



UNIVERSIDADE FEDERAL DO TOCANTINS
CÂMPUS DE PALMAS
PROGRAMA DE PÓS-GRADUAÇÃO MESTRADO PROFISSIONAL EM MODELAGEM
COMPUTACIONAL

MARCELO AUGUSTO FERRARI FACCIÓNI

**USO DA INTELIGÊNCIA COMPUTACIONAL NO JULGAMENTO DE
IMPROCEDÊNCIA LIMINAR DO PEDIDO**

Palmas /TO
2022

MARCELO AUGUSTO FERRARI FACCIÓNI

**USO DA INTELIGÊNCIA COMPUTACIONAL NO JULGAMENTO DE
IMPROCEDÊNCIA LIMINAR DO PEDIDO**

Dissertação apresentada ao Programa de Pós-Graduação em Modelagem Computacional. Foi avaliada para obtenção do título de Mestre em Modelagem Computacional e aprovada em sua forma final pelo orientador e pela Banca Examinadora.

Orientador: Prof. DSc Marcelo Lisboa Rocha

Palmas /TO
2022

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da Universidade Federal do Tocantins

F138u Faccioni, Marcelo Augusto Ferrari.

 Uso da Inteligência Computacional no Julgamento de Improcedência
 Liminar do Pedido. / Marcelo Augusto Ferrari Faccioni. – Palmas, TO, 2022.
 46 f.

 Dissertação (Mestrado Acadêmico) - Universidade Federal do Tocantins
 – Câmpus Universitário de Palmas - Curso de Pós-Graduação (Mestrado) em
 Modelagem Computacional de Sistemas, 2022.

 Orientador: Marcelo Lisboa Rocha

 1. Uso da Inteligência Computacional. 2. Julgamento de Improcedência
 Liminar do Pedido. 3. Celeridade na Prestação Jurisdicional. 4. Segurança
 Jurídica das Decisões Judiciais. I. Título

CDD 004

TODOS OS DIREITOS RESERVADOS – A reprodução total ou parcial, de qualquer
forma ou por qualquer meio deste documento é autorizado desde que citada a fonte.
A violação dos direitos do autor (Lei nº 9.610/98) é crime estabelecido pelo artigo 184
do Código Penal.

**Elaborado pelo sistema de geração automática de ficha catalográfica da UFT com os
dados fornecidos pelo(a) autor(a).**

FOLHA DE APROVAÇÃO

MARCELO AUGUSTO FERRARI FACCIONI

USO DA INTELIGÊNCIA COMPUTACIONAL NO JULGAMENTO DE IMPROCEDÊNCIA LIMINAR DO PEDIDO

Dissertação apresentada ao Programa de Pós-Graduação em Modelagem Computacional. Foi avaliada para obtenção do título de Mestre em Modelagem Computacional e aprovada em sua forma final pelo orientador e pela Banca Examinadora.

Orientador: Prof. DSc Marcelo Lisboa Rocha

Data de aprovação: ____ / ____ / ____

Banca Examinadora

Prof.^a. Dr.^a. Ângela Issa Haonat – TJTO /ESMAT – Membro Externo

Prof. Dr. David Nadler Prata – PPGMCS/UFT – Membro Interno

Prof. Dr. Gentil Veloso Barbosa – PPGMCS/UFT – Membro Interno

Prof. Dr. Marcelo Lisboa Rocha – PPGMCS/UFT – Membro Interno (Orientador)

Palmas, 2022

Dedico o presente trabalho à minha Família e em especial aos meus filhos João Augusto e Marcelo Augusto que são a minha fonte de inspiração e representam o futuro.

RESUMO

O presente trabalho tem como objetivo demonstrar que o uso da inteligência computacional na prestação jurisdicional pode trazer celeridade e segurança jurídica para as decisões judiciais. Será criada e depois avaliada a possibilidade de utilização de uma ferramenta tecnológica baseada em processamento em linguagem natural e classificação textual para aplicação simultânea dos precedentes judiciais de observância obrigatória e o instituto processual do julgamento de improcedência liminar do pedido. Com isso se buscará comprovar que a decisão judicial poderá ser proferida mais rapidamente, inclusive de forma uniforme para a mesma questão de direito discutida em várias demandas. Através do uso da inteligência computacional realizamos a leitura da petição inicial de processos distribuídos junto aos 1º e 5º Juizados Especiais da Comarca de Palmas-TO nos anos de 2019, 2020, 2021 e 2022, com foco na parte do documento denominada como pedido final ou requerimento e confrontamos com os precedentes vinculativos titulados com os TEMAS 864 e 1177 de Repercussão Geral, editados pelo STF e TEMAS 986 e 1118 de Recursos Repetitivos, gerados pelo STJ. Buscamos conhecer a possível incidência e aplicação do precedente vinculativo no caso específico e dessa forma demonstrar que o julgador poderia saber, logo após o protocolo da ação, que o objeto da pretensão inaugural já tinha manifestação do STF ou STJ e por isso caberia o julgamento de improcedência liminar, e assim comprovar a possibilidade de obter celeridade (eficácia, eficiência, efetividade) e segurança jurídica na prestação jurisdicional com o uso conjugado da tecnologia com os precedentes vinculativos.

Palavras-chaves: *Julgamento liminar de improcedência do pedido. Precedentes vinculativos. Inteligência Computacional. Classificação textual. Processamento de Linguagem Natural.*

ABSTRACT

The present work aims to demonstrate that the use of computational intelligence in judicial provision can bring speed and legal certainty to judicial decisions. The possibility of using a technological tool based on processing in natural language and textual classification will be created and then evaluated for the simultaneous application of judicial precedents of mandatory observance and the procedural institute of the preliminary injunction judgment of the request. With this, it will be sought to prove that the judicial decision can be rendered more quickly, even uniformly for the same question of law discussed in several demands. Through the use of computational intelligence, we carried out the reading of the initial petition of processes distributed with the 1st and 5th Special Courts of the District of Palmas-TO in the years 2019, 2020, 2021 and 2022, focusing on the part of the document denominated as final request or request and confronted with the binding precedents titled with THEMES 864 and 1177 of General Repercussion, edited by the STF and THEMES 986 and 1118 of Repetitive Appeals, generated by the STJ. We seek to know the possible incidence and application of the binding precedent in the specific case and thus demonstrate that the judge could know, right after the filing of the action, that the object of the inaugural claim already had a manifestation of the STF or STJ and therefore it would be up to the judgment of injunction, and thus prove the possibility of obtaining speed (effectiveness, efficiency, effectiveness) and legal certainty in the jurisdictional provision with the combined use of technology with binding precedents.

Key-words: *Preliminary judgment dismissing the application. Binding precedents. Computational Intelligence. Text classification. Natural Language Processing.*

LISTA DE ILUSTRAÇÕES

Figura 1- Figura de distribuição dos processos na Justiça Brasileira no ano de 2020.....	15
Figura 2 - Tramitação dos processos nos juizados especiais brasileiros em 2019.....	20
Figura 3 - Tramitação dos processos nos juizados especiais brasileiros em 2020	21
Figura 4 - Tramitação dos processos nos juizados especiais brasileiros em 2021.....	21
Figura 5 - Tramitação dos processos nos juizados especiais brasileiros em 2022.....	21
Figura 6 - Exemplo de funcionamento do SMOTE em um <i>dataset</i> desbalanceado.....	33
Figura 7 - Fluxograma das fases de criação da ferramenta	35
Figura 8 - Arquitetura da <i>Deep Learning</i> LSTM utilizada.....	42
Figura 1 - Arquitetura da <i>Deep Learning</i> BERT utilizada.....	42
Figura 2 - Matriz de confusão sobre os dados de teste da LSTM.....	43
Figura 31 - Relatório de classificação de teste da LSTM, mostrando as medidas <i>precision</i> , <i>recall</i> e F1-Score.....	44
Figura 12 – Matriz de confusão sobre os dados de teste da BERT.....	45
Figura 13 – Relatório de classificação de teste da BERT, mostrando as medidas <i>precision</i> , <i>recall</i> e F1-Score.....	46
Figura 14 – Resultado apresentado pela ferramenta desenvolvida, após a classificação do texto referente ao processo em questão.....	48

LISTA DE ABREVIATURAS E SIGLAS

PPGCom	Programa de Pós-Graduação em Comunicação e Sociedade
UFT	Universidade Federal do Tocantins
STF	Supremo Tribunal Federal
STJ	Superior Tribunal de Justiça
CPC	Código de Processo Civil
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CAFe	Comunidade Acadêmica Federada
CF	Constituição Federal
TO	Tocantins
LSTM	<i>Long Short-Term Memory</i>
RNR ou RNN	Rede Neural Recorrente
TJTO	Tribunal de Justiça do Estado do Tocantins
CNJ	Conselho Nacional de Justiça
NUGEP	Núcleo de Gerenciamento de Precedentes
COGES	Coordenadoria de Gestão, Projetos e Estratégia
IAC	Incidente de Assunção de Competência
RR	Roraima
SC	Santa Catarina
PLN	Processamento em Linguagem Natural
PJE	Processo Judicial Eletrônico
SAJ	Sistema de Automação da Justiça
PROJUDI	Processo Judicial Digital
BERT	<i>Bidirectional Encoder Representations from Transformers</i>
IPVA	Imposto sobre a Propriedade de Veículos Automotores

SUMÁRIO

1. INTRODUÇÃO	12
1.1 Objetivo Geral	13
1.2 Objetivos Específicos	14
1.3 Justificativa	14
1.4 Capítulos Subsequentes	17
2. EMBASAMENTO TEÓRICO	17
2.1 Foco da pesquisa	20
2.2 Julgamento de Improcedência liminar do pedido	22
2.3 Precedentes Vinculativos	23
2.4 Inteligência Artificial	26
2.5 Mineração de dados	26
2.5.1 Processamento de Linguagem Natural	28
2.6 Classificação do texto	30
2.7 Machine learning	30
2.7.1 Deep learning	31
2.7.2 Long Short-Term Memory (LSTM)	31
2.7.3 Bidirectional Encoder Representations from Transformers (BERT)	32
2.8 Balanceamento de Classes com a Técnica SMOTE	33
3. INTELIGÊNCIA COMPUTACIONAL NO JULGAMENTO DE IMPROCEDÊNCIA LIMINAR DO PEDIDO	34
3.1 Arquivos de petições iniciais	36
3.2 Extração do texto do pedido ou requerimento final	36
3.3 Pré-processamento	37
3.4 Divisão do conjunto de dados para treinamento e teste	39
3.5 Treinamento das Deep Learning LSTM e BERT	39
3.6 Avaliação do treinamento	39
3.7 Teste das Deep Learning LSTM e BERT	40
3.8 Modelo Gerado	40
4. EXPERIMENTO COMPUTACIONAL E ANÁLISE DOS RESULTADOS TRABALHOS FUTUROS	40
4.1 Configuração do Experimento Computacional Utilizado	41
4.2 Análises e Discussões	43

4.2.1 Resultados do Experimento Realizado com a Deep Learning LSTM43
4.2.2 Resultados do Experimento Realizado com a Deep Learning BERT44
4.2.3 Conclusões dos Resultados Obtidos46
5. CONCLUSÕES E TRABALHOS FUTUROS48
REFERÊNCIAS50

1. INTRODUÇÃO

Diante do expressivo crescimento de novos processos que aportam no Judiciário a cada ano e a escassez da mão de obra para solucioná-los, conforme já indicado pelo Portal do Conselho Nacional de Justiça (CNJ), conhecido como Justiça em Números (<https://www.cnj.jus.br/pesquisas-judiciarias/justica-em-numeros/>), resta comprometida a celeridade processual, conhecida pelo legislador constitucional como a razoável duração do processo, sendo salutar a busca de meios alternativos e tecnológicos para se empreender uma marcha satisfatória para solução das lides para concretizar a eficácia, eficiência e efetividade da prestação jurisdicional (DA SILVA, 2016).

No caso específico será feito o uso simultâneo da inteligência computacional com processamento de linguagem natural, criando-se um algoritmo treinado com um precedente obrigatório, previsto no Código de Processo Civil de 2015, para ajudar o magistrado a verificar prematuramente, à primeira vista dos autos, a ocorrência de situação processual que autoriza desde logo julgar liminarmente improcedente o pedido inicial, encerrando a lide, em prazo exíguo.

Para construção e uso da ferramenta tecnológica foi realizado o levantamento de dados junto à Coordenadoria de Gestão, Projetos e Estratégia do Tribunal de Justiça do Estado do Tocantins (COGES/TJTO) do total de novas ações que foram protocoladas junto aos 1º e 5º Juizados Especiais da Comarca de Palmas-TO, nos anos de 2019, 2020 e 2021. Desse total foram separados os processos que possuem como assunto data-base, regime previdenciário, devolução de contribuições previdenciárias pagas além do teto, reajuste de remuneração, soldo, proventos ou pensão e servidores inativos na autuação do processo.

O TEMA 864 publicado pelo Supremo Tribunal Federal (STF), balizou os pedidos judiciais cujo objeto seja o pagamento de data-base ao servidor público.

Já no TEMA 1177 o STF se pronunciou acerca da constitucionalidade do estabelecimento, pela Lei Federal nº 13.954/2019, de nova alíquota para a contribuição previdenciária de policiais e bombeiros militares estaduais inativos e pensionistas.

Foi conhecido, em relação aos processos que tinham como assunto data-base, com aplicação do TEMA 864 do STF, o tempo médio de tramitação dessas ações entre a data do protocolo da petição inicial até seu julgamento de improcedência.

O conhecimento desse lapso temporal não foi possível em relação aos demais processos com aplicação do TEMA 1177 do STF em razão de se encontrarem suspensos por determinação contida no Incidente de Assunção de Competência nº 6, que está tramitando no

Tribunal de Justiça do Estado do Tocantins, assim como os processos relacionados aos TEMAS 986 e 1118 do STJ, uma vez que ainda não transitou em julgado a sua decisão final.

O TEMA 986 do STJ discute a inclusão da Tarifa de Uso do Sistema de Transmissão de Energia Elétrica e da Tarifa de Uso do Sistema de Distribuição de Energia Elétrica na base de cálculo do Imposto de Circulação de Mercadorias e Serviços.

Por fim, o TEMA 1118 do STJ busca definir se o alienante de veículo automotor incorre, solidariamente, na responsabilidade tributária pelo pagamento do Imposto sobre a Propriedade de Veículos Automotores, quando deixa de providenciar a comunicação da venda do bem móvel ao órgão de trânsito competente.

Com o uso da linguagem natural, foi feita a leitura dos termos da petição inicial desses processos, em especial da parte titulada como dos pedidos ou dos pedidos e requerimentos finais, e a comparação de seu conteúdo com a redação do precedente vinculativo contida nos TEMAS 864 e 1177 do STF e TEMAS 986 e 1118 do STJ.

Com a possibilidade de se identificar os processos em que se poderiam aplicar os TEMAS 864 e 1177 do STF e TEMAS 986 e 1118 do STJ, se buscará demonstrar ao final que o julgamento liminar de improcedência do pedido pode ser realizado com o auxílio da inteligência computacional, possibilitando sua incidência e adoção logo após o protocolo da inicial, sem a necessidade das ações seguirem toda a tramitação normal de seu procedimento, o que resultará em celeridade na prestação jurisdicional, cabendo registrar, ainda, que todas as demandas idênticas vão receber uma decisão judicial uniforme, evidenciando a segurança jurídica das relações.

1.1 Objetivo Geral

Demonstrar a possibilidade do uso da inteligência computacional, abordando a mineração de dados com processamento da linguagem natural, para verificar a ocorrência da identidade entre a petição inicial do processo judicial e os precedentes dotados de eficácia vinculante consolidados pelo Supremo Tribunal Federal (Recurso Extraordinário 905357 RR, recurso com repercussão geral – TEMA 864 ou Recurso Extraordinário 1338750 SC, recurso com repercussão geral – TEMA 1177), e Superior Tribunal de Justiça (Recurso Especial 1699851/TO – TEMA repetitivo 986 e Recurso Especial 1881788/SP – TEMA repetitivo 1118) oportunizando o julgamento de improcedência liminar das demandas, o que impacta diretamente na celeridade e segurança jurídica da prestação jurisdicional.

1.2 Objetivos Específicos

- a) Criar ferramenta tecnológica com processamento de linguagem natural para classificação de texto pela *Deep Learning* do tipo *Long Short-Term Memory* (LSTM) e também pela *Deep Learning* do tipo BERT, com a matéria tratada nos TEMAS 864 e 1177 do STF e TEMAS 986 e 1118 do STJ;
- b) Utilizar a ferramenta para identificar o TEMA 864 do STF nas petições iniciais dos processos judiciais distribuídos nos anos 2019, 2020 e 2021, nos 1º e 5º Juizados Especiais da Comarca de Palmas/TO que apontaram como assunto data-base;
- c) Utilizar a ferramenta para identificar o TEMA 1177 do STF nas petições iniciais dos processos judiciais distribuídos nos anos 2020, 2021 e 2022 nos 1º e 5º Juizados Especiais da Comarca de Palmas-TO que apontaram como assuntos a devolução de contribuições previdenciárias pagas além do teto, reajuste de remuneração, soldo, proventos ou pensão e servidores inativos, sendo a parte promovente policial ou bombeiro militar inativo ou pensionista;
- d) Utilizar a ferramenta para identificar o TEMA 986 do STJ nas petições iniciais dos processos judiciais distribuídos nos anos 2020, 2021 e 2022 nos 1º e 5º Juizados Especiais da Comarca de Palmas-TO que apontaram como assuntos repetição de indébito ou exclusão de ICMS;
- e) Utilizar a ferramenta para identificar o TEMA 1118 do STJ nas petições iniciais dos processos judiciais distribuídos nos anos 2020, 2021 e 2022 nos 1º e 5º Juizados Especiais da Comarca de Palmas-TO que apontaram como assuntos IPVA - Imposto Sobre Propriedade de Veículos Automotores, perdas e danos, obrigação de fazer/não fazer e licenciamento de veículos;
- f) Avaliar a possibilidade de redução temporal na tramitação da demanda em que for identificada a aplicação do precedente vinculante com seu julgamento de improcedência liminar.

1.3 Justificativa

O Poder Judiciário recebeu em 2020, 25,8 milhões de novos processos, segundo informação contida no site do Conselho Nacional de Justiça - CNJ¹, conforme a Figura 4 a

¹ <https://www.cnj.jus.br/wp-content/uploads/2021/09/relatorio-justica-em-numeros2021-12.pdf>

seguir:

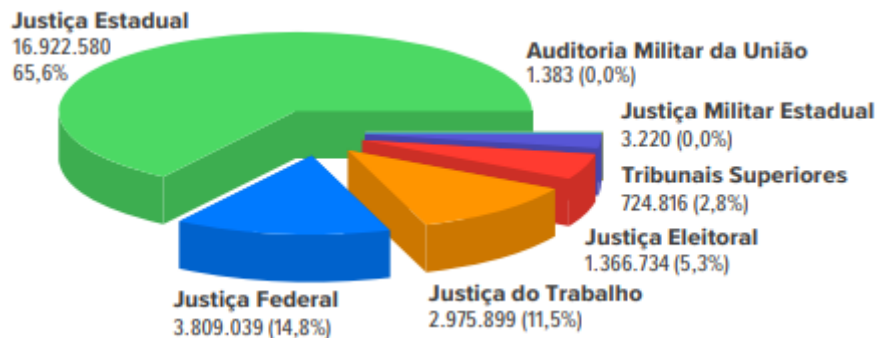


Figura 4. Figura de distribuição dos processos na Justiça Brasileira no ano de 2020.
Fonte: Conselho Nacional de Justiça. CNJ, 2021.

A existência de grande volume de novas demandas judiciais as quais já possuem entendimento consolidado e vinculante geram atraso na prestação jurisdicional até que sejam detectadas pelo magistrado, o que ocorre, normalmente, somente quando a lide chega ao seu fim, na fase de julgamento, quando já se percorreu todo procedimento processual (DE OLIVEIRA BIANCHI, 2020).

Além disso, muitas vezes essas demandas que possuem homogeneidade nas questões jurídicas e fáticas não são identificadas, gerando publicações de decisões judiciais com tratamento diferenciado para casos semelhantes, violando o princípio da segurança e estabilidade das sentenças (BARBOZA, 2017).

O sistema de precedentes judiciais, inaugurado pelo Código de Processo Civil de 2015, em seu artigo 927, tem o escopo de dar maior segurança jurídica ao jurisdicionado, a partir do julgamento igual para os casos iguais, resolvendo ineficiências que enfraquecem o sistema processual, almejando concretizar a efetividade da justiça. O que se busca é incrementar a razoável duração do processo reforçando a segurança jurídica (PASCHOAL; ANDREOTTI, 2018).

Em síntese, o Código de Processo Civil de 2015 aperfeiçoou e efetivou o sistema de precedentes a fim de acelerar a prestação jurisdicional, promover a segurança jurídica aos jurisdicionados e fornecer uma maior estabilidade e coerência ao exercício da atividade do Poder Judiciário (CARVALHO, 2017).

O indeferimento liminar do pedido concretiza os princípios da razoável duração do processo e da segurança jurídica, acelerando a extinção da lide por meio da antecipação da improcedência e, ao mesmo tempo, resguardando a autoridade do precedente judicial

vinculante, mantendo, portanto, a estabilidade, a integridade e a coerência das decisões judiciais (FREIRE JUNIOR, 2021).

Diante do expressivo número de precedentes de observância obrigatória, bem como suas alterações e revisões, é inegável a dificuldade do julgador de seu domínio por completo, surgindo a tecnologia como grande ferramenta auxiliar nessa tarefa de identificação da sua ocorrência na lide (OLIVEIRA, 2020).

Foi direcionada a pesquisa aos 1º e 5º Juizados Especiais da Comarca de Palmas-TO em razão de possuírem demandas com possibilidade da aplicação dos TEMAS 864 e 1177 do STF e TEMAS 986 e 1118 do STJ por terem competência jurisdicional natural para julgamento de ações contra a Fazenda Pública com uso dos assuntos data-base e devolução de contribuições previdenciárias pagas além do teto.

O Sistema Eproc/TJTO² para tramitação eletrônica das lides, adotado pelo Egrégio Tribunal de Justiça do Estado do Tocantins desde o ano de 2010, permite que se utilize dessa inteligência computacional para auxiliar o magistrado nessa análise do cabimento dos TEMAS 864 e 1177 do STF e TEMAS 986 e 1118 do STJ nos processos, sendo a peça fundamental para a existência do presente projeto, em razão de possuir o banco de dados para a pesquisa sistêmica e ser a ferramenta de trabalho tecnológica do julgador.

Nesse contexto perceber-se-á que a aplicação simultânea da técnica de inteligência computacional e do instituto processual da improcedência liminar do pedido poderá gerar o maior ganho de produtividade, com celeridade e segurança na prestação jurisdicional dos últimos anos, de forma simples, inicialmente havendo necessidade apenas de incremento na área tecnológica do Tribunal de Justiça, com baixíssimo custo operacional uma vez que já existe no âmbito do Sodalício Estadual o Núcleo de Gerenciamento de Precedentes (NUGEP) que deverá ficar encarregado de alimentar e manter a ferramenta atualizada.

A técnica de inteligência computacional com o uso de processamento da linguagem natural consistirá basicamente na transformação do documento para um formato de texto, aplicando-se um tratamento para identificar suas características importantes, separando os formatos mais puros para associação às ocorrências no conteúdo do arquivo (NASCIMENTO; CORRÊA; PINHO, 2019).

² <http://www.tjto.jus.br/index.php/noticias/1004-judiciario-do-tocantins-faz-primeiro-despacho-digital-em-processo-eletronico>.

Será possível complementar a capacidade do julgador não só na execução de sua tarefa principal de prestação jurisdicional, mas também na de armazenamento de dados consubstanciados nos precedentes obrigatórios.

Essa prática levará o 1º e o 5º Juizados Especiais da Comarca de Palmas-TO a alcançarem resultados elevados de produtividade com celeridade, além do que se proporcionará às partes o usufruto de uma justiça contemporânea, com decisões uniformes e atuais, confeccionadas com maior rapidez.

1.4 Capítulos Subsequentes

Inicialmente o trabalho se limitará a verificar a ocorrência dos TEMAS 864 e 1177 do STF e TEMAS 986 e 1118 do STJ, todavia, como existem vários julgados de observância obrigatória pelo magistrado, parece ser salutar o enriquecimento da ferramenta tecnológica de linguagem natural com o maior número possível de assuntos vinculativos.

É bom dizer também que não só as demandas com os assuntos data-base, devolução de contribuições previdenciárias pagas além do teto, reajuste de remuneração, soldo, proventos ou pensão e servidores inativos podem ter a incidência de algum precedente vinculativo, mas qualquer uma, ou seja, todos os processos devem passar pela tarefa de comparação feita pela ferramenta tecnológica.

Posteriormente outros precedentes vinculativos também serão inseridos na ferramenta e poderá se caminhar para outra vertente consistente na indicação ao julgador do último arquivo produzido no sistema Eproc/TJTO que se relacione com o conjunto de palavras que está sendo utilizado pelo usuário.

2. EMBASAMENTO TEÓRICO

Para o desenvolvimento deste trabalho foram realizadas, inicialmente, consultas de artigos científicos, trabalhos de conclusão de cursos, dissertações e teses, nos sites da CAPES e do Google Acadêmico, sempre direcionando a busca para responder se o uso da inteligência computacional pode ajudar na aplicação do julgamento de improcedência liminar do pedido e dar mais celeridade e segurança jurídica na prestação jurisdicional.

O acesso ao site da CAPES foi feito através da Comunidade Acadêmica Federada (CAFe), usando e-mail da Universidade Federal do Tocantins (UFT). É bom dizer que a base de dados onde é feita a busca abrange diferentes fontes de informação, sendo as mais

conhecidas e melhor conceituadas no meio científico, contudo o Google Acadêmico possui uma gama muito mais numerosa de documentos.

Cumprе mencionar, ainda, que após a seleção dos trabalhos, na referência bibliográfica de cada um, pode ser encontrado novo material para leitura e estudo uma vez que possuem ligação direta e objetiva com o mesmo tema. Esta forma de pesquisa torna a ação do conhecimento mais concisa e rápida, sem perda de tempo com trabalhos aleatórios ao objetivo da busca.

Utilizaram-se inicialmente os seguintes termos (string) para pesquisa: Julgamento liminar de improcedência pedido - Mineração de dados- Processamento de linguagem natural - *Machine learning* -Prestação jurisdicional inteligência artificial - *Deep learning*.

2.1 Foco da Pesquisa

A Constituição Federal de 1988 foi modificada pela Emenda Constitucional nº 45 de 2004, momento que foi introduzido o princípio da duração razoável do processo, com status de direito fundamental, no inciso LXXVIII do artigo 5º, inclusive como forma de trazer ao Poder Judiciário respeito e confiança. (THEODORO JÚNIOR, 2005).

Assim está disposto na Lei Maior:

- “Art. 5º Todos são iguais perante a lei, sem distinção de qualquer natureza, garantindo-se aos brasileiros e aos estrangeiros residentes no País a inviolabilidade do direito à vida, à liberdade, à igualdade, à segurança e à propriedade, nos termos seguintes: (...)
- LXXVIII - a todos, no âmbito judicial e administrativo, são assegurados a razoável duração do processo e os meios que garantam a celeridade de sua tramitação.”

Foi no Código de Processo Civil de 2015 que foram arrolados os precedentes de observância obrigatória, já com vistas a dar mais celeridade na prestação jurisdicional.

Os precedentes obrigatórios escolhidos para análise e comparação pela inteligência computacional foram os Recursos Extraordinários com Repercussão Geral nº 905357 RR e 1338750 SC, que geraram, respectivamente, o TEMA 864 e TEMA 1177, publicados pelo Supremo Tribunal Federal, e Recursos Especiais Repetitivos nº 1699851/TO e 1881788/SP que deram ensejo ao TEMA 986 e TEMA 1118 originados pelo Superior Tribunal de Justiça, cujas teses fixadas foram a seguinte:

“A existência, ou não, de direito subjetivo a revisão geral da remuneração dos servidores públicos por índice previsto apenas na Lei de Diretrizes Orçamentárias, sem correspondente dotação orçamentária na Lei Orçamentária do respectivo ano”.

“A competência privativa da União para a edição de normas gerais sobre inatividades e pensões das polícias militares e dos corpos de bombeiros militares (artigo 22, XXI, da Constituição, na redação da Emenda Constitucional 103/2019) não exclui a competência legislativa dos Estados para a fixação das alíquotas da contribuição previdenciária incidente sobre os proventos de seus próprios militares inativos e pensionistas, tendo a Lei Federal 13.954/2019, no ponto, incorrido em inconstitucionalidade”.

“Definir se é legal a Inclusão da Tarifa de Uso do Sistema de Transmissão de Energia Elétrica (TUST) e da Tarifa de Uso do Sistema de Distribuição de Energia Elétrica (TUSD) na base de cálculo do ICMS”.

“Definir se o alienante de veículo automotor incorre, solidariamente, na responsabilidade tributária pelo pagamento do Imposto sobre a Propriedade de Veículos Automotores - IPVA, quando deixa de providenciar a comunicação da venda do bem móvel ao órgão de trânsito competente”.

Foi direcionada a pesquisa aos 1º e 5º Juizados Especiais da Comarca de Palmas-TO em razão de possuírem demandas com possibilidade da aplicação dos TEMAS 864 e 1177, ambos do STF e TEMAS 986 e 1118, os dois do STJ, por terem competência exclusiva para julgamento de ações contra a Fazenda Pública.

Registro que o 1º Juizado Especial da Comarca de Palmas-TO passou a ter competência exclusiva fazendária somente a partir do dia 20 de julho de 2020, em decorrência da publicação da Resolução nº 19/2020 do Tribunal de Justiça do Estado do Tocantins, e por essa razão não foi possível utilizar seus dados anteriores a essa data.

As informações relativas ao quantitativo de processos e média de tempo de tramitação nos 1º e 5º Juizados Especiais da Comarca de Palmas foram coletados no sistema Cenarius³ e junto à Coordenadoria de Estratégica, Estatística e Projetos do Tribunal de Justiça do Estado do Tocantins (COGES/TJTO) via Sistema Eletrônico de Informação (SEI) que gerou os processos números: 20.0.000024198-0⁴ e 22.0.000033564-2⁵.

³ https://bi.tjto.jus.br/extensions/Painel_Publico_Cenarius/Painel_Publico_Cenarius.html.

⁴ https://sei.tjto.jus.br/sei/controlador.php?acao=arvore_visualizar&acao_origem=procedimento_visualizar&id_procedimento=10000003023026&infra_sistema=100000100&infra_unidade_atual=110001214&infra_hash=56a4ca7a09cd7155ad9ae94b203933563b22bfaa4c459c880231d0059d679b14.

Já os dados acerca do movimento forense dos 1º e 5º Juizados Especiais da Comarca de Palmas, nos anos de 2019, 2020 e 2021 indicam que foram distribuídos no total, 11.744 novos processos, dos quais 3.211 tinham como assunto data-base. Isso significa que a média anual de processos novos para cada juizado especial foi de 3.914, sendo 1.070 com o assunto data-base. O tempo médio de tramitação desses processos específicos, em 2020, foi de 141 dias e, em 2021, 110 dias.

O quantitativo de processos com os assuntos devolução de contribuições previdenciárias pagas além do teto, reajuste de remuneração, soldo, proventos ou pensão e servidores inativos que foram distribuídos nos 1º e 5º Juizados Especiais da Comarca de Palmas, nos anos de 2020, 2021 e até setembro de 2022, revelam que entraram 1.027 novas ações.

O Conselho Nacional de Justiça, no seu portal da internet, no painel Justiça em Números⁶, indica que no Judiciário Brasileiro, a média de processos distribuídos para cada juizado especial por ano é de 1.300 (mil e trezentos) para cada magistrado, de forma global, incluindo-se todos os assuntos e não só a data-base. Informa, ainda, que o tempo médio de tramitação dos processos nos juizados especiais brasileiros, na fase de conhecimento, tanto em 2019 como em 2020, foi de 9 (nove) meses, conforme as Figuras 2 e 3.



Figura 5. Tramitação dos processos nos juizados especiais brasileiros em 2019.

Fonte: www.cnj.jus.br.

⁵https://sei.tjto.jus.br/sei/controlador.php?acao=arvore_visualizar&acao_origem=procedimento_visualizar&id_procedimento=10000004302969&infra_sistema=100000100&infra_unidade_atual=110001214&infra_hash=7b772e84a4e4d799e2bbb58549b1dcafd29573c4562ee353382b6427a525f92c.

⁶ <https://wwwh.cnj.jus.br/pesquisas-judiciarias/justica-em-numeros/>



Figura 6. Tramitação dos processos nos juizados especiais brasileiros em 2020. Fonte: www.cnj.jus.br.

Já no ano de 2021 o tempo médio de tramitação subiu para 11 (onze) meses, conforme a Figura 7.

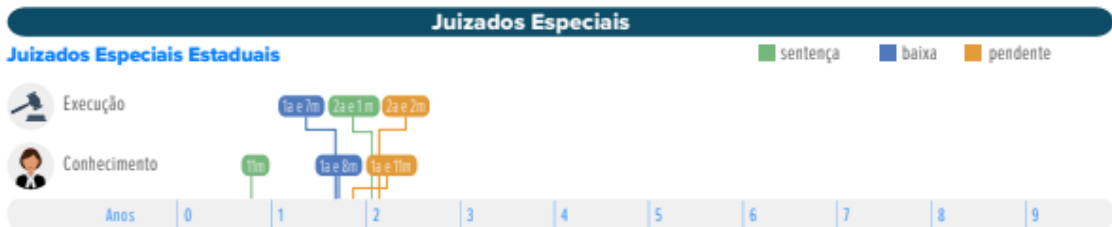


Figura 7. Tramitação dos processos nos juizados especiais brasileiros em 2021. Fonte: www.cnj.jus.br.

Agora no ano de 2022 o tempo médio de tramitação subiu para 12 (doze) meses, conforme a Figura 8.



Figura 8. Tramitação dos processos nos juizados especiais brasileiros em 2022. Fonte: www.cnj.jus.br.

O Tribunal de Justiça do Estado do Tocantins informa no seu painel Estatística de 1º Grau⁷ que o tempo médio de tramitação de um feito nos 1º e 5º Juizados Especiais da Comarca de Palmas, nos anos de 2020 e 2021 foi de 555 dias.

Com esses números já podemos afirmar que temos uma média de processos novos distribuídos no período de 2019, 2020 e 2021 aos 1º e 5º Juizados Especiais da Comarca de Palmas-TO muito acima das demais unidades judiciárias brasileiras em decorrência dos índices arrolados acima. A média de dias de tramitação geral dos processos é bem mais demorada no Judiciário Tocantinense do que nos demais juizados especiais dos outros Tribunais de Justiça Estaduais.

2.2 Julgamento de Improcedência Liminar do Pedido

Como ferramenta disponibilizada ao magistrado para coibir aquelas novas demandas cujo tema já foi objeto de análise judicial e que gerou a edição de Súmulas, Acórdãos Repetitivos, Acórdãos de Afetação, Incidentes de Resolução de Demandas Repetitivas e Julgados acerca da matéria, o Código de Processo Civil de 2015 reafirmou a aplicação do indeferimento liminar do pedido, como consta da redação do artigo 332 e incisos do CPC.

Veja-se que não se formará a relação processual, que se aperfeiçoa com a citação da parte promovida. O réu nem chegará, a saber, que existiu a ação diante da aplicação do indeferimento liminar do pedido. Somente caso seja interposto recurso de apelação dessa decisão judicial de indeferimento liminar é que se dará conhecimento ao réu sobre a existência da demanda.

O indeferimento liminar do pedido concretiza os princípios da razoável duração do processo e da segurança jurídica, acelerando a extinção do processo por meio da antecipação da improcedência e, ao mesmo tempo, resguardando a autoridade do precedente judicial vinculante, mantendo, portanto, a estabilidade, a integridade e a coerência das decisões judiciais (DE CASTRO CATHARINA, 2021).

⁷ https://bi.tjto.jus.br/extensions/Painel_Publico_Cenarius/Painel_Publico_Cenarius.html

2.3 Precedentes Vinculativos

Sem dúvida alguma o maior desafio do Judiciário moderno é dar agilidade na tramitação dos processos, e para isso ele deve adotar uma série de modificações e alternativas para execução de sua tarefa primordial, de julgamento das lides.

O sistema de precedentes judiciais, inaugurado pelo Código de Processo Civil de 2015, em seu artigo 927, tem o escopo de dar maior segurança jurídica ao jurisdicionado, a partir do julgamento igual para os casos iguais, resolvendo ineficiências que enfraquecem o sistema processual, almejando concretizar a efetividade da justiça. O que se busca é incrementar a razoável duração do processo reforçando a segurança jurídica (PASCHOAL; ANDREOTTI, 2018).

Essa necessidade de aprimoramento decorreu de um Poder Judiciário imerso em elevado número de processos e de recursos. Em análise, percebe-se que boa parte das demandas se origina de conteúdos semelhantes, exigindo, por conseguinte, decisões em mesma monta, ou seja, em prestígio ao princípio da isonomia, cabendo aí o papel didático das decisões colegiadas nos juízos ad quem (JÚNIOR, 2017).

Em síntese, o Código de Processo Civil de 2015 aperfeiçoou e efetivou o sistema de precedentes a fim de acelerar a prestação jurisdicional, promover a segurança jurídica aos jurisdicionados e fornecer uma maior estabilidade e coerência ao exercício da atividade do Poder Judiciário.

A lei deixou de ser o único paradigma obrigatório que vincula a decisão do julgador. Os precedentes judiciais também vinculam as decisões judiciais atualmente, já que o novo Código de Processo Civil de 2015 estabelece que não se considera fundamentada qualquer decisão judicial que deixar de seguir precedente ou jurisprudência invocada pela parte, sem mostrar a existência de distinção no caso em julgamento ou a superação do entendimento. O objetivo é de que as decisões judiciais sejam tomadas com coerência e integridade, ou seja, não destoam de outras decisões já prolatadas sobre o mesmo tema e envolvendo as mesmas circunstâncias.

Percebe-se que a aplicação dos precedentes obrigatórios reduz o tempo de duração dos processos, assim como aquelas demandas aventureiras e, conseqüentemente, diminuindo a grande litigiosidade brasileira, de forma a minimizar a sobrecarga do Judiciário, comprometido pela demora na entrega da prestação jurisdicional (DONIZETTI, 2017).

O sistema de precedentes judiciais organizados pelo CPC é uma manifestação do princípio da igualdade, já que garante que pessoas numa mesma situação recebam do

Judiciário o mesmo tratamento no julgamento de seus processos. Além disso, tende a promover a segurança jurídica e diminuir o volume de causas no Judiciário, já que, sabendo dos precedentes em seu caso concreto, as pessoas tendem a evitar demandas temerárias (STEMLER, 2019).

Em grande parte dos assuntos discutidos judicialmente já existe entendimento consolidado e vinculante proferido pelo Supremo Tribunal Federal, Superior Tribunal de Justiça e Tribunais de Justiça, esposados através dos seguintes atos:

I - decisões do Supremo Tribunal Federal em controle concentrado de constitucionalidade;

II - enunciados de súmula vinculante;

III - acórdãos em incidente de assunção de competência ou de resolução de demandas repetitivas e em julgamento de recursos extraordinário e especial repetitivos;

IV - enunciados das súmulas do Supremo Tribunal Federal em matéria constitucional e do Superior Tribunal de Justiça em matéria infraconstitucional;

V - orientação do plenário ou do órgão especial aos quais estiverem vinculados.

A necessidade de ampliação do domínio dos precedentes obrigatórios pelo julgador é inequívoca em face da sua relevância para a prestação jurisdicional, existindo no sítio da internet do STF, STJ e Tribunais Estaduais campo próprio de sua listagem, como por exemplo:

- <http://portal.stf.jus.br/jurisprudencia/sumariosumulas.asp?base=26>
- https://processo.stj.jus.br/repetitivos/temas_repetitivos/
- <https://www.tjto.jus.br/index.php/boletins-informativos>

Todavia, em decorrência da existência de milhares de precedentes obrigatórios, das suas constantes alterações com novas publicações e revisões de seu conteúdo, o domínio humano sobre o sistema vinculativo em sua amplitude fica prejudicado dando lugar ao uso da tecnologia para sanar essa lacuna de conhecimento.

Na atualidade se verifica a ocorrência do seguinte quantitativo de precedentes vinculativos:

- 58 (cinquenta e oito) súmulas vinculantes editadas pelo STF;
- 22 (vinte e duas) decisões do STF em controle concentrado de constitucionalidade;

- 823 (oitocentos e vinte e três) teses com repercussão geral editados pelo STF;
- 1166 (um mil cento e sessenta e seis) incidentes de assunção de competência ou de resolução de demandas repetitivas editados pelo Superior Tribunal de Justiça (STJ);
- 654 (seiscentas e cinquenta e quatro) súmulas editadas pelo STJ;
- 736 (setecentos e trinta e seis) súmulas editadas pelo STF;
- 06 (seis) incidentes de resolução de demandas repetidas em curso no TJTO;
- 04 (quatro) incidentes de assunção de competência em curso no TJTO.

Com isso, deixando o magistrado de perceber a aplicação de um precedente obrigatório logo no início da tramitação do pedido, resultará em atraso na prestação jurisdicional uma vez que vários atos processuais serão praticados e somente ao final da lide, na fase de julgamento, é que se observará que se poderia ter julgado liminarmente improcedente o pleito inaugural.

Além disso, em muitas situações não é identificado o precedente obrigatório nas demandas que possuem homogeneidade nas questões jurídicas e fáticas, gerando publicações de decisões judiciais com tratamento diferenciado para casos semelhantes, gerando insegurança jurídica.

O Conselho Nacional de Justiça publicou em 09 de setembro de 2022 a Recomendação 134/2022 para que os Tribunais e magistrados tribunais que, nos termos do que prevê o CPC, com regularidade, zelem pela uniformização das questões de direito controversas que estejam sob julgamento, utilizando-se, com a devida prioridade, dos instrumentos processuais cabíveis, ficando evidente que esse assunto ainda é novo no cotidiano forense, mas que representa uma nova concepção de jurisdição em que o Poder Judiciário procura não apenas resolver de modo atomizado e repressivamente os conflitos já instaurados, mas se preocupa em fornecer, de modo mais estruturado e geral, respostas às controvérsias atuais, latentes e potenciais, de modo a propiciar a efetiva segurança jurídica.

Apesar da razoável oferta de ferramentas tecnológicas que utilizam inteligência artificial para auxiliar na resolução de questões legais, fornecendo *insights* e subsidiando de várias outras formas os profissionais do ramo, o uso deste tipo de tecnologia ainda é bastante restrito no Brasil, sobretudo por parte dos órgãos julgadores. A adoção de novas ferramentas tecnológicas poderia trazer vários benefícios ao poder judiciário, com benefícios na velocidade e no esforço necessário à análise destes processos (FARACO; TODESCO, 2018).

Então nosso sistema processual deixou claro à parte postulante e ao magistrado que no julgamento da demanda deve se observar se não existe qualquer pronunciamento anterior que vincule a decisão a ser proferida, sendo vedado ao julgador que proceda de forma diferente, salvo se justificado por diferença no caso, superação ou modificação do entendimento.

2.4 Inteligência Artificial

Antes de aprofundarmos alguns conceitos, como processamento de linguagem natural e classificação de textos, é importante conhecer o que é inteligência artificial e onde é aplicada, pois todos os outros conceitos fazem parte dessa grande área de conhecimento.

Conforme Serrano (2012), não existe um consenso entre os autores na definição exata de inteligência artificial, porém todos concordam que para considerar que uma máquina tenha inteligência ela deve ser capaz de executar tarefas para as quais um humano precisaria pensar.

Uma das atividades que caracteriza a inteligência artificial é a compreensão da fala de um ser humano. Para tal, é necessário que a máquina seja capaz de fazer análises morfológicas, sintáticas, semânticas e de contexto (SERRANO, 2012).

Atualmente, análises desse tipo são feitas por meio do processamento de linguagem natural, que é um ramo da inteligência artificial (SERRANO, 2012).

Para o desenvolvimento deste trabalho será necessário aplicar o conceito de inteligência artificial a partir de dados de textos, o que exige a aplicação de algumas técnicas adicionais citadas anteriormente, como processamento de linguagem natural e classificação de textos.

2.5 Mineração de Dados

A mineração de dados nasceu de uma necessidade de se extrair padrões e regras a partir de um conjunto de dados. Em verdade, com o surgimento da tecnologia computacional o objetivo principal é o de guardar/agrupar/armazenar dados (CAMILO, 2009).

Pela mineração de dados tentamos prever um ponto em particular (direcionada) ou achar padrões ou criar grupos (não direcionada) nesse conjunto de dados.

Essa ferramenta avançou muito nos últimos anos, observando-se que seu crescimento obedeceu a fatores diversos, tais como a redução do valor dos equipamentos de tecnologia, aumento do poder de armazenamento e manuseio de dados pelas máquinas, necessidade

comercial de conhecimento do cliente e de adoção de técnicas mais diretas e produtivas para a atividade lucrativa, dentre outras (CAMILO, 2009).

A mineração de dados é perfeitamente utilizável na prestação jurisdicional na medida em que temos no campo jurídico uma gama crescente de dados que vão se multiplicando a cada prática de um novo ato processual.

Antes, porém, é preciso explicar que a mineração de dados se apresenta em forma de várias técnicas (DE AMO, 2004), cabendo então verificar qual delas é a mais adequada para o trabalho que queremos desenvolver.

Técnica é distinta de tarefa, sendo que nesta especificamos o que estamos querendo buscar nos dados, que tipo de regularidades ou categoria de padrão tem interesse em encontrar, ou que poderiam nos surpreender. Naquela há especificação dos métodos que nos garantam como descobrir os padrões que nos interessam (DE AMO, 2004).

Assim temos técnica para regra de associação, padrão sequencial, classificação, clusterização (agrupamentos) e de *outliers*.

Na associação temos um padrão da forma $X \rightarrow Y$, onde X e Y, sendo ambos os conjuntos de valores. Exemplo de uma loja de artigos para pescaria. O seguinte padrão: clientes que compram anzol também compram linha de pescar. Isso representa uma regra de associação que repete um padrão de comportamento dos clientes da loja. Descobrir regras de associação entre produtos comprados por clientes numa mesma compra pode ser útil para melhorar a organização das prateleiras, facilitar (ou dificultar) as compras do usuário ou induzi-lo a comprar mais.

Já na sequência ocorre uma expressão $\langle I_1, \dots, I_n \rangle$, onde cada I representa um conjunto de itens. Assim, por exemplo, a sequência $\langle \{\text{vara de pescar}\}, \{\text{anzol, isca}\} \rangle$ representa o padrão Clientes que compram vara de pescar, tempos depois compram anzol e isca. Descobrir tais padrões sequenciais em dados temporais pode ser útil na oferta de kits, por exemplo.

Pela classificação se procura buscar um conjunto de modelos que descrevem e distinguem classes ou conceitos. Os dados de treinamento estão devidamente classificados e as etiquetas das classes são conhecidas. Com isso se utiliza o modelo para prever a classe de objetos que ainda não foram classificados. Em algumas aplicações, o usuário está mais interessado em prever alguns valores ausentes em seus dados, em vez de descobrir classes de objetos. Isto ocorre, sobretudo quando os valores que faltam são numéricos. Neste caso, a tarefa de mineração é denominada Predição.

Nos agrupamentos (*clusters*) falta conhecimento, não se tem definição acerca das etiquetas de classe. Por exemplo, poderíamos aplicar análise de clusters sobre o banco de dados de um supermercado a fim de identificar grupos homogêneos de clientes e assim saber que clientes aglutinados em determinada região da cidade costumam vir ao supermercado aos domingos, enquanto clientes aglutinados em outros pontos da cidade costumam fazer suas compras às segundas-feiras.

Por fim, *outliers* ocorrem quando um banco de dados contém dados que não apresentam o comportamento geral da maioria. Muitos métodos de mineração descartam estes *outliers* como sendo ruído indesejado. Entretanto, em algumas aplicações, tais como detecção de fraudes, estes eventos raros podem ser mais interessantes do que eventos que ocorrem regularmente. Por exemplo, podemos detectar o uso fraudulento de cartões de crédito ao descobrir que certos clientes efetuaram compras de valor extremamente alto, fora de seu padrão habitual de gastos.

2.5.1 Processamento de Linguagem Natural

Para que seja possível comparar o texto escrito na petição inicial com o enunciado do precedente é necessário que o computador consiga entender o que um humano escreve; logo, é necessário entender o que é Processamento de Linguagem Natural (PLN).

As redes neurais foram originalmente projetadas por psicólogos e neurobiologistas que procuravam desenvolver um conceito de neurônio artificial análogo ao neurônio natural. Intuitivamente, uma rede neural é um conjunto de unidades do tipo input e output.

O processo de aprendizado de um certo conceito pela rede neural corresponde à associação de pesos adequados às diferentes conexões entre os neurônios. Por esta razão, o aprendizado utilizando Redes Neurais também é chamado de Aprendizado Conexionista.

A ferramenta de mineração de dados com o uso de processamento da linguagem natural consiste basicamente na transformação do documento para um formato de texto, aplicando-se um tratamento para identificar suas características importantes, separando os formatos mais puros para associação às ocorrências no conteúdo do arquivo (NASCIMENTO; CORRÊA; PINHO, 2019).

Processamento de linguagem natural é um ramo da área de inteligência artificial que tem como objetivo fazer com que o computador consiga interpretar um texto reproduzido por um humano (GOMES; EVSUKOFF, 2019).

Por uma forma lógica são codificados os possíveis sentidos de cada palavra e identificados os relacionamentos semânticos entre palavras e frases (GONZALEZ; LIMA, 2003).

Neste contexto pode ser utilizada a linguagem natural na aplicação baseada em texto para busca de documentos específicos numa base de dados (DE OLIVEIRA; TONIN; PRIETCH, 2010).

Ao converter uma sequência em um conjunto (tokenização), você perde a noção de ordem.

O Processamento em Linguagem Natural (PLN) pode ser definido como a habilidade de um computador em processar a mesma linguagem que os humanos usam no seu dia-a-dia. Embora também lide com *corpus*, seus interesses são essencialmente aplicados. Para o PLN, não é objetivo descrever e, sim, criar soluções para problemas bastante pontuais, relacionados com o reconhecimento e a reprodução da linguagem humana em alguma escala. As soluções do PLN são sempre pensadas em termos de menor custo e maior benefício (FINATTO, LOPES, CIULLA, 2015).

É certo que referida ferramenta não tomará decisões no processo judicial, mas apenas oferecerá ao julgador uma conferência prévia entre o que é pleiteado no processo e a ocorrência de algum precedente vinculativo, possibilitando que se aplique a improcedência liminar do pedido.

Para o processamento de linguagem natural é necessário passar por algumas etapas:

- Fonologia: estuda os sons que compõem as palavras. É aplicada quando algum sistema precisa escutar algum indivíduo falando;
- Análise morfológica: examina como as palavras se desmembram em componentes e como isso pode afetar sua classe gramatical (substantivos, verbos, pronomes e outras);
- Análise sintática: avalia a estrutura da frase de acordo com a gramática da linguagem utilizada;
- Análise semântica: analisa e associa um significado às palavras e às sentenças em que estão inseridas;
- Análise pragmática: analisa o conteúdo todo e não somente uma parte do texto.

2.6 Classificação de Texto

Classificação de texto é uma das tarefas mais fundamentais do processamento de linguagem natural e tem por objetivo associar um texto a uma ou mais categorias pré-estabelecidas (GOMES; EVSUKOFF, 2019).

Por exemplo: é possível associar o texto de cada e-mail à classe spam ou à classe não spam. Outro exemplo é associar mensagens SMS enviadas a uma empresa sobre um determinado produto e classificá-las como positivas ou negativas (GOMES; EVSUKOFF, 2019).

A ideia é tornar viável a classificação automática dos textos utilizando aprendizado de máquina supervisionado, ou seja, fazer com que a máquina aprenda a classificar com base no texto processado e rotulado (classificado).

Em função dessa ação e do algoritmo escolhido, a máquina será capaz de prever (com determinada margem de erro) a classificação de novos textos, automatizando o processo.

2.7 *Machine Learning*

A partir do entendimento sobre inteligência artificial este trabalho irá aprofundar-se no conceito de *Machine Learning*, de forma a mostrar como a máquina aprende com os dados coletados.

O termo *Machine Learning* começou a ser utilizado em 1959, pelo engenheiro Arthur Samuel que elaborou a seguinte conceituação: “Um campo de estudo que dá aos computadores a habilidade de aprender sem terem sido programados para tal” (BELL, 2015). Contudo, foi somente com o advento da internet e a necessidade de organizar de forma automatizada o grande volume de informações existentes na rede que esse conceito começou a ser mais abordado.

Atualmente, quando falamos de *Machine Learning* estamos falando de uma aplicação computacional capaz de ler e aprender com a experiência sem a necessidade de um programa de computador explícito (BELL, 2015). O processo de aprendizado inicia-se com a observação de dados históricos, com o objetivo de identificar padrões e tomar decisões futuras de acordo com esses padrões. A principal premissa é que o computador seja capaz de aprender sozinho.

Podemos classificar os algoritmos de aprendizado de máquina em supervisionados, não supervisionados, entre outros. Métodos supervisionados são capazes de utilizar o que foi

aprendido no passado para prever os acontecimentos futuros. Usando um conjunto de dados de treinamento conhecido, o algoritmo de aprendizado gera uma função para fazer previsões fornecendo valores de saída. Com a comparação da saída gerada e do acontecimento real são encontrados erros de predição que são utilizados para aperfeiçoar o modelo (BELL, 2015).

Os métodos não supervisionados são usados quando as informações para treinamento não são classificadas e nem rotuladas. O sistema não prevê a saída correta, mas analisa as informações e pode extrair conjuntos de dados para descrever estruturas escondidas de dados não rotulados (BELL, 2015).

2.7.1 *Deep Learning*

A *Deep Learning* é um sub-campo da *Machine Learning*, cujos algoritmos utilizam várias camadas de processamento não lineares, objetivando aprender diversos níveis de abstração (DENG, 2014). Visa o aprendizado da máquina diante da possibilidade de se detectar padrões durante o seu processo com essas diferentes camadas.

O uso da *Deep Learning* teve um expressivo aumento em decorrência do avanço tecnológico, em especial, pela evolução das máquinas e seus componentes, uma vez que trabalha com grande volume de dados.

Todavia a rede neural tradicional não tem persistência, ou seja, não guarda os dados de aprendizado da ação anterior para usar na posterior. Para resolver esse problema temos que utilizar uma rede neural recorrente.

Uma rede neural recorrente pode ser imaginada como múltiplas cópias da mesma rede, cada uma passando uma mensagem a um sucessor, as informações persistem durante o aprendizado de eventos futuros.

2.7.2 *Long Short-Term Memory (LSTM)*

As redes neurais recorrentes (RNR ou RNN — *Recurrent Neural Networks*) são uma família de *Deep Learning* voltadas para o processamento de dados sequenciais. Portanto, constituem-se num recurso bastante apropriado em situações onde há a necessidade de classificar textos. São exemplos: Rede Neural Recorrente do tipo LSTM (*Long Short-Term Memory*), Rede Neural Convolutiva (CNN), Rede Neural Recorrente do tipo *Gated Recurrent Unit* (GRU), Rede Neural Recorrente do tipo Bi-LSTM (*Bidirectional Long Short-Term Memory*) (CHARU, 2018).

Uma situação muito comum onde estas arquiteturas costumam ser utilizadas é quando surge a necessidade de prever a próxima palavra a fazer parte de uma sentença, no momento em que você está escrevendo um texto.

Uma variante das RNNs convencionais é a *Long Short-Term Memory* (LSTM), que resolve um problema conhecido chamado *vanishing gradient*, o qual ocorre quando os pesos computados nas partes iniciais da sequência perdem influência ao longo das iterações, sendo sensíveis às novas entradas. Com isto, a variação de informações contextuais que pode ser captada pelas RNNs comuns costuma ser bastante limitada. Tal limitação leva estas arquiteturas a ter uma performance bastante sofrível em sequências mais longas.

A LSTM surgiu em 1997 para resolver este problema, sendo uma RNN com subestruturas que ajudam a gerenciar a memória da rede neural recorrente. Estas subestruturas são denominadas células de memória, e são fundamentais na solução do problema que a LSTM se propõe a resolver. Assim, sendo a análise de sentimentos uma tarefa de classificação de textos, é natural a escolha das RNNs nestes casos já que a informação temporal e sequencial é extremamente importante para um possível aprendizado de maior assertividade (RONDEL e HILKNER, 2019).

A LSTM apareceu como uma alternativa à resolução dos problemas do gradiente explosivo (que tende para o infinito), e do gradiente desaparecido (que tende a zero) durante o processo de treinamento de uma rede neural.

Então podemos dizer que a *Long Short-Term Memory* (LSTM) são algoritmos que possuem uma memória de longo prazo, composta por uma rede neural recorrente, todavia ao invés de multiplicar a entrada processada, procede a sua soma. Pode associar padrões de textos mesmo com muitas repetições de palavras-chaves.

2.7.3 Bidirectional Encoder Representations from Transformers (BERT)

O algoritmo de aprendizado profundo para processamento de linguagem natural denominado de *Bidirectional Encoder Representations from Transformers* (BERT), de código aberto, desenvolvido por pesquisadores do Google IA Language em 2019 e apresentado em DEVLIN et al. (2019), também foi utilizado para treinamento da ferramenta tecnológica, em busca de um melhor desempenho do que o apresentado pela rede LSTM.

O BERT treina os modelos de linguagem natural observando o conjunto completo de palavras, para que seja capaz de compreender uma língua em uma consulta ou frase conhecida como treinamento bidirecional, enquanto os modelos de PLN tradicionais agem na ordem da

sequência de palavras (da direita para a esquerda ou da esquerda para a direita). O BERT facilita os modelos de linguagem a discernir o contexto das palavras com base nas palavras circundantes, em vez de palavras que o seguem ou precedem (GONZÁLEZ-CARVAJAL e GARRIDO-MERCHÁN, 2020).

2.8 Balanceamento de Classes com a Técnica SMOTE

Na prática podem ocorrer *datasets* que possuam categorias/classes com muito mais dados (mais frequentes) que outras, os quais podem prejudicar a etapa de aprendizado da técnica de inteligência computacional de classificação utilizada e consequentemente os resultados obtidos provocando superajustamento - *overfitting* (HE e GARCIA, 2009). Por exemplo, um *dataset* de imagens de raio-X de tórax, em geral tem-se 90% ou mais dos dados que não apresentam e 10% ou menos que apresentam a COVID-19 (CHAMSEDDINE et al., 2002), mostrando que em uma série de problemas práticos ocorrem o desbalanceamento (não manter a proporção) entre as categorias/classes.

O problema de desbalanceamento de classes também afeta o caso em estudo neste trabalho, com a ocorrência de novos TEMAS/classes (jurisprudências) para os quais ainda se têm poucas exordiais (exemplos) no *dataset*.

A técnica mais utilizada na literatura até o momento para balanceamento de classes é a SMOTE (*Synthetic Minority Oversampling Technique* ou Técnica de Superamostrar a Minoria Sinteticamente) criada por CHAWLA et al. (2002).

A técnica SMOTE consiste da criação automática de exemplos sintéticos das classes minoritárias (aumentando as suas quantidades) de modo a que o *dataset* se torne balanceado, conforme apresentado na Figura 9.

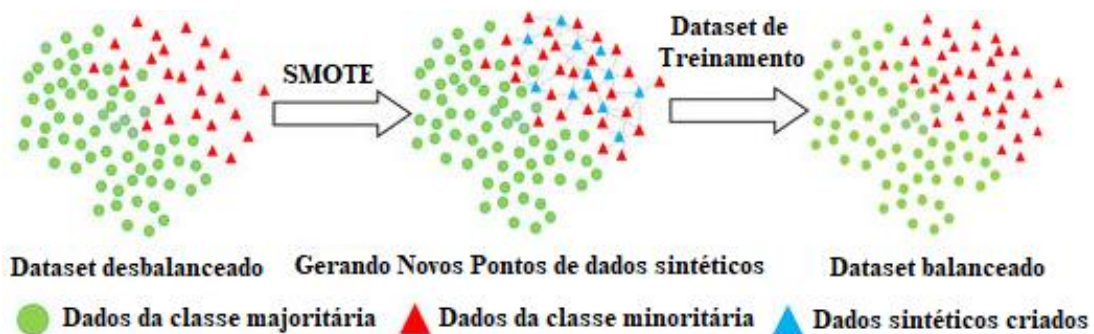


Figura 9. Exemplo de funcionamento do SMOTE em um *dataset* desbalanceado.

Na Figura 9, verifica-se visualmente a aplicabilidade da técnica de SMOTE. Tem-se inicialmente (esquerda) duas classes, onde a classe verde é a majoritária e a vermelha é a minoritária. Posteriormente (centro), tem-se a criação via técnica SMOTE de dados sintéticos em azul. Finalmente (direita), tem-se os dados sintéticos inseridos na classe vermelha, tornando o *dataset* balanceado.

3. INTELIGÊNCIA COMPUTACIONAL NO JULGAMENTO DE IMPROCEDÊNCIA LIMINAR DO PEDIDO

Valendo-se da aprendizagem profunda de máquina, o presente trabalho viabiliza a automação de análises textuais de processos jurídicos, que no caso são as petições iniciais, para identificação dos TEMAS 864 e 1177, ambos do STF e TEMAS 986 e 1118, do STJ.

É importante destacar que eventual tomada de decisão no processo judicial partirá de ação humana e não da ferramenta tecnológica a qual apenas irá auxiliar o julgador em identificar a aplicabilidade de algum precedente vinculativo à situação processual posta em análise na lide.

Neste trabalho, serão focados nos modelos de aprendizagem profunda (*deep learning*) conhecidos como LSTM (*Long Short-Term Memory*) e BERT (*Bidirectional Encoder Representations from Transformers*).

Vale ressaltar que as redes LSTM e BERT não apresentam restrição de uma saída para cada entrada, não sendo um modelo estático de dados, já que produzem classificações dependentes do contexto a que estão expostas.

Para adoção da classificação de múltiplos casos, após a extração dos textos das petições iniciais, foi realizada a transformação dos mesmos, fazendo uso de uma *Toolkit* para processamento de linguagem natural, processando a linguagem simbólica e estatística para a escrita. Dessa forma, foram utilizados os modelos de LSTM e BERT para análise dos textos e dados, a qual é uma ampla rede de caracteres que se sobrepõem utilizada para classificar, agrupar e reconhecer o processamento de linguagem natural.

A última fase é a de decisão para efetivamente concluir se um documento é classificado ou não como sendo compatível com os TEMAS 864 e 1177, do STF e TEMAS 986 e 1118, do STJ.

Em suma, partiremos de um volume de dados e informações sobre os TEMAS 864 e 1177, do STF e TEMAS 986 e 1118, do STJ para a extração de modelos, resultados e

tendências referentes aos dados averiguados.

A pesquisa será desenvolvida observando os seguintes métodos para atingir os objetivos propostos:

- quanto à abordagem será quantitativa, uma vez que levar-se-á em conta o resultado numérico e percentual obtido por meio de procedimento controlado, observando-se a delimitação do problema exposto acima;
- quanto à natureza será aplicada, pois vai gerar conhecimento para aplicação dirigida à resolução do problema específico da aplicabilidade os TEMAS 864 e 1177, do STF e TEMAS 986 e 1118, do STJ nas demandas protocoladas nos 1º e 5º Juizados Especiais da Comarca de Palmas-TO nos anos de 2019, 2020, 2021 e 2022;
- quanto aos objetivos será descritiva, uma vez que se poderá observar, após o uso da ferramenta tecnológica, que será possível alcançar a celeridade, eficácia e segurança jurídica da prestação jurisdicional;
- quanto aos procedimentos será de caso, já que dirigida exclusivamente às demandas protocoladas nos 1º e 5º Juizados Especiais da Comarca de Palmas-TO nos anos de 2019, 2020, 2021 e 2022.

Podemos representar as fases de desenvolvimento da ferramenta tecnológica pelo fluxograma da Figura 10.



Figura 10. Fluxograma das fases de criação da ferramenta. Fonte: Faccioni, M. (2022).

Nas seções 3.1 a 3.8 serão explicadas cada uma das etapas/fases apresentadas na Figura 7.

3.1 Arquivos de Petições Iniciais

Inicialmente, faz-se necessária a elaboração de um *dataset* com 719 (setecentos e dezenove) linhas e duas colunas. Onde cada linha representa uma petição inicial e as duas colunas simbolizam respectivamente o texto da petição inicial e se ela classifica positivamente ou negativamente com os TEMAS 864 e 1177, do STF e TEMAS 986 e 1118, do STJ. Foram utilizadas 300 (trezentas) petições iniciais relacionadas com o TEMA 864 do STF, 30 (trinta) petições iniciais relacionadas com o TEMA 1177 do STF, 67 (sessenta e sete) do TEMA 986 do STJ, 22 (vinte e duas) do TEMA 1118 do STJ e outras 300 (trezentas) com assuntos e temas diferentes, sendo que todas foram extraídas do sistema Eproc/TJTO, de demandas que já tramitaram ou ainda estão em curso nos respectivos 1º e 5º Juizados Especiais da Comarca de Palmas.

As 300 (trezentas) petições iniciais relacionadas com o TEMA 864 do STF tinham data-base como assunto principal, informado na autuação. Já as 30 (trinta) petições iniciais relacionadas com o TEMA 1177 do STF tinham como discussão o pedido de reembolso do desconto previdenciário do policial ou bombeiro inativo como assunto principal, informado na autuação. Nas 67 (sessenta e sete) exordiais relacionadas ao TEMA 986 do STJ o assunto principal era repetição de indébito ou exclusão de ICMS. Nas 22 (vinte e duas) iniciais do TEMA 1118, o assunto principal era IPVA – Imposto Sobre a Propriedade de Veículo Automotor. As outras 300 (trezentas) exordiais possuíam assuntos variados, diversos dos assuntos principais dos TEMAS 864 e 1177, do STF e TEMAS 986 e 1118, do STJ.

Não obstante, para colocar essa ferramenta tecnológica em produção no E-Proc/TJTO, com o uso de um hardware mais poderoso, esse limite do *dataset* do presente trabalho é facilmente transponível.

Deve ser ressaltado que a referida ferramenta pode ser utilizada em conjunto ou integrada a qualquer sistema de processo eletrônico, como por exemplo no PJe (Processo Judicial Eletrônico), SAJ (Sistema de Automação Judicial), PROJUDI (Processo Judicial Digital), dentre outros.

3.2 Extração do Texto do Pedido ou Requerimento Final

Do teor das petições iniciais verificou-se que interessa a parte denominada de *pedido final* ou *requerimento*, ocasião em que se descreve o pleito da lide, o próprio objeto da demanda e que contém o texto que mais se assemelha à redação dos TEMAS 864 e 1177 do STF e TEMAS 986 e 1118 do STJ.

3.3 Pré-processamento

Essa etapa trabalha o texto referente a cada petição inicial de cada linha de dados, transformando-o de maneira a ser possível ser trabalhado pela técnica de aprendizado profundo, sendo utilizadas para isso, as bibliotecas de processamento de linguagem natural Spacy (<https://spacy.io/>) e NLTK (<https://www.nltk.org/>), que são as principais existentes (LAURIOLA, LAVELLI e AIOLLI, 2022).

Esse método usa uma sequência de texto e separa todas as suas palavras, removendo algumas pontuações. Também, é preciso levar em conta, que algumas palavras no *corpus* (conjunto de textos) de treino estarão muito presentes, como é o caso de preposições e artigos, são palavras que em geral não tem um significado.

Assim deve se proceder à exclusão de números, artigos, símbolos, pontuação, interjeições, e tudo mais que não contribui no resultado final. Finalmente, são coletadas as frequências de cada palavra e a frequência da co-ocorrência entre duas palavras no texto.

O objetivo é melhorar a qualidade dos dados para assim alcançar melhores resultados na classificação dos documentos.

As técnicas são:

- Padronização da coleção textual: foram removidos caracteres irrelevantes, como sinais de pontuação, caracteres especiais e alfanuméricos, assim como padronização dos textos em caixa baixa, ou seja, passaram a ter apenas caracteres minúsculos, para que palavras como “Subiu” e “subiu” não sejam vistas como palavras diferentes.
- Remoção de *stopwords*: nesta etapa ocorre a remoção de palavras pouco discriminativas, como artigos, pronomes, e advérbios, visto que estas palavras não expressam nenhum sentimento, e não contribuem para o significado de uma frase. Neste contexto, foi utilizado a lista de *stopwords* disponibilizada pela

biblioteca de processamento de linguagem natural Spacy, que contém as palavras consideradas irrelevantes em diversos idiomas, inclusive o português.

- Lematização: este processo tem como objetivo agrupar as variantes de um termo em um único lema, ou seja, transformar verbos para sua forma no infinitivo, e substantivos e adjetivos para o masculino singular. Por exemplo, as palavras “casa”, “casas”, “casinha” e “casarão” são todas formas do mesmo lema “casa”. Assim como as palavras “tiver”, “tinha”, “tenho” e “terei” são do mesmo lema “ter”. Outra técnica complementar é a stemização (do inglês *stemming*), que auxilia na redução da dimensionalidade dos atributos. O processo de *stemming* consiste em reduzir uma palavra ao seu radical, sua raiz, seu *stem*. A palavra “configurar” se reduziria a “confi”, assim como “configurando”, “configurado” e “configuração”. Para aplicar a técnica de lematização (*lemmatisation*), também foi utilizada a biblioteca Spacy e para a técnica de stemização (*stemming*) foi utilizada a biblioteca NLTK.
- Word embeddings: essa técnica objetiva representar uma palavra por meio de um vetor numérico de baixa dimensionalidade. A obtenção desse vetor é feita utilizando o contexto das palavras, isto é, palavras ao redor, de forma que palavras que ocorrem em contexto parecidos serão representadas por vetores parecidos.
- Tokenização é o processo que transforma frases em uma sequência de *tokens*, em outras palavras criamos um vetor que é a representação da frase em números.
- Finalmente, é preciso converter os tokens, que estão em formato de texto, para um formato numérico. Dados textuais representam valores discretos, e o classificador “só entende números”. Isto deve acontecer antes de se passar os dados para o classificador.
- Label encoding: modelos de Aprendizado de Máquina e Aprendizado Profundo exigem que todas as variáveis de entrada e saída da rede sejam numéricas. Dessa forma, foi utilizada a biblioteca Sklearn para aplicar a técnica *label encoding* na coluna de dados referente se o processo classifica ou não com o TEMA 1188 do

STJ, ou com o TEMA 986 do STJ, TEMA 864 do STF ou com o TEMA 1177 do STF, para um valor numérico (0, 1, 2, 3 e 4).

3.4 Divisão do Conjunto de Dados para Treinamento e Teste

Nessa etapa, o conjunto de dados será dividido em duas partes, uma para treino e outra para validação (teste) do modelo. O percentual dessa divisão será de 90% do *dataset* para treinamento e 10% do *dataset* para teste. Aqui também é importante manter o balanceamento (representatividade) de cada classe (atende ao TEMA 864 do STF ao TEMA 1177 do STF ao TEMA 986 do STJ ao TEMA 1118 do STJ ou nenhum dos quatro). Para isso é utilizada a técnica SMOTE apresentada na seção 2.8.

3.5 Treinamento das *Deep Learning* LSTM e BERT

Na fase de treinamento das redes LSTM e BERT é onde é realizado o aprendizado das técnicas. Os dados são apresentados à rede por um número de vezes (número de épocas) e espera-se que ao final, tenha-se atingido um determinado nível de aprendizado desejado, ou seja, que se obtenha um valor mínimo de uma medida de avaliação de aprendizado selecionada.

Existem uma série de hiperparâmetros que são especificados nesta fase, que serão apresentados com os respectivos valores no capítulo referente aos resultados computacionais e discussões.

3.6 Avaliação do Treinamento

Para realizar a avaliação da performance do treinamento, utilizou-se métricas e esquemas de avaliação comumente utilizados em problemas de classificação, como matriz de confusão. Uma matriz de confusão é um artefato usado para avaliar a performance de modelos destinados à realização de classificação em Machine Learning (Aprendizado de Máquina). Os valores que compõem a matriz são obtidos após a fase de treinamento do modelo, e a eficiência é medida ao fornecer os segmentos do conjunto de testes ao método de classificação e comparando as predições com os rótulos corretos dos exemplos de teste.

Caso, o treinamento tenha atingido um valor maior ou igual a métrica de desempenho estabelecida, passa-se para a próxima fase, que é testar a rede. Caso contrário, retorna-se à

etapa anterior de treinamento das redes LSTM e BERT de modo a buscar aperfeiçoar o seu aprendizado/treinamento (mudar a arquitetura da rede e/ou seus hiperparâmetros).

3.7 Teste das *Deep Learning* LSTM e BERT

Nesta etapa as *Deep Learning* LSTM e BERT são avaliadas com dados diferentes daqueles que foram utilizados na fase de aprendizado, ou seja, deseja-se verificar o desempenho das redes LSTM e BERT para dados desconhecidos com o intuito de conhecer se o processo de aprendizado foi efetivo.

Nesse caso, tendo as redes LSTM e BERT apresentado desempenho maior ou igual a um determinado valor, no processo de classificação correta de casos, passa-se para a próxima fase que é a de aplicação em produção. Caso contrário, sendo baixo o desempenho, retorna-se à etapa 3.5 de treinamento da rede em busca de aperfeiçoar o treinamento da rede e consequentemente o desempenho na fase de teste.

3.8 Modelo Gerado

Estando pronto o modelo, tanto a LSTM quanto a BERT, se aplica a toda nova demanda ajuizada, procurando o magistrado encontrar a similitude entre o novo pedido e algum dos TEMAS em questão (TEMAS 864 e 1177 do STF e TEMAS 986 e 1118 do STJ), com a finalidade de descobrir se o precedente obrigatório tem aplicabilidade ao caso concreto e assim proceder ao julgamento liminar de improcedência, ganhando com isso a celeridade na prestação jurisdicional, com segurança jurídica.

4. EXPERIMENTO COMPUTACIONAL E ANÁLISES DE RESULTADOS

Será possível complementar a capacidade do julgador não só na execução de sua tarefa principal de prestação jurisdicional, mas também na de armazenamento de dados consubstanciados nos precedentes obrigatórios.

Veja que estamos falando que a ferramenta tecnológica vem complementar, auxiliar o julgador até porque nenhum sistema tecnológico é autossuficiente a ponto de substituir o conhecimento e a experiência que o julgador adquire ao longo do exercício de sua função.

Por esse projeto será possível demonstrar, com o uso da inteligência computacional abordando a mineração de dados com processamento de linguagem natural, a ocorrência da

identidade entre a petição inicial do processo judicial com os precedentes dotados de eficácia vinculante, todavia essa ação de identificação será executada logo no início do procedimento, no nascedouro do processo judicial, quando for protocolada a petição inicial e por isso possibilitará a maior celeridade e segurança da prestação jurisdicional (ARGONDIZO, 2018).

A comprovação decorre objetivamente da possibilidade de redução do tempo médio de tramitação do feito judicial com o assunto data-base. Por exemplo, em 2021, o tempo médio de tramitação dos processos relacionados ao TEMA 864 do STF foi de 110 dias segundo informações contidas no processo SEI nº 20.0.000024198-0⁸, sendo que com o uso da ferramenta esse mesmo prazo poderia ser reduzido para apenas 01 dia. O mesmo ocorreria com os processos com aplicação do TEMA 1177 do STF ou com os TEMAS 986 e 1118 do STJ. Com isso, a média geral de 555 dias de duração (painel Estatística de 1º Grau⁹), englobando todo acervo em tramitação na vara, seria menor, proporcionando um redutor temporal, reforçando o compromisso da prestação jurisdicional célere e segura.

Registro que o julgador não estará obrigado a tomar nenhuma decisão caso seja encontrada a similitude de algum precedente vinculativo ao caso *sub judice*, entretanto estará ciente que acerca da matéria posta em discussão no processo já existe manifestação vinculativa de um órgão *ad quem* lhe competindo observar seus parâmetros no momento de proferir a solução final à lide.

4.1 Configuração do Experimento Computacional Realizado

Aqui, foi utilizado um conjunto de dados com 711 processos (300 com o TEMA 864, 28 com o TEMA 1177, 66 com o TEMA 986, 21 com o TEMA 1118 e 296 não relacionados a quaisquer dos TEMAS), totalizando 711 exemplos.

Nesse conjunto de dados foi executado o pré-processamento de texto, conforme descrito na seção 3.3.

Após o pré-processamento de texto, foi aplicada a técnica SMOTE conforme apresentada na seção 2.8, para equilibrar a quantidade de exemplos entre os diferentes temas/classes. Ao final, cada tema/classe tinha 300 exemplos, totalizando 1500 exemplos. Desses 1500 exemplos, foram divididos em 90% para conjunto de treinamento e 10% para

⁸https://sei.tjto.jus.br/sei/controlador.php?acao=arvore_visualizar&acao_origem=procedimento_visualizar&id_procedimento=10000003023026&infra_sistema=100000100&infra_unidade_atual=110001214&infra_hash=56a4ca7a09cd7155ad9ae94b203933563b22bfaa4c459c880231d0059d679b14.

⁹ https://bi.tjto.jus.br/extensions/Painel_Publico_Cenarius/Painel_Publico_Cenarius.html

conjunto de teste. Ressalta-se que foi mantida a proporção de ocorrência dos dados (estratificação) nos conjuntos de treinamento e teste.

Nas Figura 11 e Figura 12 são apresentadas seguidamente as arquiteturas da *Deep Learning LSTM* e da *Deep Learning BERT* utilizadas nesse trabalho.

Todos os experimentos foram realizados em uma máquina com sistema operacional Windows 10, processador Intel(R) Core(TM) i7-11800H de 2.30Ghz, 16 Gb de RAM DDR4-3200Mhz e GPU RTX3070.

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 100, 100)	5000000
spatial_dropout1d (SpatialD ropout1D)	(None, 100, 100)	0
lstm (LSTM)	(None, 100)	80400
dense (Dense)	(None, 5)	505
=====		
Total params: 5,080,905		
Trainable params: 5,080,905		
Non-trainable params: 0		

Figura 11. Arquitetura da *Deep Learning LSTM* utilizada. Faccioni, M. (2022).

Layer (type)	Output Shape	Param #
input_ids (InputLayer)	[(None, 30)]	0
bert (TFBertMainLayer)	TFBaseModelOutputWithPoolingAndCrossAttentions(last_hidden_state=(None, 30, 768), pooler_output=(None, 768), past_key_values=None, hidden_states=None, attentions=None, cross_attentions=None)	108923136
pooled_output (Dropout)	(None, 768)	0
tipodecisao02 (Dense)	(None, 25)	19225
tipodecisao01 (Dropout)	(None, 25)	0
tipodecisao (Dense)	(None, 5)	130
=====		
Total params: 108,942,491		
Trainable params: 108,942,491		
Non-trainable params: 0		

Figura 12. Arquitetura da *Deep Learning BERT* utilizada. Faccioni, M. (2022).

4.2 Análises e Discussões

Aqui são apresentados e analisados os resultados da aplicação da *Deep Learning LSTM* e *Deep Learning BERT* na classificação dos processos como sendo pertinentes ou não aos TEMAS 864, 986, 1118 e 1177.

4.2.1 Resultados do Experimento Realizado com a Deep Learning LSTM

Para uma análise do esforço computacional (tempo de execução), a *Deep Learning LSTM* levou 127.10 segundos na etapa de pré-processamento, 88.54 segundos na etapa de treinamento e 0.62 segundos na etapa de testes para os conjuntos de dados apresentados na Seção 4.1.

Para a *Deep Learning LSTM* com a configuração apresentada na Seção 4.1 e com a arquitetura apresentada na Figura 11, a mesma apresentou uma acurácia de 97% na fase de treinamento e uma acurácia de 90% na fase de testes. Para corroborar essa informação, na Figura 13 é apresentada a matriz de confusão (é uma tabela que descreve o desempenho do modelo de classificação) para os dados de teste (dados não vistos na etapa de treinamento).

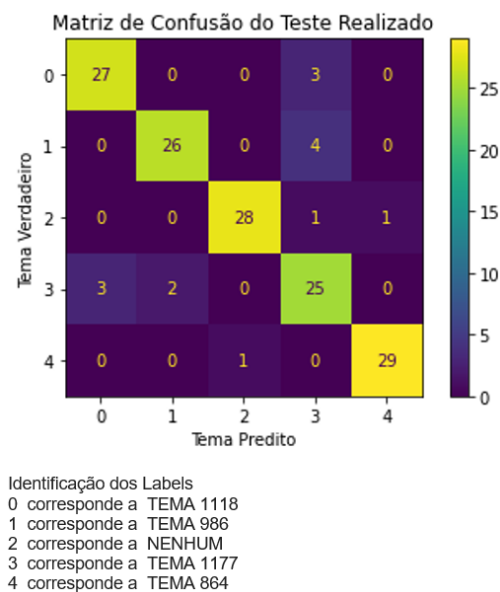


Figura 13. Matriz de confusão sobre os dados de teste da LSTM. Faccioni, M. (2022).

A matriz de confusão da Figura 13 mostra que para as 5 classes com 150 exemplos (processos), apenas 15 são classificados erroneamente (aproximadamente 10%), levando a uma precisão da classificação de 97% nos dados de teste. Aqui, os seguintes erros aconteceram: três exemplos da classe 0 (TEMA 1118) foram classificados erroneamente como da classe 3 (TEMA 1177), quatro exemplos da classe 1 foram classificados erroneamente como da classe 3 (TEMA 1177), dois exemplos da classe 2 (NENHUM) foram classificados erroneamente como sendo um da classe 3 (TEMA 1177) e outro da classe 4 (TEMA 864), cinco exemplos da classe 3 (TEMA 1177) foram classificados como três sendo da classe 0 (TEMA 1118) e outros dois sendo da classe 1 (TEMA 986) e finalmente 1 exemplo da classe 4 (TEMA 864) foi classificado erroneamente como sendo da classe 2 (NENHUM).

A partir dos dados da Figura 13, é gerado o relatório de classificação da Figura 14 com as métricas mais utilizadas para avaliação de desempenho de técnicas de classificação de inteligência computacional, que são *precision* (precisão), *recall* (sensitividade) e F1-Score.

Relatório de Classificação					
	precision	recall	f1-score	support	
0	0.90	0.90	0.90	30	
1	0.87	0.93	0.90	28	
2	0.93	0.97	0.95	29	
3	0.83	0.76	0.79	33	
4	0.97	0.97	0.97	30	
micro avg	0.90	0.90	0.90	150	
macro avg	0.90	0.90	0.90	150	
weighted avg	0.90	0.90	0.90	150	
samples avg	0.90	0.90	0.90	150	

Figura 14. Relatório de classificação de teste da LSTM, mostrando as medidas *precision*, *recall* e F1-Score. Faccioni, M. (2022).

Da Figura 14, no primeiro bloco de valores, são apresentadas as medidas de desempenho para cada uma das classes em separado. Já no segundo bloco de valores, são apresentadas as medidas médias sobre todas as classes. Todas essas métricas variam entre 0 e 1 e quanto mais próxima de 1 melhor. No caso em tela, o valor médio de 0.90 (*samples avg*) para todas as classes, mostra um bom desempenho da ferramenta proposta.

4.2.2 Resultados do Experimento Realizado com a *Deep Learning* BERT

Para uma análise do esforço computacional (tempo de execução), a *Deep Learning* BERT levou 198.28 segundos na etapa de pré-processamento, 652.30 segundos na etapa de treinamento e 0.92 segundos na etapa de testes para os conjuntos de dados apresentados na Seção 4.1.

Para a *Deep Learning* BERT com a configuração apresentada na Seção 4.1 e com a arquitetura apresentada na Figura 12, a mesma apresentou uma acurácia de 99.70% na fase de treinamento e uma acurácia de 99% na fase de testes. Para corroborar essa informação, na Figura 15 é apresentada a matriz de confusão (é uma tabela que descreve o desempenho do modelo de classificação) da *Deep Learning* BERT para os dados de teste (dados não vistos na etapa de treinamento).

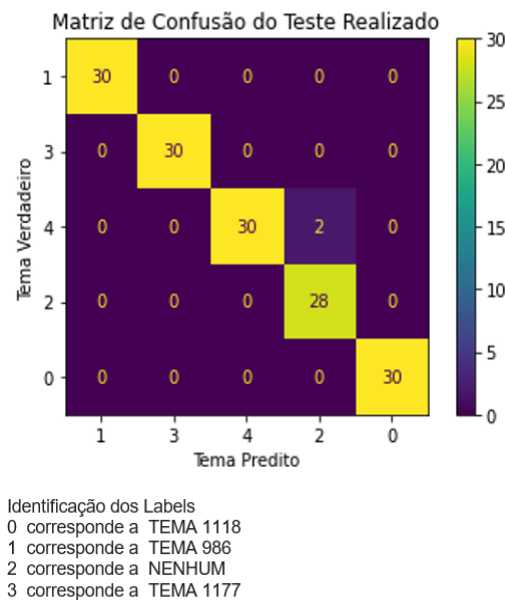


Figura 15. Matriz de confusão sobre os dados de teste da BERT. Faccioni, M. (2022).

A matriz de confusão da Figura 15 mostra que para as 5 classes com 150 exemplos (processos), apenas 2 são classificados erroneamente (aproximadamente 1%), levando a uma precisão da classificação de 99% nos dados de teste. Aqui, os seguintes erros aconteceram: dois exemplos da classe 2 (NENHUM) foram classificados erroneamente como da classe 4 (TEMA 864).

A partir dos dados da Figura 15, é gerado o relatório de classificação da Figura 16 com as métricas mais utilizadas para avaliação de desempenho de técnicas de classificação de inteligência computacional, que são *precision* (precisão), *recall* (sensitividade) e F1-Score.

Relatório de Classificação				
	precision	recall	f1-score	support
0	1.00	1.00	1.00	30
1	1.00	1.00	1.00	30
2	0.93	1.00	0.97	28
3	1.00	1.00	1.00	30
4	1.00	0.94	0.97	32
accuracy			0.99	150
macro avg	0.99	0.99	0.99	150
weighted avg	0.99	0.99	0.99	150

Figura 16. Relatório de classificação de teste da BERT, mostrando as medidas *precision*, *recall* e F1-Score. Faccioni, M. (2022).

Da Figura 16, no primeiro bloco de valores, são apresentadas as medidas de desempenho para cada uma das classes em separado. Já no segundo bloco de valores, são apresentadas as medidas médias sobre todas as classes. Todas essas métricas variam entre 0 e 1 e quanto mais próxima de 1 melhor. No caso em tela, o valor médio de 0.99 (*samples avg*) para todas as classes, mostra um excelente desempenho da ferramenta proposta.

4.2.3 Conclusão a Respeito dos Resultados Obtidos

Observando os resultados das *Deep Learning* LSTM e BERT apresentados nas Seções 4.2.1 e 4.2.2 respectivamente, obtém-se os seguintes apontamentos:

- A *Deep Learning* BERT apresenta maior tempo computacional que a *Deep Learning* LST, em todas as fases (treinamento, teste e previsão/estimação), principalmente na fase de treinamento.
- Já na fase de previsão/estimação *Deep Learning* LSTM apresenta tempo computacional de 15ms e a *Deep Learning* BERT apresenta tempo computacional de 98ms.
- A *Deep Learning* BERT apresentou melhor desempenho que a *Deep Learning* LSTM em todas as métricas de avaliação, em especial, apresentando um rendimento médio 99% da BERT contra 90% da LSTM na fase de testes, sendo a BERT 9% melhor, caracterizando um desempenho superior.

A partir dos apontamentos, verifica-se que a *Deep Learning* BERT, apresenta maior tempo computacional que a *Deep Learning* LSTM em todas as fases (treinamento, teste e previsão/estimação), todavia, na fase de previsão/estimação, ocasião na qual se identifica na

petição inicial a ocorrência ou não de algum precedente judicial, constata-se uma diferença desprezível entre as redes (98ms para BERT e 15ms para LSTM = 83ms) para o usuário final. Não obstante, o desempenho da Deep Learning BERT (99%) apresentou desempenho 9% superior ao da Deep Learning LSTM (90%) para determinar se há ou não a ocorrência de algum precedente vinculativo e, caso exista, indicará qual.

Assim sendo, tem-se como conclusão que a *Deep Learning* BERT apresenta resultado superior a *Deep Learning* LSTM para o problema em questão e sendo apropriada para previsão/estimação de TEMAS a partir de exordiais apresentadas.

Posteriormente, foi realizada uma avaliação adicional com 100 processos diferentes dos utilizados nas fases de treinamento e de teste. Nesse caso, a acurácia obtida foi de 99%, demonstrando a alta taxa de acerto da ferramenta proposta. Ressaltando que o tempo de processamento de cada processo individual, leva em média 3.57 segundos, mostrando a alta eficiência da ferramenta proposta na fase de análise e identificação de similitude de algum TEMA com a petição inicial.

A ferramenta tecnológica proposta, dada uma petição inicial, a mesma primeiramente passa por uma fase de pré-processamento de texto, conforme descrito na Seção 3.3. Após, esses dados pré-processados do texto são repassados à ferramenta tecnológica para a *Deep Learning* utilizada realizar a previsão/estimação. Ao final, é apresentado se NENHUM ou qual TEMA é aplicado ao processo analisado. Caso algum TEMA seja aplicado, também é oferecido o *link* de consulta que leva ao site oficial do STF ou STJ com todas as informações correspondentes ao precedente. Na sequência é transcrita na íntegra a seção dos pedidos, requerimentos ou requerimentos finais do processo em questão.

Tomando como base um processo/petição inicial qualquer, a mesma foi submetida à ferramenta tecnológica proposta e ao final, apresentou resultado de previsão/estimação, conforme a Figura 17.

```

1/1 [=====] - 0s 344ms/step
Tempo de Execução da Estimacão = 0.41942644119262695

==>Tema ou Súmula Aplicada ao Processo Analisado= TEMA 864

==>Link STF referente ao TEMA 864 = https://portal.stf.jus.br/jurisprudenciaRepercussao/verAndamentoProcesso.asp?incidente=4820176&numeroProcesso=905357&classeProcesso=RE&numeroTema=864

==>TEXTO DO PEDIDO= Ante todo o exposto, requer de Vossa Excelência: a) Nos moldes do Art. 319, inciso VII, o requerente desde já opta pela não realização de audiência de conciliação ou de mediação, visto a inquestionabilidade dos direitos pleiteados, e por não haver previsão em legislação estadual para que a Administração Pública possa fazer acordo judicial; b) Seja o Estado do Tocantins, representado pela Procuradoria Geral do Estado, citado para, querendo, apresentar resposta à presente demanda, nos termos da lei adjetiva civil, sob pena de revelia e confissão; ---- Edifício JK Business Center, 103 Sul, Av. JK, ACSO 1, CONJ. 01, Lote 41, Sala 1.205, Plano Diretor Sul ---- Palmas - TO - CEP: 77.015-012 - Telefone: (63) 3322-1829 Abreu, Cardoso & Gomes Advogados Associados S/S - OAB/TO 220 Abel Cardoso de Souza Neto - OAB/TO 4156 Bernardino de Abreu Neto - OAB/TO 4232 Rogério Gomes Coelho - OAB/TO 4155

c) Que no mérito seja julgada procedente a ação para que o Requerido seja condenado a pagar o valor das diferenças salariais devidas ao requerente relativo ao período de abril de 2015 a dezembro de 2016 (tendo em vista lapso prescricional), sendo que estas diferenças de salários são entre os valores efetivamente pagos e os que deveriam ter sido pagos, na conformidade das tabelas de subsídio previstas nos Anexos I a VI da Lei Estadual nº 2.884/2014, sendo as tabelas dos anexos II e V, referente ao período de janeiro a dezembro de 2015 e ainda, nas tabelas dos anexos III e VI, referente ao período de janeiro a dezembro de 2016, com reflexos sobre 13º Salário, 1/3 de férias, tendo em vista o lapso prescricional cobrar-se-á o período de abril de 2015 a dezembro de 2016, no valor total de R$ 59.991,76 (cinquenta e nove mil, novecentos e noventa e um reais e setenta e seis centavos), conforme se observa da Planilha de Cálculo em anexo; custas processuais e honorários de sucumbência, nos termos dos artigos 82 d) Ainda, que o Requerido condenado ao pagamento das e seguintes do NCPC, do total em atraso; e) Requer que seja concedido por Vossa Excelência os benefícios da Justiça Gratuita, caso não seja esse o entendimento, que o pagamento das custas processuais e taxa judiciária seja feito ao final do processo como medida JUSTA. Por fim, caso não seja este o entendimento, que seja parcelado na forma do art. 98, § 6º do CPC; f) Ainda, requer que a demanda seja julgada no estado em que se encontra, uma vez que não se encontram mais provas a serem apresentadas e ser demanda apenas de direito, na forma do art. 355, inciso I, do CPC e art. 5º, inciso LXXVIII da Constituição Federal. Processo Civil, declaro autênticos todos os documentos acostados na Nos termos do inciso IV, do art. 425 do Novo Código de presente peça, sob responsabilidade pessoal. Protesta a requerente por todos os meios de provas ---- Edifício JK Business Center, 103 Sul, Av. JK, ACSO 1, CONJ. 01, Lote 41, Sala 1.205, Plano Diretor Sul ---- Palmas - TO - CEP: 77.015-012 - Telefone: (63) 3322-1829 Abreu, Cardoso & Gomes Advogados Associados S/S - OAB/TO 220 Abel Cardoso de Souza Neto - OAB/TO 4156 Bernardino de Abreu Neto - OAB/TO 4232 Rogério Gomes Coelho - OAB/TO 4155 admitidos em direito, especialmente através de prova documental, depoimento pessoal e demais provas admitidas em direito (art. 369 CPC), desde já ficam requeridas. nove mil, novecentos e noventa e um reais e setenta e seis centavos). Dá-se à causa o valor de R$ 59.991,76 (cinquenta e termos em que pede e espera deferimento.

```

Figura 17. Resultado apresentado pela ferramenta desenvolvida, após a classificação do texto referente ao processo em questão. Faccioni, M. (2022).

Na Figura 17, são apresentados o tempo de execução da estimacão em segundos (aproximadamente 0.42 segundos), o TEMA que se aplica à petição inicial em questão, assim como o link do STF ou STJ referente ao precedente, com todas suas informações, inclusive redação da tese. Após, apresenta também a seção dos pedidos, requerimentos ou requerimentos finais da petição inicial em análise.

A apresentação do link do STF ou STJ, como também a seção dos pedidos, requerimentos ou requerimentos finais da petição inicial se justifica, para facilitar a análise e assertividade da ferramenta tecnológica pelo julgador, lembrando que a mesma busca tornar o julgamento mais célere e eficiente e não substituir o magistrado na prática de seus atos.

5. CONCLUSÕES E TRABALHOS FUTUROS

Restou demonstrado que o uso da inteligência computacional ajudará o julgador a descobrir a incidência de algum precedente de observância obrigatória logo no nascedouro do processo, bem após o protocolo da petição inicial. A ferramenta proposta busca auxiliar o magistrado na tarefa de encontrar a incidência de algum precedente na demanda que está sob sua análise de maneira célere e eficiente. A mesma proporciona isso em menos de 1 segundo

na máquina utilizada. Apresentando, caso exista, o TEMA incidente, o *link* oficial da origem da consulta da jurisprudência no respectivo tribunal, como também a seção Pedidos, Requerimentos ou Requerimentos Finais da exordial. Não obstante, a ferramenta proposta (baseada na *Deep Learning* BERT) apresentou uma margem de acerto de 99% em dados não vistos (teste), apresentando um excelente desempenho.

Aqui não queremos criar uma ferramenta tecnológica que irá proferir decisões judiciais em substituição ao julgador, mas uma forma de auxílio tecnológico na busca de uma prestação jurisdicional célere e segura. Ensinamos a máquina a analisar o corpo da petição inicial e compará-lo com o texto dos precedentes de observância obrigatória para que o magistrado não deixe de aplicar ao caso decisão já pacificada ou que profira resultado conflitante para situação de direito análoga.

A ferramenta apenas auxilia o magistrado em encontrar a similitude entre o pedido inicial e a existência de algum precedente de observância obrigatória, não substituindo o julgador na análise da lide, muito menos impondo a prática de algum ato decisório sobre a demanda. Sempre caberá ao magistrado analisar a ocorrência de algum precedente vinculante ao caso concreto e assim adotar, se for o caso, o julgamento de improcedência liminar do pedido, tornando mais segura, célere e produtiva a prestação jurisdicional.

Como foi salientado anteriormente, a vasta gama de precedentes dificulta o domínio do seu conteúdo de forma geral pelo julgador, ainda mais quando se trata de membro novo na carreira ou quando ocorre alteração de competência acerca da matéria na vara onde atua o magistrado, situações em que o auxílio da ferramenta tecnológica se mostra imprescindível para o bom e célere exercício da jurisdição.

Nesse momento foram utilizados apenas quatro precedentes obrigatórios, devendo a ferramenta ser configurada/carregada com os outros milhares que existem, não sendo necessário que se direcione a comparação a determinado assunto indicado na autuação, mas que a conferência seja feita de maneira global.

A cada nova inserção de algum precedente caberá o aprendizado da máquina a fim de se buscar a melhor resposta possível, havendo a necessidade de se atentar aos termos mais similares.

Nesse sentido, o desenvolvimento de uma ferramenta multidisciplinar que mescla direito, inteligência artificial e design de sistemas contribui, certamente, para que outras áreas do Judiciário, e da Administração Pública como um todo, possam aprimorar seus métodos de trabalho, tornando-os mais ágeis e eficientes.

Como trabalhos futuros, pretende-se implementar efetivamente a ferramenta tecnológica desenvolvida nesse trabalho ao sistema EProc/TJTO tanto em primeiro como também em segundo graus de jurisdição, fazendo os ajustes necessários para que sejam incluídos na ferramenta tecnológica todos os milhares de precedentes vinculativos, para que auxilie efetivamente os magistrados na distribuição de uma jurisdição mais célere, sendo importante conscientizar o Tribunal de Justiça do Estado do Tocantins acerca da necessidade de implantação e funcionamento de um laboratório de inteligência artificial para desenvolvimento e criação de outros instrumentos digitais.

Essa inserção e atualização constante dos precedentes na ferramenta tecnológica deverá ser feita e acompanhada pelo Núcleo de Gerenciamento de Precedentes (NUGEP), sendo o órgão responsável pela catalogação e divulgação dos precedentes judiciais formados pelo julgamento de demandas repetitivas e incidentes de assunção de competência, bem como pela padronização de procedimentos administrativos destinados à aplicação desses precedentes nos processos judiciais pendentes e futuros.

No Tribunal de Justiça do Estado do Tocantins, o Núcleo de Gerenciamento de Precedentes (NUGEP-TJTO) foi criado por meio da Resolução n. 16/2017, de 22 de junho de 2017, publicada no Diário da Justiça eletrônico n. 4063 de 26 de junho de 2017. A Resolução foi editada em observância à Resolução n. 235, de 13 de julho de 2016, do Conselho Nacional de Justiça.

Sugere-se também como trabalho futuro, a proposta de dividir os precedentes vinculativos (conforme descrito na Seção 2.3) de acordo com os assuntos dos processos, em cada vara judicial. Com isso se buscará adaptar a ferramenta tecnológica, treinando a *Deep Learning* BERT, somente com os precedentes vinculativos correspondentes à matéria da competência daquela vara, buscando, assim, um desempenho aprimorado, dado que o número de precedentes vinculativos aplicáveis a uma determinada vara é bem menor que o total existente.

REFERÊNCIAS

ARGONDIZO, Luís Fernando Centurião; DE LIMA, Wellington Henrique Rocha. **A improcedência liminar do pedido como garantidor da celeridade processual, uma análise crítica ao novo código de processo civil**. Revista de Processo, Jurisdição e Efetividade da Justiça, v. 4, n. 2, p. 100-114, 2018.

BARBOZA, Estefânia Maria de Queiroz. **Precedentes judiciais e segurança jurídica: fundamentos e possibilidades para a jurisdição constitucional brasileira**. Saraiva Educação SA, 2017.

BELL, J. Machine Learning: **Hands-On for Developers and Technical Professionals**. Indianápolis: John Wiley & Sons, Inc, 2015.

CAMILO, Cássio Oliveira; SILVA, João Carlos da. **Mineração de dados: Conceitos, tarefas, métodos e ferramentas**. Universidade Federal de Goiás (UFG), v. 1, n. 1, p. 1-29, 2009.

CARVALHO, Samantha de Araújo. **Do julgamento liminar de improcedência e os precedentes obrigatórios no CPC de 2015**. Direito UNIFACS–Debate Virtual, n. 209, 2017.

CHAMSEDDINE, Ekram et al. **Handling class imbalance in COVID-19 chest X-ray images classification: Using SMOTE and weighted loss**. Applied Soft Computing, v. 129, p. 109588, 2022.

CHARU, C Aggarwal et al. **Neural networks and deep learning**. Springer, 2018.

CHAWLA, Nitesh V. et al. **SMOTE: synthetic minority over-sampling technique**. Journal of artificial intelligence research, v. 16, p. 321-357, 2002.

DA SILVA, Fernanda Júnia Santana; DE OLIVEIRA, Fabiana Noronha. **Juizados especiais cíveis: Auxiliando a celeridade do judiciário**. ANAIS SIMPAC, v. 6, n. 1, 2016.

DE AMO, Sandra. **Técnicas de mineração de dados**. Jornada de Atualização em Informática, p. 26, 2004.

DE CASTRO CATHARINA, Alexandre. **Julgamento liminar e a duração razoável do processo: ponderação entre efetividade da atividade jurisdicional e celeridade processual**. Revista Interdisciplinar do Direito-Faculdade de Direito de Valença, v. 19, n. 1, p. 11-27, 2021.

DE OLIVEIRA BIANCHI, Giovana. **O princípio do contraditório é afetado quando o juiz profere a improcedência liminar do pedido?** Revista do Curso de Direito, v. 15, n. 15, p. 133-142, 2020.

DE OLIVEIRA NETO, João Mendes; TONIN, Sávio Duarte; PRIETCH, Soraia Silva. **Processamento de Linguagem Natural e suas Aplicações Computacionais**. Acesso em, v.12, 2010.

DE OLIVEIRA NETO, João Mendes; TONIN, Sávio Duarte; PRIETCH, Soraia Silva. **Processamento de Linguagem Natural e suas Aplicações Computacionais**. Acesso em, v. 12, 2010.

DENG, Li; Yu, Dong. **Aprendizagem profunda: métodos e aplicações**. Fundamentos e tendências no processamento de sinais, v. 7, n. 3-4, pág. 197-387, 2014.

DEVLIN, Jacob; CHANG, Ming-Wei; KENTON, Lee; TOUTANOVA, Kristina. **Bert: Pre-training of deep bidirectional transformers for language understanding**. In: Proceedings of naacL-HLT. 2019. p. 4171-4186.

DONIZETTI, Elpídio. **Curso didático de direito processual civil**. 20. Ed. rev. atual. e ampl. São Paulo: Atlas, 2017.

FARACO, Fernando Melo; TODESCO, José Leomar. **Inteligência artificial aplicada ao direito: uma revisão sistemática da literatura**: Uma revisão sistemática da literatura. In: Anais do Congresso Internacional de Conhecimento e Inovação–ciki. 2018.

FINATTO, Maria José Bocorny; LOPES, Lucelene; SILVA, Alena Ciulla. **Processamento de Linguagem Natural, Linguística de Corpus e Estudos Linguísticos**: uma parceria bem-sucedida. Domínios de linguagem. Uberlândia, MG. Vol. 9, n. 5 (dez. 2015), p.[41]-59, 2015.

FREIRE JUNIOR, Nilson de Souza. **O sistema de precedentes judiciais no ordenamento jurídico brasileiro sob a ótica dos princípios da segurança jurídica e celeridade processual**. 2021.

GOMES, D. D.; EVSUKOFF, A. G. **Processamento de linguagem natural em Português e aprendizagem profunda para o domínio de óleo e gás**. 2019.

GONZALEZ, Marco; LIMA, Vera Lúcia Strube. **Recuperação de informação e processamento da linguagem natural**. In: XXIII Congresso da Sociedade Brasileira de Computação. 2003. p. 347-395.

GONZÁLEZ-CARVAJAL, Santiago; GARRIDO-MERCHÁN, Eduardo C. **Comparing BERT against traditional machine learning text classification**. arXiv preprint arXiv:2005.13012, 2020.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep learning**. MIT press, 2016.

GREFF, Klaus et al. (2016). «LSTM: A search space odyssey. Em: IEEE transactions on neural networks and learning systems 28.10, pp. 2222–2232.

HE, Haibo; GARCIA, Edwardo A. **Learning from imbalanced data**. IEEE Transactions on knowledge and data engineering, v. 21, n. 9, p. 1263-1284, 2009.

JÚNIOR, Antônio Pereira Gaio. **Considerações acerca da compreensão do modelo de vinculação às decisões judiciais:** os precedentes no novo código de processo civil brasileiro. *Revista Interdisciplinar de Direito*, v. 13, n. 2, 2017.

LAURIOLA, Ivano; LAVELLI, Alberto; AIOLLI, Fabio. **An introduction to deep learning in natural language processing:** Models, techniques, and tools. *Neurocomputing*, v. 470, p. 443-456, 2022.

NASCIMENTO, Felipe Mozart Santana; CORRÊA, Renato Fernandes; PINHO, Fábio Assis. **Percorso metodológico para construção de Ontologias Jurídicas.** *Informação & Sociedade*, v. 29, n. 4, p. 135, 2019.

OLIVEIRA, Fabiana Luci de; CUNHA, Luciana Gross. **Os indicadores sobre o Judiciário brasileiro:** limitações, desafios e o uso da tecnologia. *Revista Direito GV*, v. 16, 2020.

PASCHOAL, Gustavo Henrique; ANDREOTTI, Paulo Antonio Brizzi. **Considerações sobre o Sistema de Precedentes Judiciais no Novo Código de Processo Civil.** *Revista Juris UniToledo*, v. 3, n. 04, 2018.

RONDEL, Giovani Garuffi; HILKNER, Rodrigo Dahlstrom. **Técnicas de Redes Neurais para Análise e Previsão do Mercado de Ações.** Relatório Técnico, Projeto Final de Graduação, Unicamp, 2019.

SERRANO, A. G. **Inteligência Artificial. Fundamentos, prática e aplicações.** Madrid: RC Libros, 2012.

SOUZA, Fábio; NOGUEIRA, Rodrigo; LOTUFO, Roberto. **BERTimbau: pretrained BERT models for Brazilian Portuguese.** In: *Brazilian conference on intelligent systems*. Springer, Cham, 2020. p. 403-417.

STEMLER, Igor Tadeu Silva Viana. **Identificação de precedentes judiciais por agrupamento utilizando processamento de linguagem natural.** Dissertação. Mestrado Profissional em Computação Aplicada. Universidade de Brasília, Brasília. 2019.

THEODORO JÚNIOR, Humberto. **Celeridade e efetividade da prestação jurisdicional.** Insuficiência da reforma das leis processuais. *Revista Síntese de Direito Civil e Processual Civil*, Porto Alegre, v. 1, n. 36, p.19-37, Jul. /Ago. p. 27, 2005.

Pesquisas eletrônicas:

- <http://www.stf.jus.br>
- <http://www.stj.jus.br>
- <http://www.tjto.jus.br>
- <http://portal.acm.org>
- <http://www.sciencedirect.com/>
- <http://iee.org/web/publications/xplore/>

- <http://www.emeraldinsight.com>
- <http://scholar.google.com.br>
- <http://www.interscience.wiley.com>
- <http://www-periodicos-capes-gov-br.ez361.periodicos.capes.gov.br/>
- <https://spacy.io/>
- <https://www.nltk.org/>