas feitas, mas ainda existe uma variância entre respostas feitas em diferentes momentos. Adicionando esse contexto, fornecemos mais ferramentas ao modelo, mais informações, que auxiliam nessa padronização buscada.

Uma possível melhoria (que vai ser citada no próximo capítulo) é a de conseguir essas leis citadas na decisão através de um regex, não utilizando outro modelo de LLM, pois assim vai ser possível conseguir a melhoria observada com o contexto adicional sem outro custo de modelo GPT e ainda é possível conseguir maior consistência na obtenção das leis citadas.

Possíveis Melhorias

- Autenticação: Adicionar autenticação de usuários para maior controle de acesso;
- Validação de Entrada: Melhorar a validação dos inputs do usuário nas rotas /search e /chat;
- Mensagens de Erro: Exibir mensagens mais detalhadas para falhas de busca ou problemas com a API;
- Busca por Tema: Implementar busca por tema jurídico nas possibilidades de escolha da decisão, aumentando a facilidade e a gama de opções na hora de buscar decisão desejada para consulta ao bot;
- Regex de busca de leis: Implementar regex na busca das leis citadas na decisão, reduzindo custos e aumentando a consistência no contexto ao deixar de depender de LLM para a obtenção das leis citadas;
- Enriquecimento da base de leis: Utilização contínua do bot para utilizar mais decisões, ver as leis principais aplicáveis e causar enriquecimento orgânico da base de leis, adicionando apenas uma vez essas principais leis citadas;
- Enriquecimento do contexto jurídico: Muitas vezes em decisões, são citadas outras
 decisões passadas, de diversos tribunais. Uma forma de tornar o modelo ainda mais
 robusto seria conseguir fornecer como contexto essas decisões citadas, pois assim ele
 poderia ter os argumentos relevantes que foram utilizados nessas decisões passadas
 (que muitas vezes são apenas citados nas decisões presentes).

Conclusão

Este trabalho apresentou o desenvolvimento de uma plataforma inteligente que integra Ciência de Dados e Direito, com o objetivo de oferecer uma solução inovadora para a análise de decisões judiciais, especialmente aquelas oriundas do Supremo Tribunal Federal. A proposta alia competências técnicas e jurídicas para criar um sistema capaz de explorar grandes volumes de informações jurídicas de forma eficiente, respondendo a perguntas específicas e contextualizando as decisões com base em leis e artigos mencionados.

A plataforma desenvolvida se mostrou promissora ao abordar desafios como o acesso à informação jurídica, a interpretação de textos legais complexos e a integração de conhecimento multidisciplinar. A funcionalidade do chatbot jurídico e a automação na extração de leis e artigos oferecem ferramentas poderosas para profissionais do Direito, estudantes e cidadãos que buscam compreender melhor as decisões judiciais e seus impactos. Ao mesmo tempo, a possibilidade de personalizar consultas e baixar informações formatadas amplia as aplicações práticas do sistema.

Além de destacar a aplicabilidade da Ciência de Dados no Direito, o projeto reforça a importância da interdisciplinaridade na resolução de problemas reais. Ele não apenas visa auxiliar uma forma moderna de prática jurídica, mas também contribui para o acesso à justiça, ao democratizar o entendimento de temas constitucionais complexos e de alta relevância social.

Como trabalhos futuros, é possível ampliar o escopo da plataforma para incluir decisões de outros tribunais, integrar novas fontes de dados legislativos e refinar os modelos de linguagem utilizados para tornar o chatbot ainda mais eficiente e preciso. Dessa forma, o projeto tem a possibilidade de continuar evoluindo como uma ferramenta inovadora que busca conectar tecnologia e Direito, promovendo eficiência e acessibilidade.