

ARQUIVO JURÍDICO

REVISTA JURÍDICA ELETRÔNICA DA UFPI



ISSN
2317-918X

V. 11, N. 1
JAN./JUN.
2024

QUALIS
B2

©PPGD/UFPI

Uma publicação do Programa de Pós-Graduação em Direito da UFPI. Todos os direitos reservados. Proibida a reprodução parcial ou total sem consentimento expresso dos editores. As opiniões emitidas nos artigos assinados são de total responsabilidade de seus autores.

Artigos para possível publicação devem ser encaminhados exclusivamente pelo portal de periódicos da UFPI (<https://revistas.ufpi.br>), com o prévio cadastramento do autor.

Dados Internacionais de Catalogação na Publicação (CIP)

Arquivo Jurídico – Revista Jurídica Eletrônica da Universidade Federal do Piauí / Programa de Pós-Graduação em Direito da UFPI, v. 1, n. 1 (jul./dez. 2011).
Teresina: Programa de Pós-Graduação em Direito da UFPI, 2011-

Trimestral

ISSN: 2317-918X (versão digital)

1. Direito – periódicos. I. Programa de Pós-Graduação em Direito da UFPI.

SOBRE A REVISTA

A Arquivo Jurídico – Revista Jurídica Eletrônica da Universidade Federal do Piauí, em circulação desde 2011, é o periódico acadêmico digital semestral do Programa de Pós-Graduação em Direito da UFPI, cujo objetivo é fomentar e difundir o intercâmbio de conhecimentos das áreas jurídicas e afins. Acesso eletrônico livre pelo portal <https://revistas.ufpi.br>. Avaliado no estrato B2 pela Qualis / CAPES (2020).

Solicita-se permuta.
Pídese canje.
On demande l'échange.
Si richiede lo cambio.
We ask for Exchange.
Wir bitten um austausch.

ARQUIVO JURÍDICO
Revista Jurídica Eletrônica da
Universidade Federal do Piauí
Periódico acadêmico oficial do Programa de
Pós-Graduação em Direito
Semestral
ISSN 2317-918X
<https://revistas.ufpi.br/>

CLASSIFICAÇÃO DE PRONUNCIAMENTOS JUDICIAIS UTILIZANDO PROCESSAMENTO DE LINGUAGEM NATURAL

AUTOMATIC CLASSIFICATION OF JUDICIAL STATEMENTS USING NATURAL LANGUAGE PROCESSING

Fernando Maurilio Bento

Mestrando na Universidade do Vale do Itajaí
<http://lattes.cnpq.br/0098491037950843>

Alexandre Moraes da Rosa

Doutor pela Universidade Federal do Paraná
Professor da Universidade do Vale do Itajaí
<http://lattes.cnpq.br/4049394828751754>

Raimundo Celeste Ghizoni Teive

Doutor pela Universidade Federal de Santa Catarina
Professor da Universidade do Vale do Itajaí
<http://lattes.cnpq.br/6094793991932450>

Resumo: Processamento de Linguagem Natural (PLN) tem sido aplicado à área jurídica como uma forma de otimizar recursos, utilizando a Inteligência Artificial para executar tarefas que até pouco tempo eram feitas apenas por humanos. Na área jurídica, composta basicamente por textos não estruturados, a classificação pode ser um aliado na organização e na gestão destes documentos. Neste trabalho, foram classificados pronunciamentos judiciais nas classes previstas no artigo 203 do Código de Processo Civil: sentenças, decisões interlocutórias e despachos. A principal contribuição deste artigo reside na demonstração que o desempenho do classificador está diretamente relacionado com a seleção prévia de palavras e o balanceamento de classes utilizadas. Resultados preliminares apontam aumento em mais de 20% da acurácia, quando é considerada a representatividade de cada palavra no corpus e a frequência da palavra em cada classe, alcançando resultados superiores a 80%.

Palavras-chave: Classificação de textos jurídicos. Processamento de linguagem natural. Inteligência artificial.

Abstract: Natural Language Processing (NLP) has been applied to the legal field as a way of optimizing resources, using Artificial Intelligence to perform tasks that until recently were only done by humans. In the legal area, basically composed of unstructured texts, classification can be an ally in the organization and management of these documents. In this work, judicial pronouncements were classified in the classes provided for in article 203 of the civil procedure code: sentences, interlocutory decisions and orders. The main contribution of this article lies in the demonstration that the performance of the classifier is directly related to the previous selection of words and the balancing of classes used. Preliminary results point to an increase of more than 20% in accuracy, when considering the representativeness of each word in the corpus and the frequency of the word in each class, reaching results above 80%.

Keywords: Legal texts classification. Natural language processing. Artificial intelligence.

Submetido em 12 de dezembro de 2022. Aprovado em novembro de 2023.

1 INTRODUÇÃO

Segundo Alves (2017), pronunciamentos judiciais são declarações verbais e públicas de juízes no intuito de resolver questões, de fato ou direito, ou, simplesmente, sem nada decidir, impulsionam a marcha processual para frente até a solução definitiva da controvérsia. Conforme o novo Código de Processo Civil, Lei nº 13.105/2015, atos decisórios, tratados neste trabalho como pronunciamentos judiciais, consistem em sentenças, decisões interlocutórias e despachos. Considerando que todo o sistema recursal é dependente da interpretação correta de um pronunciamento judicial, a classificação entre os três tipos previstos no Código de Processo Civil é crucial para o bom andamento de um processo, especialmente porque há restrição recursal e incidência do fenômeno preclusivo no caso de inércia da parte.

A classificação de textos, além de outras diversas atividades, que, segundo Araújo e Simioni (2021) têm sido substituídas pela Inteligência Artificial, passou a ser uma das finalidades do Processamento de Linguagem Natural (PLN). PLN é conhecido como a tecnologia que estuda os problemas da geração e compreensão automática de línguas humanas naturais pelos dispositivos tecnológicos. Segundo Mcroy (2021), PLN é a disciplina de projeto e uso de programas de computador para analisar ou gerar linguagem humana.

Na área jurídica, o PLN, dentre outras finalidades, tem sido utilizado para classificação de documentos, ou seja, com base em um texto, é possível que o computador identifique qual é a sua classe ou categoria, a exemplo do trabalho de Polo, Giochetti e Bertolo (2021). Considerando que documentos jurídicos não são estruturados, a exemplo de leis, decisões judiciais e petições, a aplicação do PLN pode ser um importante aliado na otimização de recursos, agilizando tarefas.

A utilização de modelos preditivos no domínio do Direito é aspiração relativamente antiga, conforme Lawlor (1963), com o impulso decorrente da ampliação de acesso aos dados conferida pela internet e a web, associada à capacidade de processamento das máquinas, tanto no ocidente, quanto no oriente, especialmente na China consoante indicam XU, Jianfeng; SUN, Fuhui; CHEN, Qiwei; WANG, Yingfei; YU, Jia; LIU, Zhenyu. Segundo Hartmann Peixoto, os modelos de Inteligência Artificial incidentes no domínio do Poder Judiciário foram regulamentados pelo Conselho Nacional de Justiça por meio da Resolução 332/2020, reconhecendo-se apenas modelos supervisionados de classificação ou de regressão. A preocupação quanto aos aspectos éticos das decisões judiciais, centradas na dignidade da pessoa humana, levam em consideração questões de transparência algorítmica e explicabilidade dos provimentos jurisdicionais.

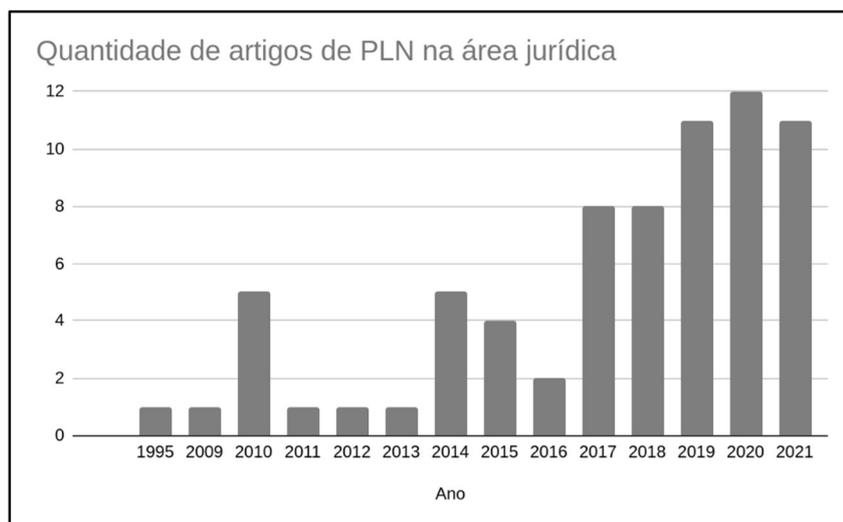
Neste contexto, o objetivo deste estudo é explorar as etapas e técnicas de PLN para classificar, de forma automática, pronunciamentos judiciais,

resultando nos tipos possíveis: sentença, decisão interlocutória e despacho. As próximas seções apresentarão os trabalhos relacionados, os procedimentos metodológicos, a base de dados, as etapas de pré-processamento e treinamento, a análise dos resultados e as conclusões.

2 TRABALHOS RELACIONADOS

Conforme observado nas bases de dados indexadas pelo Portal de Periódicos da CAPES, pesquisas aplicando PLN na área jurídica vêm sendo realizadas desde 1995. Desde então, o tema foi explorado de forma esporádica até final de 2016, quando então passou a figurar em maior volume e consistência a partir de 2017, em especial a partir de 2019, apresentando claramente uma tendência de crescimento de pesquisa sobre o tema, conforme ilustrado pelo gráfico 1.

Gráfico 1: Evolução de trabalhos científicos do PLN aplicado à área jurídica



Os trabalhos relacionados agrupam-se com objetivos distintos, podendo ser diferenciados pelas aplicações que podem ser dadas para o PLN. Galani et al (2015), Aletras (2016), Katz (2017), Shulayeva (2017), Petrovic (2018), Medvedeva (2019), Ji (2020) I, Li (2021) e Polo, Ciochetti e Bertolo (2021) trabalharam com classificação, em geral, buscando prever decisões judiciais em processos judiciais. Fernandes (2019), Ji (2020) II, Adewoyin (2020) e Trias (2021) trabalharam sobre a extração de informação, transformando textos em dados. Anand (2019) trabalhou em sumarização, reduzindo textos aos seus conteúdos mais relevantes. Por fim, as pesquisas de Wagh (2020), Huang (2021) e El Jelali (2015) trataram sobre a similaridade entre processos.

Também é possível agrupar os trabalhos correlatos pelas técnicas e algoritmos utilizados, sendo Bidirectional Long Short-Term Memory (Bi-

LSTM), Support Vector Machine (SVM) e algoritmos baseados em Transformers as tecnologias mais utilizadas. Deste grupo vale destacar que o ano mais recente, 2021, foi marcado pela utilização de algoritmos baseados em Transformers, indicando um possível novo estado da arte no que diz respeito ao PLN.

Em linhas gerais, os resultados encontrados nestes trabalhos de pesquisa são positivos, demonstrando que há espaço para aprimoramento das abordagens utilizadas, num tema de pesquisa que apresenta uma tendência de alta no seu volume de pesquisas.

3 PROCEDIMENTOS METODOLÓGICOS

Esta seção apresenta como o trabalho foi desenvolvido para atingir os seus objetivos.

Inicialmente, foi aplicado o pré-processamento em toda a base de dados, originando a base de dados pré-processada para a aplicação dos algoritmos.

A base pré-processada, ordenada em uma sequência aleatória, foi distribuída para suportar três simulações, sendo $\frac{2}{3}$ dos dados utilizados para treinamento e $\frac{1}{3}$ para teste, de forma que a base de teste não tivesse nenhum registro repetido entre as simulações.

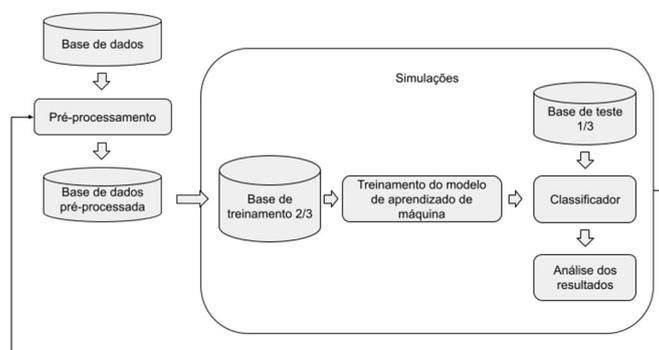
Foram então aplicadas as três simulações.

Após a análise dos resultados, novos ciclos de pré-processamento foram realizados, utilizando critérios para seleção das características consideradas no treinamento dos modelos, até que os critérios não produzissem melhora nos resultados em pelo menos um algoritmo.

Por fim, realizou-se o comparativo entre os resultados antes e depois da aplicação dos critérios de seleção.

A figura 1 ilustra o encadeamento entre os procedimentos e as bases de dados geradas ao longo do trabalho.

Figura 1. Procedimento metodológico



4 BASE DE DADOS

A base de dados utilizada para este estudo, em PLN também conhecida como Corpus, significa um conjunto de textos ou documentos, composta por pronunciamentos judiciais de diversos Tribunais. A base de dados foi obtida diretamente dos Tribunais por meio de *web crawling*. Aleatoriamente foram selecionados 235 pronunciamentos judiciais de diversos órgãos julgadores, conforme exemplificado no quadro 1.

Quadro 1. Exemplos de pronunciamentos judiciais

Exemplo 1 (Decisão Interlocutória):

“Sr. Advogado, TRIBUNAL DE JUSTICA DE SANTA CATARINA - COMARCA DE JOINVILLE JUÍZO DE DIREITO DA VARA CÍVEL JUIZ(A) DE DIREITO C.T. ESCRIVÃ(O) JUDICIAL S.F. EDITAL DE INTIMAÇÃO DE ADVOGADOS RELAÇÃO N.9 - ADV: J.C, M.M., B.C., P.M. Processo - Procedimento Sumário - Acidente de Trânsito - Requerido: S.S. Autor: G.A. Réu: M.P. - Isso posto, defiro o pedido da ré S.S. e determino a reunião destes autos ao processo de n. 799-8.6.8, em trâmite nesta Vara Cível. Em consequência: I - Translade-se esta decisão para os autos mencionados. II - Publique-se, com prazo de dias para eventual manifestação das partes. III - Decorrido, dê-se vista dos autos ao Ministério Público pelo prazo de (quinze) dias, haja vista haver interesse de incapaz relativamente aos autos de n. 799-8.6.8. (CPC, art. 78, II). IV - Com o retorno, venham ambos os feitos conclusos para sentença”.

Exemplo 2 (Despacho):

“Sr. Advogado, Notificação - Processo N. ROT- Relator R. M. RECORRENTE O.B. S/A ADVOGADO B.T. RECORRENTE A.E. ADVOGADO H.J. ADVOGADO V.P. RECORRIDO G.C. ADVOGADO F.M. TERCEIRO INTERESSADO J.D. Intimado(s)/Citado(s): - A.E. Destinatário: A.E. Fica(m) intimado(s) o(s) agravado(s) para contraminutar(em) o(s) agravo(s) de instrumento e contrarrazoar(em) o(s) recurso(s) de revista conforme a Instrução Normativa n. 6 do TST. Em 9/10/2019. Secretaria de Apoio Jurisprudencial e Recursal”.

Exemplo 3 (Sentença):

“ADV: R.S.F. Processo - Procedimento do Juizado Especial Cível / Juizado Especial Cível - Autora : N. Z - Réu : U. F.- Isso posto, JULGO PARCIALMENTE PROCEDENTES os pedidos formulados na inicial da AÇÃO DE RESSARCIMENTO C/C INDENIZAÇÃO POR DANOS MORAIS ajuizada por N. Z. em desfavor de U.F. para, em consequência, DECLARAR nula de pleno direito a Cláusula , alínea do contrato de assistência à saúde perfectibilizado pelas partes no específico caso concreto impingido nos autos; e CONDENAR a demandada a arcar com o importe equivalente a R\$ 3.888,9 (três mil oitocentos e oitenta e oito reais e noventa e cinco centavos), com incidência de correção monetária a fluir a partir de 01/04/2020 e juros moratórios a contar da citação. Sem custas e honorários advocatícios. P. R. I. Oportunamente, ao arquivo”.

Como classes para realização da classificação, selecionou-se o tipo de publicação, seguindo uma definição formal contida no artigo 203 do Código de Processo Civil, que classifica os pronunciamentos do juiz entre decisões

interlocutórias, despachos e sentenças. Ressalvadas as disposições expressas dos procedimentos especiais, sentença é o pronunciamento por meio do qual o juiz põe fim à fase cognitiva do procedimento comum, bem como extingue a execução, com ou sem julgamento de mérito. Decisão interlocutória é todo pronunciamento judicial de natureza decisória que não se enquadre como sentença, resolvendo questão incidental. Despachos são todos os demais pronunciamentos do juiz praticados no processo, de ofício ou a requerimento da parte.

As 235 publicações selecionadas foram submetidas a uma classificação, realizada por uma profissional da área do direito, resultando em um Corpus com a seguinte distribuição de classes.

Tabela 1. Registros da base de dados

Classe	Registros	%
Decisão Interlocutória	32	13,62%
Despacho	137	58,30%
Sentença	66	28,09%
Total	235	100,00%

Tabela 1. Registros da base de dados

5 PRÉ-PROCESSAMENTO

No PLN, a etapa de pré-processamento tem a finalidade de criar um conjunto de palavras que terão alguma relevância para criação do modelo de classificação e é dividida em algumas tarefas sequenciais, conforme apresentado nos próximos itens. Esta técnica de separar e contar as palavras dos textos é conhecida como *bag of words* (sacola de palavras). Como exemplo, para ilustrar a transformação que o texto sofre após cada tarefa, será utilizada a frase “Edital de intimação de sentença da vítima Fulano. Prazo do EDITAL: 30 (trinta) dias”.

a) Tokenização: Transformação de caracteres para minúsculos, troca de caracteres acentuados pelo mesmo caractere sem a acentuação e retirada de caracteres que não sejam letras, além da exclusão dos números. Com a execução destas tarefas, o texto fica “edital de intimacao de sentenca da vitima fulano de tal prazo do edital trinta dias”.

b) Remoção de *stopwords*: Remoção de palavras irrelevantes para a classificação, como as preposições, por exemplo; remoção de nomes próprios

e remoção de palavras com menos de 3 caracteres. Com a execução destas tarefas, o texto fica “edito intimacao sentenca vitima prazo edital dias”.

c) Stemmização: Redução das palavras ao seu radical, mantendo apenas a parte mais significativa de cada palavra. Como exemplo, as palavras “intimacao”, “intimo” e “intimamos”, passa a se apresentar apenas como “intim”. Com a execução destas tarefas, o texto fica “edito intim sentenc vitim praz edit dia”.

6 TREINAMENTO

A etapa de treinamento consiste na aplicação do algoritmo escolhido para criação dos modelos de aprendizado de máquina, para os quais foi possível a aplicação dos classificadores e análise dos resultados, conforme já ilustrado na figura 1. Neste trabalho foi utilizado o algoritmo Naive Bayes, o qual já tinha apresentado resultados melhores em trabalhos anteriores dos autores referenciados.

6.1 Divisão da base de dados: Treinamento e Teste

Para treinamento e teste dos modelos foi utilizada a técnica de *holdout*, separando aleatoriamente 2/3 da base principal para treinamento e 1/3 para testes. O processo foi repetido por três vezes, alternando os registros de forma que a base de testes fosse completamente distinta a cada simulação. As quantidades de registros para cada simulação são apresentadas na tabela 2.

Tabela 2. Registros da base de dados

Classe	Simulação 1		Simulação 2		Simulação 3	
	Base de Treino	Base de Teste	Base de Treino	Base de Teste	Base de Treino	Base de Teste
Decisão Interlocutória	18 11,46	14 17,95%	24 15,29%	8 10,26%	22 14,01%	10 12,82%
Despacho	96 61,15%	41 52,56%	86 54,78%	51 65,38%	92 58,60%	45 57,69%
Sentença	43 27,39%	23 29,49%	47 29,94%	19 24,3	43 27,39%	23 29,49%
Total	157 100%	78 100%	157 100%	78 100%	157 100%	78 100%

6.2 Algoritmo Naive Bayes

Naive Bayes é um classificador probabilístico baseado na Regra de Bayes, que pressupõe a independência entre os dados, sendo essa a sua principal vantagem, pois consegue obter bons resultados com poucos dados. Conforme

Raschka (2014), a fórmula que rege a Regra de Bayes é:

$$P(\text{classe} | \text{características}) = \frac{P(\text{classe}) \times P(\text{características} | \text{classe})}{P(\text{características})}$$

Onde lê-se que a probabilidade de um texto ser de uma determinada classe dado um conjunto de palavras é igual ao produto da probabilidade geral da classe pela probabilidade de ocorrer o mesmo conjunto de palavras dada a classe, dividido pela probabilidade de ocorrer o mesmo conjunto de palavras.

6.3 Aplicação dos algoritmos: treinamento dos modelos e classificação

Para verificar o resultado dos algoritmos foram utilizadas as métricas de acurácia, precisão, recall e F β -Score. A acurácia é a medida geral de desempenho do algoritmo, sendo a divisão dos acertos pela quantidade total de registros classificados. Precisão é a taxa de textos corretamente classificados como positivos de cada classe. Recall é a relação entre a quantidade de exemplos classificados corretamente como positivos e o número total de exemplos de cada classe. O F β -Score é a média harmônica entre precisão e recall. O resultado apresentado é a média ponderada dos resultados, considerando o peso de cada classe.

Tabela 3. Resultados obtidos com Naive Bayes

Simulação 1				Simulação 2				Simulação 3			
Matriz confusão				Matriz confusão				Matriz confusão			
	DI	DE	SE		DI	DE	SE		DI	DE	SE
DI	5	5	4	DI	.	7	1	DI	.	6	4
DE	4	33	4	DE	1	48	2	DE	.	45	.
SE	.	11	12	SE	1	11	7	SE	.	9	14
Acurácia: 64,10% Precisão: 63,06% Recall: 64,10% F β -Score: 62,81%				Acurácia: 70,51% Precisão: 64,60% Recall: 70,51% F β -Score: 65,41%				Acurácia: 75,64% Precisão: 66,20% Recall: 75,64% F β -Score: 69,59%			

Após treinamento e classificação do algoritmo Naive Bayes foram obtidos os resultados apresentados na tabela 3. As classes estão representadas por siglas, para melhorar a forma de visualização, sendo Decisão Interlocutória = DI, Despacho = DE e Sentença = SE.

6.4 Ajuste na seleção das palavras (aplicação de critérios de seleção)

Após a execução das simulações, foi levantada a hipótese de que o desequilíbrio entre a quantidade de registros para cada classe pudesse estar influenciando negativamente os resultados. No intuito de equilibrar a quantidade de características, visando melhorar os resultados obtidos, os modelos foram gerados novamente, agora com novos critérios de seleção das características para geração dos modelos. Em 2005, Doam e Horiguchi, haviam constatado em seu artigo que a aplicação de um ranking, considerando multicritérios de seleção, podem reduzir o número de termos e melhorar a performance do classificador.

Além dos tratamentos que já haviam sido executados na etapa de pré-processamento, foi levando em consideração a representatividade de cada palavra no total de palavras e a frequência da palavra em cada classe. O critério estipulado de forma arbitrária para o experimento foi de que para fazer parte das características do modelo a palavra deveria constar apenas em uma das classes. O resultado foi uma redução da matriz de palavras, conforme apresentado na tabela 4.

Tabela 4. Redução da matriz de palavras

Classe	Palavras antes da aplicação do critério de seleção	Palavras após da aplicação do critério de seleção
Decisão Interlocutória	396	4
Despacho	1198	6
Sentença	1297	103
Total	2891	113

Considerando a nova matriz de termos, os modelos foram gerados novamente, gerando os resultados apresentados na tabela 5.

Tabela 5. Resultados obtidos com Naive Bayes após aplicação de critérios de seleção

Simulação 1				Simulação 2				Simulação 3			
Matriz confusão				Matriz confusão				Matriz confusão			
	DI	DE	SE		DI	DE	SE		DI	DE	SE
DI	8	6	.	DI	4	4	.	DI	5	5	.
DE	.	40	1	DE	1	49	1	DE	.	45	.
SE	.	5	18	SE	.	7	12	SE	.	6	17
Acurácia: 84,62% Precisão: 87,11% Recall: 84,62% F β -Score: 84,04%				Acurácia: 83,33% Precisão: 84,09% Recall: 83,33% F β -Score: 82,31%				Acurácia: 85,90% Precisão: 88,67% Recall: 85,90% F β -Score: 85,02%			

7 ANÁLISE DOS RESULTADOS

A análise dos resultados foi feita com base em dois principais eixos: um comparativo entre os algoritmos e um comparativo antes e depois da aplicação do critério de seleção de características para geração dos modelos. Os resultados apresentados na tabela 6 são as médias aritméticas entre as iterações de cada caso.

Tabela 6. Resultados gerais

	Resultado antes da aplicação do critério de seleção	Resultados após a aplicação do critério de seleção
Acurácia	70,09%	84,62%
Precisão	64,62%	86,62%
Recall	70,09%	84,62%
Fβ-Score	65,94%	83,79%

Na comparação entre a aplicação do algoritmo, pôde-se registrar uma melhora significativa nos resultados, chegando o Naive Bayes a 84,62% de acurácia, ou seja, um resultado 20,73% melhor que o resultado obtido sem a aplicação dos critérios.

8 CONCLUSÃO

Com estes experimentos, concluímos que é possível criar classificadores com acurácia acima de 80% para pronunciamentos judiciais, identificando de forma automática, por meio do PLN, decisões interlocutórias, sentenças e despachos, tendo o algoritmo Naive Bayes demonstrado uma melhora significativa após a aplicação de critérios de seleção de características. Também é possível afirmar que o desequilíbrio entre a quantidade de classes na base de treinamento pode implicar em um viés nos resultados, sendo a aplicação de critérios para seleção de palavras, conforme demonstrado neste trabalho, uma alternativa para melhorar a acurácia dos algoritmos.

Uma hipótese com relação a essa melhora nos resultados após a aplicação de critérios de seleção, é o desequilíbrio entre a quantidade de registros de cada classe no corpus original, que é eliminada ao realizar-se a seleção. O teste desta hipótese pode ser objeto de estudo em trabalhos futuros.

REFERÊNCIAS

- ADEWOYIN, Oluwande; PADAYACHY, Thashen; SCHOLTZ, Brenda. Pilot Testing of an Information Extraction (IE) Prototype for Legal Research. *The African Journal of Information and Communication* 2020.25. 2020.
- ALETRAS, Nikolaos; TSARAPATSANIS, Dimitrios; PREOȚIUC-PIETRO, Daniel; LAMPOS, Vasileios. *Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective*. PeerJ. Computer Science 2. 2016.
- ALVES, Rodrigo Lucas da Silva Pereira da Gama. *O conteúdo do pronunciamento judicial como elemento de harmonia do sistema recursal brasileiro*. 2017. 154 f. Dissertação (Mestrado em Direito) - Programa de Estudos Pós-Graduados em Direito, Pontifícia Universidade Católica de São Paulo, São Paulo, 2017.
- ANAND, Deepa; WAGH, Rupali; Effective Deep Learning Approaches for Summarization of Legal Texts. *Journal of King Saud University. Computer and Information Sciences*. 2019.
- ARAÚJO, Erik da Silva; SIMIONI, Rafael Larazzotto. A inteligência artificial e ponderação: o procedimentalismo de robert alexy na forma de algoritmo. In: *Arquivo Jurídico - Revista Jurídica Eletrônica da Universidade Federal do Piauí*, Volume 8, N1. 2021.
- DOAN, Son; Horiguchi, Susumu. An efficient feature selection using multi-criteria in text categorization for Naïve Bayes classifier”. In: *Fourth International Conference on Hybrid Intelligent Systems*. 2005.
- EL JELALI, Soufiane; FERSINI, Elisabetta; MESSINA, Enza. Legal Retrieval as Support to EMediation: Matching Disputant’s Case and Court Decisions. *Artificial Intelligence and Law* 23.1. 2015.
- FERNANDES, William Paulo Ducca; SILVA, Luiz José Schirmer; FRAJHOF, Isabella Zalberg; DE ALMEIDA, Guilherme Da Franca Couto Fernandes; KONDER, Carlos Nelson; NASSER, Rafael Barbosa; DE CARVALHO, Gustavo Robichez; BARBOSA, Simone Diniz Junqueira; LOPES, Hélio Côrtes Vieira. Appellate Court Modifications Extraction for Portuguese. In: *Artificial Intelligence and Law*. 2019.
- GALGANI, Filippo; COMPTON, Paul; HOFFMAN, Achim. LEXA: Building Knowledge Bases for Automatic Legal Citation Classification. *Expert Systems with Applications* 42.17-18. 2015.
- HARTMANN PEIXOTO, Fabiano. *Direito e Inteligência Artificial*. Coleção Inteligência Artificial e Jurisdição. Volume 2. DR.IA. Brasília, 2020.
- HUANG, Zihan; LOW, Charles; TENG, Mengqiu, ZHANG, Zhang; HO, Daniel E.; KRASS, Mark S.; GRABMAIR, Matthias. *Context-Aware Legal Citation Recommendation Using Deep Learning*. Association for Computing Machinery. 2021.

- JI, Donghong; GAO, Jun; FEI, Hao; TENG, Chong; REN, Yafeng. *A Deep Neural Network Model for Speakers Coreference Resolution in Legal Texts*. Information Processing & Management. 2020.
- JI, Donghong; TAO, Peng; FEI, Hao; REN, Yafeng. *An End-to-end Joint Model for Evidence Information Extraction from Court Record Document*. Information Processing & Management. 2020.
- LAWLOR, Reed C. What Computers Can Do: Analysis and Prediction of Judicial Decisions. *American Bar Association Journal*, vol. 49, n. 4, april 1963, p. 337-344 (<https://www.jstor.org/stable/25722338>)
- KATZ, Daniel Martin; BOMMARITO 2nd, Michael J.; BLACKMAN, Josh. *A General Approach for Predicting the Behavior of the Supreme Court of the United States*. PloS One. 2017.
- LI, Jiamin; LIU, Xingbo; NIE, Xiushan; MA, Lele; LI, Peng; ZHANG, Kai; YIN, Yilong. Weighted-Attribute Triplet Hashing for Large-Scale Similar Judicial Case Matching. *Computational Intelligence and Neuroscience*. 2021.
- MCROY, Susan. *Principles of Natural Language Processing*. 2021.
- MEDVEDEVA, Masha; VOLS, Michel; WIELING, Martijn. sing Machine Learning to Predict Decisions of the European Court of Human Rights. *Artificial Intelligence and Law* 28.2. 2019.
- MORAIS DA ROSA, Alexandre. A questão digital: o impacto da inteligência artificial no Direito. *Revista de Direito da Faculdade Guanambi*, Guanambi, v. 6, n. 2, p. 1- 18, jul./dez. 2019.
- NUNES, Dierle; MARQUES, Ana Luiza Pinto Coelho. Inteligência artificial e direito Processual: vieses algorítmicos e os riscos de atribuição de função decisória às máquinas. *Revista de Processo*, Distrito Federal, v. 285, n. 43, p. 1-19, nov. 2018.
- PETROVIC, Djordje; STANKOVIC, Milena. Use of Linguistic Forms Mining in the Link Analysis of Legal Documents. *Computer Science and Information Systems* 15.2. 2018.
- POLO, Felipe Maia; CIOCHETTI, Itamar; BERTOLO, Emerson. Predicting legal proceedings status: approaches based on sequential text data. In: *Anais da the Eighteenth International Conference on Artificial Intelligence and Law*. Association for Computing Machinery, New York, NY, USA, 264–265. 2021.
- RASCHKA, Sebastian. Naive Bayes and Text Classification I - Introduction and Theory. *ArXiv abs/1410.5329*. 2014.
- SHULAYEVA, Olga; SIDDHARTHAN, Advait; WYNER, Adam. Recognizing Cited Facts and Principles in *Legal Judgements*. *Artificial Intelligence and Law* 25.1. 2017.
- TRIAS, Fernando. *Temporal Social Network Analysis Using Harvard Caselaw Access Project*. 2021.
- WAGH, Rupali; ANAND, Deepa; Legal Document Similarity: A Multi-criteria Decision-making Perspective. *PeerJ. Computer Science* 6. 2020.

XU, Jianfeng; SUN, Fuhui; CHEN, Qiwei; WANG, Yingfei; YU, Jia; LIU, Zhenyu. Uma nova estrutura de padrões e práticas inovadoras no projeto de engenharia do sistema de sistemas de quadra inteligente da China. *Estudo Estratégico da Academia Chinesa de Engenharia*, 2022, 24(4): 105\u2012120
<https://doi.org/10.15302/J-SSCAE-2022.04.005>.

