Gomes, Lucas Moreira; Jader Martins Camboim de Sá; Peng, Yaohao

Working Paper

Línguas naturais e máquinas artificiais: Aplicação de técnicas de mineração de texto para a classificação de sentenças judiciais brasileiras

Texto para Discussão, No. 2612

Provided in Cooperation with:
Institute of Applied Economic Research (ipea), Brasília

Suggested Citation: Gomes, Lucas Moreira; Jader Martins Camboim de Sá; Peng, Yaohao (2020) : Línguas naturais e máquinas artificiais: Aplicação de técnicas de mineração de texto para a classificação de sentenças judiciais brasileiras, Texto para Discussão, No. 2612, Instituto de Pesquisa Econômica Aplicada (IPEA), Brasília, https://doi.org/10.38116/td2612

This Version is available at:
https://hdl.handle.net/10419/240806

TEXTO PARA DISCUSSÃO

2612

LÍNGUAS NATURAIS E MÁQUINAS ARTIFICIAIS: APLICAÇÃO DE TÉCNICAS DE MINERAÇÃO DE TEXTO PARA A CLASSIFICAÇÃO DE SENTENÇAS JUDICIAIS BRASILEIRAS

Lucas Moreira Gomes
Jader Martins Camboim de Sá
Peng Yaohao

ipea

2612
TEXTO PARA DISCUSSÃO
Brasília, novembro de 2020

LÍNGUAS NATURAIS E MÁQUINAS ARTIFICIAIS: APLICAÇÃO DE TÉCNICAS DE MINERAÇÃO DE TEXTO PARA A CLASSIFICAÇÃO DE SENTENÇAS JUDICIAIS BRASILEIRAS

Lucas Moreira Gomes[1]
Jader Martins Camboim de Sá[2]
Peng Yaohao[3]

1. Auxiliar de Pesquisa do Subprograma Nacional de Pesquisa para o Desenvolvimento (PNPD) na Assessoria Técnica (Astec) do Ipea. *E-mail*: <lucasmoreira@lamfo.unb.br>.
2. Auxiliar de Pesquisa do PNPD na Astec/Ipea. *E-mail*: <jader.martins@ipea.gov.br>.
3. Pesquisador Visitante na Astec/Ipea e Assessor Especial na Secretaria de Política Econômica (SPE) da Secretaria Especial da Fazenda do Ministério da Economia. *E-mail*: <peng.yaohao@fazenda.gov.br>.

Governo Federal

Ministério da Economia
Ministro Paulo Guedes

ipea **Instituto de Pesquisa Econômica Aplicada**

Fundação pública vinculada ao Ministério da Economia, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiros – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

Presidente
Carlos von Doellinger

Diretor de Desenvolvimento Institucional
Manoel Rodrigues Junior

Diretora de Estudos e Políticas do Estado, das Instituições e da Democracia
Flávia de Holanda Schmidt

Diretor de Estudos e Políticas Macroeconômicas
José Ronaldo de Castro Souza Júnior

Diretor de Estudos e Políticas Regionais, Urbanas e Ambientais
Nilo Luiz Saccaro Júnior

Diretor de Estudos e Políticas Setoriais de Inovação e Infraestrutura
André Tortato Rauen

Diretora de Estudos e Políticas Sociais
Lenita Maria Turchi

Diretor de Estudos e Relações Econômicas e Políticas Internacionais
Ivan Tiago Machado Oliveira

Assessor-chefe de Imprensa e Comunicação (substituto)
João Cláudio Garcia Rodrigues Lima

Ouvidoria: http://www.ipea.gov.br/ouvidoria
URL: http://www.ipea.gov.br

Texto para Discussão

Publicação seriada que divulga resultados de estudos e pesquisas em desenvolvimento pelo Ipea com o objetivo de fomentar o debate e oferecer subsídios à formulação e avaliação de políticas públicas.

Texto para discussão / Instituto de Pesquisa Econômica Aplicada.- Brasília : Rio de Janeiro : Ipea , 1990-

ISSN 1415-4765

1.Brasil. 2.Aspectos Econômicos. 3.Aspectos Sociais. I. Instituto de Pesquisa Econômica Aplicada.

CDD 330.908

JEL: C55, C44, O33, C63, K40, C81.

DOI: http://dx.doi.org/10.38116/td2612

SUMÁRIO

SINOPSE

ABSTRACT

1 INTRODUÇÃO 7

2 JURIMETRIA.....9

3 *BIG DATA* E MINERAÇÃO DE TEXTO PARA POLÍTICAS PÚBLICAS 13

4 ANÁLISE DE TEXTOS JURÍDICOS 14

5 METODOLOGIA.....16

6 RESULTADOS.....22

7 CONCLUSÃO 34

REFERÊNCIAS 36

SINOPSE

Este trabalho investigou o uso de técnicas de inteligência artificial e de mineração de texto para a classificação de sentenças judiciais quanto à procedência do pedido do autor da ação, bem como discutiu potenciais aplicações alternativas no âmbito da formulação e avaliação de políticas públicas. Ademais, o trabalho construiu um levantamento de estudos relativos à jurimetria provenientes da literatura científica especializada e detalhou a operacionalização do tratamento de dados textuais, bem como definindo conceitos e métodos básicos de mineração de texto. Por fim, uma análise empírica de classificação de textos jurídicos em quatro categorias foi executada utilizando-se dados do Tribunal Regional Federal da 2ª região coletados pelo IpeaJus, banco de dados do Ipea sobre o sistema de justiça do Brasil, com os resultados sendo discutidos à luz de diversas métricas quantitativas de avaliação e prospectos de desenvolvimentos futuros em contextos diversos.

Palavras-chave: processamento em linguagem natural; classificação de documentos; procedência jurídica; jurimetria; *Big Data*.

ABSTRACT

This paper investigated the usage of artificial intelligence and text mining techniques for classification of court judgments and discussed potential alternative applications in formulation and evaluation of public policies. Besides, we built a survey of studies related to Jurimetry based on the specialized scientific literature and detailed the operationalization of the of textual data treatment, as well as basic concepts and methods of text mining. Finally, we performed an empirical analysis of classification of legal texts into four categories using real data from the Brazilian 2nd Federal Regional Court collected by IpeaJus, the database about the Brazilian Justice System from Ipea, discussing the results in light of various quantitative evaluation metrics and prospects for future developments in different contexts.

Keywords: natural language processing; document classification; legal proceedings; jurimetry; Big Data.

1 INTRODUÇÃO

Muito se discute acerca da abundância de dados e da necessidade do desenvolvimento de tecnologias e estruturas de gerenciamento para lidar com as fontes de informações – estruturadas ou não –, para se "minerar" padrões relevantes presentes no *Big Data* e para subsidiar a tomada de decisões estratégicas. Conforme discutido em *Digital workplaces: vision and reality* (White, 2012), mais que meramente uma questão de volume – bases de dados com mais observações e com maior periodicidade –, *Big Data* implica potencial de se gerar valor com as informações disponíveis, bem como em uma necessidade de adaptar os métodos de análise de acordo com a natureza do dado. Dessa forma, o desafio do relativamente recente campo de estudos "ciência de dados" envolve uma interseção entre a tecnicidade – seja matemática, seja computacional – de se lidar com a crescente abundância de dados e a sensibilidade em identificar em que sentido a massa de dados pode ajudar a melhor resolver um problema levantado, para de fato se "minerar" o valor latente de textos e tabelas.

Tal qual discutido em estudos como *The future of employment: How susceptible are jobs to computerisation?* (Frey e Osborne, 2017), *El futuro del trabajo en América Latina y el Caribe* (Bosch e Ripani, 2018) e *Na era das máquinas, o emprego é de quem?* (Albuquerque *et al.*, 2019), a emergência da "era do *Big Data*" coincide com o desenvolvimento de tecnologias que permitem a automação de tarefas tipicamente rotineiras, pouco intensivas em cognição, e de baixa exigência de capacitação técnica. Por outro lado, ocupações que demandam habilidades criativas e de inovação ainda possuem baixa propensão de automação. Nesse sentido, a automação de tarefas pode contribuir para o aumento da eficiência de trabalhos rotineiros que servirão de insumos para a tomada de decisões, deslocando a energia de gestores e de acadêmicos para problemas mais desafiadores.

Nesse âmbito, este trabalho discute a aplicação da automação especificamente em relação à classificação de sentenças jurídicas em categorias distintas de procedência do pedido realizado pelo autor no texto, desenvolvendo um sistema computacional especializado em "ler" os textos e classificá-los com base em textos passados cujo conteúdo foi inspecionado, total ou parcialmente, por humanos. Baseado em técnicas de mineração de texto e processamento em linguagem natural (*natural language processing*, doravante NLP), o sistema proposto neste trabalho constitui-se em uma

estrutura de análise de dados que pode potencialmente ser refinado metodologicamente e estendido para aplicações mais generalistas, fornecendo índices e indicadores que podem ser utilizados posteriormente para, por exemplo, mensurar a eficiência jurídica no Brasil, o que por sua vez pode auxiliar na formulação de políticas públicas direcionadas para o aprimoramento do aparato judicial brasileiro. Foi conduzida uma revisão da literatura científica especializada no tocante à aplicação de técnicas matemáticas e computacionais para análise de textos jurídicos e a estudos recentes para o cenário brasileiro.

Ademais, este texto descreve procedimentos básicos de tratamento de dados textuais, discutindo com detalhes a transposição de textos, compostos por palavras e caracteres, para uma estrutura tipicamente numérica, bem como conceitos básicos de importância relativa de termos ou combinações de termos, métodos conhecidos de classificação textual e principais métricas de avaliação da qualidade de classificação. Com isso, este trabalho almeja fornecer um guia de diretrizes para aplicações alternativas de mineração de texto, notadamente para fins de avaliação de políticas públicas. Um estudo de caso foi realizado utilizando dados de sentenças judiciais coletadas do Tribunal Regional Federal da 2ª região (TRF2), datadas entre 2015 e 2017, classificando-as em cinco categorias: *procedentes*, *improcedentes*, *parcialmente procedentes*, *acordos* e outras classificações.

Este trabalho está estruturado da seguinte maneira: seguidas desta Introdução, a seção 2 apresenta trabalhos clássicos em jurimetria, enquanto que a seção 3 trata sobre aplicações de técnicas de mineração de texto no planejamento de políticas públicas. Logo depois, a seção 4 apresenta estudos específicos de textos jurídicos, seguida pela seção 5, que detalha as metodologias empregadas neste estudo, desde a conversão de textos para estruturas numéricas, até métricas de frequência e relevância de n-gramas e a especificação matemática do classificador. A seção 6 apresenta os resultados dos experimentos empíricos, com os resultados das classificações fora da amostra e os verbetes mais relevantes para cada classe. Por fim, a seção 7 apresenta as conclusões do trabalho, bem como limitações e recomendações futuras de pesquisa e aplicações práticas no âmbito de políticas públicas e avaliação do sistema judicial.

2 JURIMETRIA

Entende-se como jurimetria o uso de métricas estatísticas e técnicas de análise sistemática da área jurídica, conceito inicialmente definido em 1948 (Loevinger, 1948). À época, já se tinha a noção de que a interpretação das informações disponíveis seria de suma importância para o avanço da eficiência jurídica, pois não se poderia respeitar aquilo que não se conhece ou não se vê o funcionamento de fato. No entanto, não se imaginava que os computadores poderiam realizar esse papel com a profundidade que hoje o fazem.

> There will always be those who will scoff at the idea that law can be put on a rational basis as visionary. Let us admit that this proposal is improbable. The most that can be said is that it is just as improbable as the possibility that the weak and stupid creature we call man will survive much longer on an insignificant speck of dust whirling madly about in the finite but unbounded reaches of a vast and expanding universe. It is just exactly as improbable as that, for it is the indispensable condition of such survival (Loevinger, 1948, p. 493).

Com o avanço da computação e tecnologias auxiliares, muito progresso foi feito na área da jurimetria desde sua concepção. Para alguns autores como Haddad (2010) e Milena e Serra (2013), a jurimetria é vista como uma ferramenta de auxílio aos magistrados, mas não como uma ferramenta capaz de traduzir a verdade absoluta. Já outros autores dão um passo à frente ao tentar não só levantar indicadores que possam auxiliar a interpretação jurídica, mas também fazer uma análise da estrutura textual gramatical. Em *Linguistic feature analysis on judicial decisions based on keyword extraction and high-frequency word statistics – taking paper of sentence for example* (Yuan *et al*., 2019), por exemplo, são analisadas sentenças da suprema corte (chinesa, no caso), e identificados fatores determinantes que precisariam estar mais claros no texto para possibilitar análises mais precisas, considerando as especificidades do idioma.

A jurimetria pode ser analisada sob diferentes óticas. Alguns estudos procuram compreender o impacto que determinadas informações podem ter sobre o resultado de uma sentença. Em *Mapping the science of law: a jurimetrics analysis* (Nishavathi e Jeyshankar, 2018) é feita uma análise dos precedentes usados na argumentação jurídica para classificar a influência que cada um deles teria sob a decisão do juiz, criando então um mapa de calor capaz de identificar o grau de importância dos precedentes. Análises como essas são relevantes porque uma boa argumentação e fundamentação adequada são tarefas que requerem custo e conhecimento específico elevados, e impactam substancialmente a decisão final do juiz.

Se por um lado a jurimetria possibilita análises automatizadas em grande escala, o próprio processo de pesquisa e desenvolvimento dessa área pode significar um dos maiores desafios. Por se tratar de uma área multidisciplinar, faz-se necessário que uma equipe de análise envolva, pelo menos, conhecimentos multidisciplinares que lidem com o processamento computacional e com análises estatísticas, além de conhecimentos necessários da área jurídica para garantir a significância dos estudos elaborados (Zabala e Silveira, 2014). Esse tipo de abordagem faz com que pesquisas dessa natureza sejam de difícil execução, já que além da dificuldade natural de se construir uma equipe com essas qualidades, pode haver também perda de interpretação entre as áreas de conhecimento.

No Brasil, a jurimetria conta com alguns estudos que empregam técnicas de análise de dados e mineração de texto, similares ao modelo proposto neste estudo. Na tese *Mensurando a eficiência no sistema judiciário: métodos paramétricos e não-paramétricos*, é destacado o potencial da jurimetria como instrumento para se alcançar ganhos significativos de eficiência no gasto público, reduzindo custos e aumentando a oferta e a qualidade de serviços públicos (Schwengber, 2006).

O Ipea, em especial, tem se dedicado ao estudo do judiciário brasileiro sob diversos pontos de vista e com o emprego de diferentes métodos, quantitativos e qualitativos. Considerando as últimas duas décadas, observamos que diferentes matérias judiciais e entes da organização do judiciário, bem como distintos aspectos do processo judicial foram objetos de estudo pelo instituto.

Ilustrativamente, os processos judiciais estudados pelo instituto foram delimitados aos assuntos de execução fiscal (Cunha *et al.*, 2011); execuções judiciais trabalhistas (Campos e Di Benedetto, 2015); judicialização de pedidos de benefício de prestação continuada e aposentadoria rural, (Castro e Jesus, 2018; Jesus *et al.*, 2018); execução de títulos extrajudiciais (Castro, 2015; Castro, Romeiro e Cavalcanti, 2019); monitórias, busca e apreensão em alienação fiduciária, despejo por falta de pagamento, pedidos de recuperação judicial e ações de alimentos (Castro, Romeiro e Cavalcanti, 2019).

Quanto à organização, foram abordados juizados especiais federais da 1ª e 2ª regiões (1ª instância da Justiça Federal nessas regiões); TRFs da 1ª e 2ª regiões (2ª instância da Justiça Federal nessas regiões) (Castro e Jesus, 2018; Jesus *et al.*, 2018); a justiça

estadual de primeiro grau (Castro e Coelho, 2011); serviços locais de notários e registro (Castro e Coelho, 2011; Castro, 2014); Justiça do Trabalho (Campos e Di Benedetto, 2015; Campos, 2018); e a reforma da justiça de 2004 (Pinheiro, 2003; Pinheiro, 2005; Castro e Cunha, 2014).

Os estudos do Ipea também examinaram, dentre outros aspectos dos processos judiciais, o tempo de tramitação e o custo unitário de processos de execução fiscal (Cunha *et al.*, 2011; Cunha, Klin e Gomes, 2011; Cunha, Medeiros e Silva, 2012); o acesso à justiça especial federal (Aquino e Colares, 2013); o passivo acumulado de execuções trabalhistas (Campos e Di Benedetto, 2015); o desempenho e a previsibilidade da justiça brasileira em geral (Pinheiro, 2003), bem como sua gestão, produtividade e qualidade (Pinheiro, 2005; Castro, 2012), eficiência, independência, acesso, desenho institucional e responsabilização judiciais (Castro, 2012; Castro e Cunha, 2014); a eficiência produtiva da justiça estadual de primeiro grau e seus indicadores de eficiência e celeridade (Castro e Coelho, 2011); o impacto da justiça trabalhista na economia e na sociedade (Campos, 2017); a efetividade resolutiva da justiça trabalhista (Campos, 2018); a territorialização dos dados sobre o sistema de justiça (Silva, 2013); e os papéis alternativos do sistema judicial para melhor resolução de conflitos laborais (Campos, 2018). Adicionalmente, a pesquisa empírica em direito no Brasil, envolvendo suas reflexões e métodos, foi reunida e discutida pelo Ipea (Cunha e Silva, 2013).

Apesar dessa diversidade e volume de pesquisa, o Ipea ainda não havia explorado o uso de técnicas de inteligência artificial, aprendizagem de máquina e processamento de linguagem natural em análises sobre processos judiciais, como feito neste estudo.

Embora a jurimetria seja uma área ainda pouco explorada (Menezes e Barros, 2018), o Brasil passa por um intenso processo de modernização no setor público (Cavalcante e Camões, 2017), sendo esse um importante fator na viabilização de estudos dessa área em território nacional.

Ainda no cenário nacional, alguns autores apontam especificidades culturais e tecnológicas que impedem o avanço rápido da jurimetria. Fatores como o operador de direito não estar acostumado a conceber análises estatísticas no processo de decisão ou, até mesmo, o receio da perda do protagonismo do magistrado com o desenvolvimento de técnicas de tais naturezas podem estar freando esse avanço (Menezes e Barros, 2018).

Enquanto isso, fatores como a existência de diferentes padrões de disponibilização da informação e, até mesmo, bloqueios de acesso em massa como o Captcha[1] (Completely Automated Public Turing test to tell Computers and Humans Apart) poderiam dificultar a criação de bancos de dados para análises robustas dos processos judiciais (Armonas Colombo, Buck e Miana Bezerra, 2017).

Por outro lado, o judiciário brasileiro vem evoluindo na disponibilização de dados e no engajamento em pesquisas. O Conselho Nacional de Justiça (CNJ) e o Conselho da Justiça Federal (CJF) cooperaram com o Ipea em estudos sobre duração e custos de processos (Cunha *et al.*, 2011) e acesso à justiça (Aquino e Colares, 2013), a título de exemplo.

A base de dados Justiça Aberta, organizada pelo CNJ, contém relatórios de produção de 8.495 serventias judiciais estaduais de primeira instância, incluindo o número de juízes e de funcionários, o estoque de processos pendentes de julgamento e os fluxos de processos distribuídos e resolvidos (sentenças e homologações de acordo), além do volume de despachos e decisões interlocutórias (Castro e Coelho, 2011).[2] Já a base de microdados do Banco Nacional de Autos Findos de Ações Trabalhistas (BNAFT) contém informações sobre a composição e os valores de execuções trabalhistas (Campos, 2017; Ipea, 2020).

Por sua vez, nos últimos anos, o Ipea vem consolidando, gradativamente, um banco de dados sobre processos judiciais e a estrutura do sistema de justiça no Brasil, denominado IpeaJus, extraindo dados de sistemas de acompanhamento processual – quando disponibilizados por órgãos da justiça – e da raspagem de textos de Diários Oficiais da Justiça. Atualmente, o IpeaJus contém cerca de 30 milhões de registros relativos a processos da esfera federal (1ª e 2ª regiões) e da esfera estadual (Tribunal de Justiça de São Paulo – TJSP) (Castro e Jesus, 2018; Jesus *et al.*, 2018). Além disso, a base de dados já foi utilizada na análise de processos de judicialização de pedidos de benefício de prestação continuada e aposentadoria rural, falências empresariais, relações trabalhistas, improbidade administrativa e judicialização de políticas públicas (Castro e Jesus, 2018; Jesus *et al.*, 2018).

1. Teste de Turing público completamente automatizado para diferenciação entre computadores e humanos.
2. Informações referentes ao ano de 2008.

A virtualização e automação de diversos processos geram uma grande quantidade de dados – o *Big Data* –, o que tem propiciado grandes revoluções em diferentes áreas da indústria. Esse recurso, quando aplicado a processos judiciais, pode promover melhores análises jurimétricas e outros benefícios além dos citados neste trabalho, como já se fez em campos como saúde e administração financeira (Raja *et al.*, 2008; Kumar e Ravi, 2016).

3 *BIG DATA* E MINERAÇÃO DE TEXTO PARA POLÍTICAS PÚBLICAS

O termo *Big Data*, no contexto de tecnologia da informação, caracteriza-se pela grande geração de dados, estruturados ou não, originados da digitalização em larga escala de serviços e transações, a chamada "revolução digital". Por sua vez, a captura, o processamento e o armazenamento desse grande volume de dados passou a ocorrer em velocidade superior àquela de produção dos dados. Assim, o *Big Data* representou grande avanço tecnológico, descrito pela expressão "3 Vs": volume, velocidade e variedade (Mukherjee e Shaw, 2016).

Para o processamento de linguagem natural, esse avanço tecnológico possibilitou uma melhoria significativa das análises, uma vez que esse ramo envolve elevada complexidade matemática. Em meados de 2009, com a recente digitalização dos processos judiciais no Brasil, tornou-se possível realizar análises de *Big Data* nessa área para todo o país. Nesse sentido, tendo em vista as tendências de modernização e integração de tecnologias capazes de lidar com um grande volume de dados estruturados ou não-estruturados, há também amplo espaço para a aplicabilidade prática de técnicas de análise de dados em larga escala no contexto da administração pública. A literatura especializada já aponta que áreas de planejamento, saúde pública, trânsito, segurança pública e distribuição de alimentos são as que possuem maior potencial para se beneficiarem do uso de técnicas dessa área (Demarzo, 2018).

Na literatura, são apontados diversos exemplos de aplicação de análise de texto em grandes volumes de dados referentes a diversas áreas da administração pública. Na saúde brasileira, por exemplo, a fim de otimizar os recursos de um hospital, foram analisados os dados gerados no primeiro encontro entre o médico e paciente no setor

de emergência, prevendo com 77% de precisão a demanda por internação (Lucini *et al.*, 2017). Em outro estudo, também na área da saúde brasileira, foram analisados os preços de compra de medicamentos disponíveis no portal da transparência e, posteriormente, classificados por agrupamento em categorias de preços máximos, médios e mínimos para um mesmo medicamento (Correa e Leal, 2018). O objetivo foi identificar superfaturamento nas compras desses medicamentos com NLP.

Já no cenário político, pesquisas relacionadas ao grau de satisfação da população com o governo se mostram como fator interessante para o entendimento e antecipação das necessidades da população. Em relação às manifestações de 2013, as quais resultaram no *impeachment* da presidente Dilma Rousseff, foram analisadas 130 mil mensagens de redes sociais em período anterior e posterior a essas manifestações. Para os autores da pesquisa, seria possível ter previsto a então polarização da população à época apenas com esse tipo de análise (Oliveira e Bermejo, 2017).

Na área da educação, é possível identificar na literatura recente que as pesquisas quantitativas se fazem majoritariamente presentes no ensino superior. Esse cenário é de se esperar, já que a oferta de dados sobre o ensino se dá em boa parte via plataformas de educação a distância. Assim, é possível afirmar que ainda existe um potencial latente na análise de dados dessa natureza na educação básica (Maschio *et al.*, 2018).

Mesmo nos casos em que não existem bases de dados consolidadas, ainda é possível gerar análises em grande escala. Na região da bacia hidrográfica do Pantanal, por exemplo, foi desenvolvido um banco de dados com notícias jornalísticas disponíveis na internet. Esses dados foram, à época, analisados para prever a safra agrícola da região (Vendrusculo *et al.*, 2006).

4 ANÁLISE DE TEXTOS JURÍDICOS

Na área jurídica, boa parte dos métodos de análise de texto tem por objetivo classificar o texto em alguma categoria já conhecida. Como exemplo dessa abordagem, 210 sentenças criminais da justiça chinesa foram analisadas com diferentes técnicas de aprendizado de máquina. Embora não seja tratado neste trabalho, o melhor resultado para esse exemplo foi obtido com redes neurais (Chou e Hsing, 2010).

Uma outra abordagem de classificação de sentenças envolve classificar determinados textos jurídicos binariamente de acordo com uma variável que se deseja estudar e, além disso, analisar fatores que possivelmente teriam impacto nessa classificação. Em um exemplo dessa estratégia, foram analisados textos de sentenças judiciais, as quais foram classificadas em relação à existência de violação dos direitos humanos em cada caso. Com essa abordagem, além de se atingir o resultado de 79% de acurácia na classificação, foi identificado que os antecedentes do processo (escritos pela própria corte) tinham o maior impacto na acurácia da classificação (Aletras *et al.*, 2016).

Já outras técnicas dão um passo à frente ao tentar não só classificar os textos, mas também compreender como determinadas estruturas sintáticas se organizam no corpo das sentenças judiciais. Dessa forma, é possível criar uma árvore de argumentação capaz de considerar as funções gramaticais das palavras no contexto de uma sentença, bem como avaliar a importância de cada uma no processo de argumentação (Wyner *et al.*, 2010)

Embora não seja tratado neste trabalho, outra abordagem razoavelmente comum para análise de textos é a classificação não supervisionada (mais precisamente chamada de agrupamento ou clusterização), usada comumente quando não se conhecem as categorias as quais se deseja definir. Um caso interessante envolvendo análise do vocabulário jurídico no processamento de linguagem jurídica foi apontado por pesquisadores da Eslováquia. Na ocasião, foi analisado um sistema da justiça local, o qual transcrevia eletronicamente para texto os diálogos jurídicos da corte. Como resultados dessa análise, foram apontados fatores de sucesso para o funcionamento adequado da técnica, por exemplo, identificação do gênero das pessoas do diálogo e vocabulário jurídico previamente determinado (Rusko *et al.*, 2011).

Se, por um lado, existe dificuldade de se fazerem análises semânticas dos textos judiciais com estatística descritiva simples – devido às especificidades sintáticas e ao vocabulário típico da área – , por outro, em concordância com o que foi possível identificar na literatura, boa parte do que vem sendo feito atualmente tenta contornar essas dificuldades com o uso de ferramentas que se tornaram destaque nos últimos anos, como a aplicação de técnicas de aprendizado de máquina e métodos de identificação de padrões baseado em inferências indutivas e interações de alta dimensionalidade (Palau e Moens, 2009).

5 METODOLOGIA

Língua natural, ou língua humana, refere-se a qualquer forma de comunicação – simbólica, oral ou escrita – desenvolvida naturalmente de forma não premeditada pelo ser humano como necessidade de se relacionar com aspectos de sobrevivência ou reprodutivos (Lyons, 1991). Para o Brasil, temos a língua portuguesa e a Libras como exemplos de linguagens naturais adotadas no território.

Na matemática ou na ciência da computação, é presente o surgimento de línguas artificiais, construídas e premeditadas, que antagonizam a língua natural. São formais, com regras extremamente rígidas quanto à sua composição sintática e semântica, de fácil e rápida compreensão por mecanismos autônomos e, ainda, costumam não sofrer variações para diferentes regiões. Para um computador, a linguagem natural é como uma linguagem alienígena, fora de qualquer estrutura a qual ele está habituado. Assim, um dos campos da ciência da computação – o processamento de linguagem natural – buscou especializar-se e desenvolver técnicas que permitem aos computadores a interpretação e manipulação de tais linguagens.

A parte mais crítica desse processo é a representação das palavras. Para nós humanos, palavras carregam sentimentos e memórias. Essas unidades de significado têm diferentes semânticas conforme seu contexto, entonação e demais características subjetivas da comunicação humana. Para um computador que segue uma estrutura formal objetiva, nada disso é representável, sendo apenas capaz de interpretar números e operações entre eles. Devemos então, a despeito dessas limitações, retratar toda essa subjetividade de forma objetiva por representações matriciais.

Um dos primeiros esforços para tentar reconstruir esses significados de forma numérica foi o modelo bolsa de palavras (*bag-of-words*), que consiste em criar uma lista com todo o vocabulário dos documentos analisados e, para cada um deles, a contagem das palavras do vocabulário presente no documento (Wang e Manning, 2012). Considere as duas sentenças *"*meu cachorro está doente*"* e *"*meu amigo e meu irmão têm alergia a cachorro". Criando o vocabulário e realizando a contagem de cada palavra para a respectiva sentença chegamos ao resultado demonstrado na tabela 1.

TABELA 1
Exemplo de *bag-of-words*

	a	alergia	amigo	cachorro	doente	e	está	irmão	meu	têm
Sentença 1	0	0	0	1	1	0	1	0	1	0
Sentença 2	1	1	1	1	0	1	0	1	2	1

Elaboração dos autores.

Essa foi uma das primeiras abordagens a ter grande sucesso para a representação estruturada de textos, mas os especialistas rapidamente perceberam um grande problema dessa abordagem: a ausência de contexto. Se, por exemplo, a palavra cachorro fosse seguida da palavra quente (*e.g.,* cachorro-quente) o significado desses vocábulos seriam completamente diferente. Porém, o modelo *bag-of-words* não capturaria essa simples composição de significado.

Assim, considerando esse fato, propõe-se a representação n-grama (*n-gram)*. Nela, ao invés de contar as palavras separadamente (aqui chamado 1-grama), contaríamos os pares/trios/etc. de ocorrências. Por exemplo, com *n* igual a 2, contaríamos os pares de palavras (para a sentença-1 teríamos a lista: "meu cachorro", "cachorro está", "está doente"), mantendo o contexto no qual aparecem, tornando-se uma forma muito mais rica de representação (Wang e Manning, 2012).

A representação n-grama embora tenha apresentado uma melhora significativa sobre o modelo *bag-of-words*, ainda se trata apenas de uma escala que não representa sua importância para a sentença – por exemplo, artigos e conectivos têm alta frequência de aparição, porém em nada contribuem para a essência do texto. Sentenças com temas diferentes terão muitas palavras insignificantes em comum, sendo então necessária uma forma de filtrar ou escalar essas palavras que não contribuem para a representação. São introduzidas, então, as palavras vazias *(stop-words)* e o ponderamento de palavras.

A expressão palavras vazias se refere à técnica de remover palavras sem importância para o processamento de linguagem natural. Para o exemplo da tabela 1, temos como palavras vazias *a*, *e*, *está* etc. Essa remoção contribui positivamente tanto em custos de armazenamento (o computador pode descartar essas representações) quanto em dimensão da representação (o modelo matemático não incorporaria essa palavra). Com menos termos para lidar, o processamento fica simplificado, melhorando a eficiência e eficácia da análise.

Essa técnica, embora resulte em ganhos sensíveis em qualidade de análise, requer uma cuidadosa seleção dos termos considerados irrelevantes. Como de costume, certas palavras vazias são disponibilizadas de maneira já pronta em pacotes de análise de texto. Se identificarmos uma ou outra palavra a ser adicionada a essa lista, podemos adicioná-la sem grandes problemas. Porém, em textos na área da física, por exemplo, a palavra *energia* pode ser uma palavra vazia, tendo em vista que tal vocábulo está presente em muitos campos da física e, por isso, não teria poder discriminativo entre os campos. Analogamente, no contexto jurídico, a palavra *lei* não possui um significado específico para discriminar documentos; em vez disso, o teor de que dispõe a referida lei contém a informação relevante para a análise semântica. Outro exemplo é a palavra *juiz* em documentos jurídicos. Adicionar manualmente tais palavras ao conjunto de palavras vazias é algo simples, mas que pode se tornar muito trabalhoso caso feito dessa maneira em todas as situações, sendo necessário uma forma inteligente e automática de resolver esse problema.

O ponderamento de palavras, junto com a frequência do termo (*term frequency*), busca atribuir a importância estatística considerando as frequências relativas das palavras. O método TFIDF (*term frequency-inverse document frequency*) pondera as palavras conforme suas frequências relativas, que é a contagem da palavra na sentença individual em razão da contagem da palavra em todas as sentenças (Katz, 1987).

Existem diversas fórmulas para a frequência de termos (Liu e Zhang, 2018; Bengfort, Bilbro e Ojeda, 2018). Aqui, adotamos a dupla normalização em 0.5 (sendo $f_{t,d}$ a contagem do termo t no documento atual d) (Liu e Zhang, 2018), que empiricamente gerou melhores *performances*:

$$\mathrm{tf}(t,d) = 0.5 + 0.5 \cdot \frac{f_{t,d}}{\max\{f_{t',d} : t' \in d\}} \tag{1}$$

Para o inverso da frequência nos documentos (*inverse document frequency*), sendo D o conjunto de todos os documentos e N a quantidade total de documentos, calculamos da seguinte forma:

$$\mathrm{idf}(t,D) = \log \frac{N}{|\{d \in D : t \in d\}|} \tag{2}$$

Por fim, multiplicamos a frequência dos termos pelo inverso da frequência nos documentos:

$$\mathrm{tfidf}(t, d, D) = \mathrm{tf}(t, d) \cdot \mathrm{idf}(t, D) \tag{3}$$

O modelo *bag-of-words*, n-grama e TFIDF apresentaram diversas melhorias na robustez representativa e, junto a modelos preditivos, trouxeram o estado da arte (Bengio, Goodfellow e Courville, 2017) em diversas tarefas de processamento de linguagem natural. Exemplo de melhoria envolve a área de classificação de textos (Wang e Manning, 2012), na qual foi possível superar em muito a *performance* humana. Entretanto, para outras tarefas mais complexas dessa classe – como tradução, resposta a perguntas, e resumo automático –, não foi possível obter melhoria significativa. Era necessária uma abordagem diferente. Com resultados de sucesso em diversos campos de visão computacional, redes neurais chamaram atenção de pesquisadores de diversas áreas. Pesquisadores de NLP também decidiram ingressar na exploração dessas técnicas. Redes neurais que consistem na busca de uma melhor representação para os dados têm um papel muito claro nesse contexto: ao invés de provermos *a priori* como os textos deverão ser representados, ficará a cargo da rede neural descobrir isso.

Com o *word embedding* (Mikolov *et al.*, 2013), uma lista de números é atribuída para cada palavra. Por meio de otimização, esses números convergem para uma região no espaço semântico (representações de linguagem natural que agregam termos e construtos latentes que possuem valor semântico similar) em que as palavras ocorrem em mesmo contexto e com variação dimensional relativa aos seus significados. Esse modelo criou um novo paradigma (Bengio, Courville e Vincent, 2013) a respeito do que e como se processar textos, em vista que a representação distribuída apresenta diversas propriedades interessantes, atingindo novos patamares em tradução automática, busca de respostas, sumarização automática, e outras tarefas de grande dificuldade para o NLP.

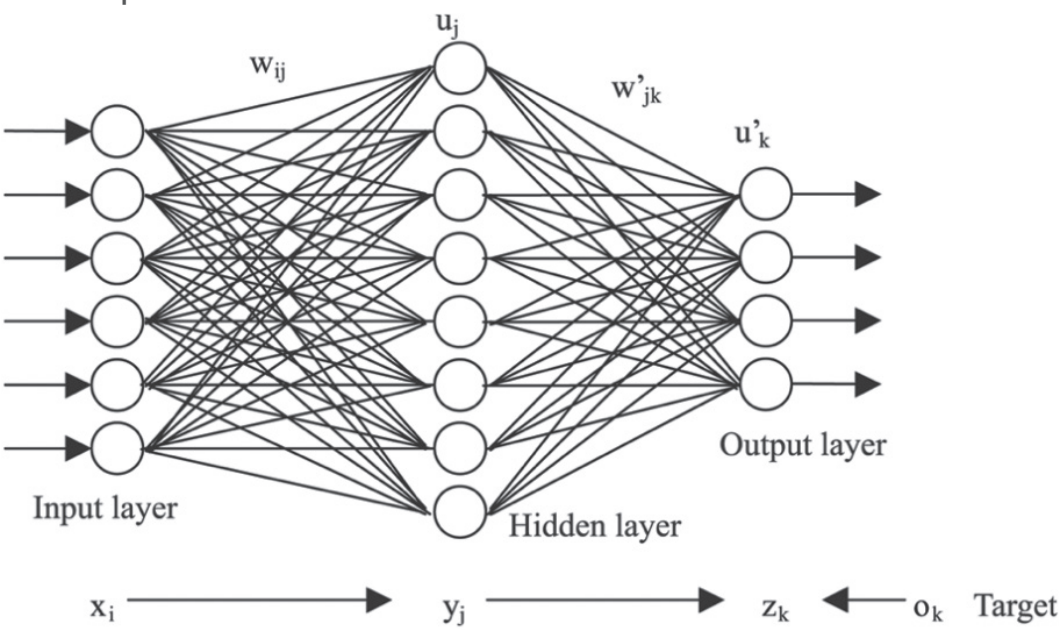
Com representações computacionais bem definidas, podemos então aplicar modelos estatísticos aos dados. Um dos modelos mais clássicos (Abu-Mostafa, Magdon-Ismail e Lin, 2012; Shalev-Shwartz e Ben-David, 2014) com objetivo de classificação é a regressão logística. De forma intuitiva, podemos pensar nela como um neurônio simples: ela recebe estímulos iniciais (no nosso caso, as palavras representadas

computacionalmente) e, então, dependendo desses estímulos e de como os dendritos do neurônio estão condicionados (importância estatística das palavras), realiza-se ou não uma sinapse (pertencer ou não a uma classe).

Porém, esse modelo é limitado à família das funções lineares. Em problemas mais complexos, costumamos ter dependências não lineares para o sinal analisado, assim, a regressão logística não apresenta, de modo geral, um bom desempenho nesse caso. Precisamos então construir um modelo que tenha como espaço de hipótese as funções não lineares que podem modelar tais problemas.

As redes neurais artificiais buscam empilhar a saída de regressões, com função de ativação semelhante à regressão logística, criando dependências não lineares entre os sinais de entrada – podendo, assim, modelar, na medida do representável, relações extremamente complexas.

FIGURA 1
Exemplo de rede neural



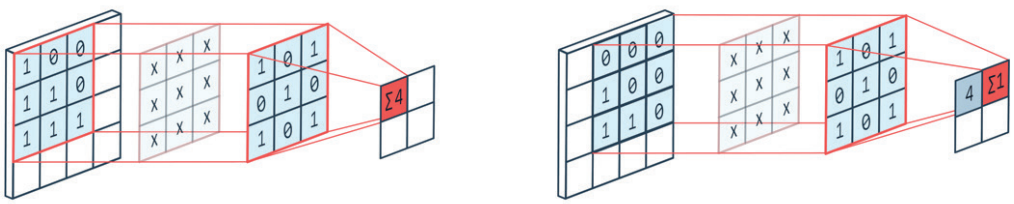
Fonte: Stack Exchange/Mathematics.[3]

3. Disponível em: <https://tinyurl.com/yb8opmew>. Acesso em: set. 2019.

Há limitações que devem ser impostas à complexidade de tais funções. Considerando que os dados normalmente são amostrais e apresentam ruídos e vieses, não podemos modelar perfeitamente os dados, pois caso contrário poderíamos ensinar até mesmo as imperfeições ao computador (Abu-Mostafa, Magdon-Ismail e Lin, 2012). Assim, algumas restrições e regularizações precisam ser impostas.

Convoluções são uma forma de representação em conjunto com a regularização. Elas apresentam interações esparsas, compartilhamento de parâmetros e representação equivariante. Essa combinação de características impõe a esse tipo de rede restrições sobre superadaptação, limitando a capacidade de serem aprendidas as imperfeições dos dados, melhorando o desempenho do modelo para o mundo real. A convolução de redes neurais é definida pela operação representada na figura 2.

FIGURA 2
Operação de convolução nas CNNs



Fonte: 2D Convolution Block.[4]

Para um aprendizado supervisionado, é mandatório a existência de etiquetas (*labels*) para cada registro (atributos os quais são preditos pelos modelos). Por exemplo, se queremos supervisionar um algoritmo no aprendizado de classificação de imagens para cães ou gatos, precisamos passar as imagens em conjunto com a etiqueta (resposta correta) daquela imagem (*e.g.*, cão ou gato).

Para supervisionar dados jurídicos, há um alto custo. No caso de imagens, rapidamente conseguimos dizer se uma figura mostra um cão ou gato, porém para textos (especialmente textos jurídicos), atribuir uma etiqueta demanda uma leitura cautelosa e interpretação de cada caso. Para contornar isso, foi utilizada uma abordagem semi-supervisionada para geração automática das etiquetas.

4. Disponível em: <https://tinyurl.com/y5sxrv37>. Acesso em: set. 2019.

Primeiro foram feitas avaliações estatísticas quanto a padrões básicos nos dados – estes extraídos por expressão regular – e, então, com uma validação amostral, confirmamos se poderíamos ou não designar certas etiquetas para os conjuntos de dados extraídos. Após isso, treinamos a regressão logística para prever novos dados e reutilizamos aqueles em que o modelo tinha alta confiança para um novo treinamento. Neste novo treinamento, palavras usadas na busca da expressão regular foram retiradas, forçando os modelos a aprenderem novos termos correlatos às etiquetas. Na seção dos resultados, são discutidos termos aprendidos que colaboram na determinação de uma classe.

6 RESULTADOS

6.1 Terminologia de classificadores

Para esta análise, são consideradas informações de três classificadores: *i*) a classificação Regex Complexo – já existente no sistema IpeaJus, se trata do classificador atual, do qual se extraíram os dados para essa pesquisa – usa somente expressões regulares[5] complexas; *ii*) a classificação Regex Trivial, que usa expressões regulares triviais, é utilizada como passo intermediário. *iii*) a classificação por Aprendizado de Máquina utiliza aprendizado de máquina baseado nas informações do classificador Regex Trivial. Essa última abordagem é o foco da análise deste trabalho, conforme explanado a seguir.

6.2 Espaço amostral e classificador Regex Complexo

Os dados utilizados nesta pesquisa provêm do banco de dados IpeaJus, desenvolvido pelo Ipea, no qual são armazenadas informações relativas a processos judiciais extraídas diretamente de alguns portais de acompanhamento processual de tribunais e da raspagem de textos de Diários Oficiais da Justiça (Jesus *et al.*, 2018).

O espaço amostral escolhido para a análise deste texto se baseia, exclusivamente, nos processos do TRF2, primeiro grau, com dados referentes aos assuntos de benefício de prestação continuada e aposentadoria rural. É importante ressaltar que os resultados,

5. Expressões regulares é o termo dado para expressões (letras, números, espaços, símbolos etc.) que aparecem de maneira regular e estruturada em um texto.

como era esperado, apresentam vocabulário e viés condizente com esse espaço amostral. Isso, porém, não deve interferir nos resultados, uma vez que nenhuma especificidade desse tribunal é tratada na análise.

A escolha desse tribunal e dos assuntos comentados acima se deve ao fato de que o banco de dados utilizado ainda está em processo de consolidação, apresentando resultados mais consistentes e contínuos para esses casos. Além disso, o TRF2 apresentou menor consistência de classificação de sentenças (com o uso do classificador Regex Complexo) em análises internas produzidas pelo projeto piloto, representando bom potencial de melhoria.

Apesar de o próprio sistema IpeaJus ser capaz de realizar as classificações das sentenças por meio de expressões regulares (busca por termos específicos presentes no corpo do texto da sentença), essa informação apresenta uma limitação importante, já que não se avaliou a precisão desse método, ou seja, não se conhece a proporção de classificações que são feitas de maneira correta ou incorreta.

Após ajustar manualmente o classificador já existente para o vocabulário e estrutura gramatical do TRF2, é possível classificar 79% das sentenças para o primeiro grau deste tribunal, enquanto os outros 21% não entram em nenhuma das dez classes. O volume de classificação e *performance* pode ser visto na tabela abaixo.

TABELA 2
Volume e percentual do classificador Regex Complexo

	Número de processos	Percentual (%)
Classificados	35309	79
Não classificados	9664	21
Total	**44973**	**100**

Elaboração dos autores.

Além disso, até que se chegue a um volume razoável de sentenças classificadas com esse classificador, são necessários ajustes manuais dessas expressões para cada tribunal (nesse caso, 79%).

Apesar de não ser conhecida a acurácia das classificações das sentenças (referente aqui aos 79%), são testados tais dados para treinamento por meio de teste empírico, sem sucesso. A acurácia de treinamento chegou a valores médios próximos de 16%, o

que mostra que os dados não apresentam estrutura regular em volume razoável para o treinamento. Em linhas gerais, com dez classes, 10% de acurácia poderia ser encarado como totalmente aleatório e 16% se torna insatisfatório. Assim, desconsideramos os dados do classificador Regex Complexo para nossa análise.

6.3 Classificador Regex Trivial

Antes de criarmos o classificador por Aprendizado de Máquina, precisamos ensinar ao computador, por meio de exemplos, quais casos pertencem a uma categoria ou outra. Assim, para criarmos essa amostra de dados, utilizamos uma classificação trivial e ingênua para termos uma direção de classificação.

Dessa forma, após tratarmos e limparmos os dados principais, e prepararmos um modelo *bag-of-words*, classificamos todos os processos em cinco principais classes. O método usado para fazer essa classificação é bastante simples: é procurado no texto o aparecimento exato das palavras *procedente*, *improcedente*, *parcialmente procedente,* e *acordo*. Quando a busca não retorna nenhum positivo, o texto é tido como não classificado (outros). Na tabela 3 é possível visualizar essas classificações.

TABELA 3
Classificação trivial com "outros"

Classificação	Número de processos	Percentual (%)
Procedente	10314	23
Improcedente	12497	28
Parcialmente procedente	1867	4
Acordo	4033	9
Outros	16262	36
Média/total	44973	100

Elaboração dos autores.

Essa informação é utilizada como palpite inicial e serve para treinarmos a máquina, com o objetivo de que esta aprenda outras características que possam ser importantes para essa classificação, e que não são tão evidentes.

Como queremos ensinar a classificar as sentenças em alguma das categorias de procedente, improcedente; parcialmente procedente e acordo, retiramos as sentenças classificadas como *outros* dessa etapa (tabela 4). Esses processos, no entanto, são analisados separadamente mais adiante.

TABELA 4
Classificação trivial sem "outros"

Classificação	Número de processos	Percentual (%)
Procedente	10314	36
Improcedente	12497	44
Parcialmente procedente	1867	7
Acordo	4033	14
Média/total	**28711**	**100**

Elaboração dos autores.

Como nosso objetivo não é ensinar o computador a classificar as sentenças com eficiência similar às expressões regulares, mas sim torná-lo capaz de considerar outros elementos não triviais que possam fazer uma análise complexa dos textos, é esperado atingir valores diferentes dos apresentados na tabela 4.

6.4 Métricas de avaliação

Para avaliar a efetividade do aprendizado da máquina, usamos as análises de *precisão*, *recall*, *F1-score* e *suporte*. No momento de treinamento da máquina, são usados 70% das sentenças para esse processo. O volume restante, correspondente a 30% (ou 8614 processos), é usado para testar o aprendizado da máquina. Esse volume (30%) é chamado de suporte. No geral, quanto maior é o suporte (em volume total, e não percentual), mais próximo de uma afirmação conclusiva chegamos (Padula *et al.*, 2017).

O conceito de precisão trata do volume de sentenças que são classificadas em uma categoria, e que são de fato dessa categoria. Essa métrica, no entanto, não considera aquelas sentenças que deveriam estar nessa classe, mas foram classificadas de outra forma. Tal fato motiva o uso do indicador recall, que mede o volume total de processos que foram classificados corretamente em uma determinada categoria, em questões absolutas. A métrica mais adequada, portanto, depende do objetivo. No entanto, podemos afirmar que usualmente uma análise consistente deve levar em conta as duas métricas. É justamente nesse sentido que o F1-score funciona, ao fazer uma média harmônica dos resultados para gerar um classificador de interpretação mais geral (Padula *et al.*, 2017).

As fórmulas para as três métricas supracitadas são: precisão, recall e F1-score.

$$Precisão = \frac{Verdadeiropositivo}{Verdadeiropositivo + Falsopositivo} \quad (4)$$

$$Recall = \frac{Verdadeiropositivo}{Verdadeiropositivo+Falsonegativo} \quad (5)$$

$$F1 = 2 \times \frac{Precisão*Recall}{Precisão+Recall} \quad (6)$$

6.5 Treinamento e classificador por aprendizado de máquina

Utilizando as informações provenientes da classificação da etapa anterior, treinamos o algoritmo para classificar essas classes, baseando-nos na categorização apresentada na seção anterior. Essa classificação usa 70% de todas as sentenças para treinar o modelo e 30% (8.614 sentenças) para calcular a precisão dessa previsão.

O treinamento e teste desse processo é realizado tanto para a análise de Unigramas (palavras individuais) quanto de Bigramas (todos os conjuntos de duas palavras vizinhas). Os resultados podem ser vistos nas tabelas 5 e 6.

TABELA 5
Resultado classificador por aprendizado de máquina baseado no aprendizado Regex Trivial, com unigramas

	Unigrama			
Classificação	Precisão (%)	Recall (%)	F1-score (%)	Suporte
Procedente	98	96	97	3749
Improcedente	94	95	94	3095
Parcialmente procedente	70	77	73	560
Acordo	99	99	99	1210
Média/total	**95**	**95**	**95**	**8614**

Elaboração dos autores.

É possível perceber que para os casos de classificação procedente, improcedente e acordo, os resultados atingem resultado muito satisfatório, com os valores da média harmônica (F1-score) em 97%, 94% e 99%.

Para as sentenças de parcialmente procedente (F1-score), o resultado de 73% reflete a expectativa que existe por um valor abaixo das outras classificações, devido ao fato intrínseco da maior subjetividade que existe nesse tipo de classe. Além disso, corrobora o fato de existir um volume menor dessa classe tanto para classificação (1.867 sentenças) quanto para teste (560 sentenças).

Realizando a mesma análise, porém dessa vez considerando o conjunto de duas palavras vizinhas também (bigramas), os resultados de precisão são apresentados na tabela 6.

TABELA 6
Resultado classificador por aprendizado de máquina baseado no aprendizado Regex Trivial, com bigramas

	Bigrama			
Classificação	Precisão (%)	Recall (%)	F1-score (%)	Suporte
Procedente	99	95	97	3749
Improcedente	97	85	90	3095
Parcialmente procedente	47	90	62	560
Acordo	99	99	99	1210
Média/total	**94**	**91**	**92**	**8614**

Elaboração dos autores.

É possível perceber que os resultados não apresentam grande mudança, com exceção das sentenças de parcialmente procedente. Estas tiveram o seu valor de precisão reduzido de 70% para 47%. Esse resultado aponta que outras sentenças (nesse caso, referentes à procedente, improcedente e acordo) estão sendo erroneamente classificadas como parcialmente procedentes.

Por outro lado, ainda para as sentenças de parcialmente procedente, o valor de recall sobe de 77% para 90%. Isso indica, em valores absolutos, que estão passando menos sentenças desse tipo (apenas 10% das sentenças parcialmente procedentes foram classificadas de alguma outra forma). Isso mostra que, potencialmente, o uso de expressões nessa análise ampliou os critérios de classificação dessa categoria.

De modo geral, no entanto, os resultados globais se mostraram muito parecidos, o que quantitativamente não representa uma grande diferença para casos gerais. Essa variação, no entanto, pode fazer diferença em casos específicos em que se priorize uma classificação sem falsos positivos ou sem falsos negativos.

Na seção seguinte (classificação de importância semântica) é feita uma análise quantitativa e qualitativa para tentar interpretar possíveis diferenças nos dois modelos de aprendizado (unigramas e bigramas) e como isso poderia impactar uma análise.

6.6 Classificação de importância semântica

Todas as palavras (tanto uni quanto bigramas) recebem um valor de importância global para a classificação de uma sentença judicial, que pode ser medido e apresentado em ordem de relevância. As tabelas 7, 8 e 9 apresentam as vinte palavras (ou conjunto de palavras, para osbigramas) obtidos nesse estudo. As palavras aparecem sem acento ou letras maiúsculas, pois foram normalizadas para análise. Na tabela 7, são apresentados os resultados para as sentenças judiciais classificadas como procedentes.

TABELA 7
Importância semântica das sentenças judiciais procedentes, em ordem decrescente

Procedente			
Unigrama		Bigrama	
Palavra	Importância	Palavra	Importância
procedente	5,20	julgo procedente	8,28
tutela	3,61	procedente	3,90
juros	3,03	data requerimento	3,12
condeno	2,97	procedente pedido	3,11
condenar	2,95	tutela	2,69
desde	2,79	condeno	2,33
requerimento	2,69	juros	2,27
pagar	2,64	condenar	2,19
parcelas	2,58	desde	2,03
alimentar	2,44	pagar	2,01
vencidas	2,36	parcelas	1,91
embargos	2,31	efeitos tutela	1,88
valores	2,30	vencidas	1,83
inss	2,10	condeno inss	1,82
administrativo	2,08	alimentar	1,79
carater	2,06	beneficio	1,67
antecipo	2,05	administrativo fl	1,61
calculos	2,00	mora	1,60
beneficio	1,91	inss	1,60
calculados	1,90	calculados	1,59

Elaboração dos autores.

Como era de se esperar, a palavra procedente é a mais importante para classificar as sentenças judiciais, *a priori*, como procedentes. Essa observação não só é válida para a abordagem de bigramas, mas também para as análises das classificações de improcedente, parcialmente procedente e acordo, as quais apresentam os próprios nomes das

classes em primeiro lugar. No entanto, nosso maior interesse (indireto) está justamente no aprendizado da máquina de outros fatores considerados não triviais, como os apresentados junto a esses citados.

Palavras como *tutela*, *juros*, *condeno*, *pagar*, *requerimento*, entre outras, provavelmente não seriam de identificação fácil por pesquisadores da área do direito ou de outros campos. Outro ponto interessante a se observar é o aparecimento de termos específicos dos assuntos analisados. Palavras como *inss*, *alimentar* e *condeno inss* poderiam parecer palavras aleatórias para qualquer análise de sentenças de aposentadoria rural e benefício de aposentadoria especial. Porém, esses termos só aparecem para a lista da classe procedente, mostrando, assim, o quanto contra intuitivo pode ser a influência de cada termo em uma classificação. Na tabela 8 são apresentados os resultados para as sentenças improcedentes.

TABELA 8
Importância semântica das sentenças judiciais improcedentes, em ordem decrescente

Improcedente			
Unigrama		Bigrama	
Palavra	Importância	Palavra	Importância
improcedente	10,77	improcedente	8,76
nao	5,82	improcedente pedido	6,70
julgada	2,65	julgo improcedente	6,14
reclamacao	2,49	nao	5,18
pedido	2,24	julgada	2,54
baixa	2,16	reclamacao	2,47
concessao	2,12	julgado baixa	2,47
miserabilidade	1,92	autora nao	1,84
improcedencia	1,84	pedido	1,80
improcedentes	1,76	miserabilidade	1,74
arquivem	1,59	improcedentes	1,65
julgo	1,57	concessao	1,52
extinguindo	1,54	improcedencia	1,49
superior	1,51	nao restou	1,47
deficiencia	1,47	baixa	1,46
comprovacao	1,43	controle	1,42
alem	1,42	pedido custas	1,42
maioria	1,41	constitucionalidade	1,38
controle	1,39	superior	1,26
transitada	1,36	concessao beneficio	1,26

Elaborado pelos autores.

Pode-se perceber que para o caso das sentenças judiciais improcedentes, por exemplo, a máquina é capaz de aprender outras variações da palavra improcedência, bem como determinar a importância relativa elevada para estas.

Fica evidente, também, a importância da palavra *não*. Nos unigramas, ela aparece como a segunda palavra mais importante, enquanto nos bigramas apareceu de três formas diferentes. Isso mostra que, apesar de não se utilizar a palavra *não* de forma explícita em nenhum momento do treinamento, a máquina consegue aprender o sentido de negatividade que essa palavra traz e a importância que isso tem na negação de um pedido considerado improcedente. Na tabela 9, são apresentados os resultados para as sentenças parcialmente procedentes.

TABELA 9
Importância semântica das sentenças parcialmente procedentes, em ordem decrescente

Parcialmente procedente			
Unigrama		Bigrama	
Palavra	Importância	Palavra	Importância
parcialmente	24,97	parcialmente	23,29
data	3,33	data	3,08
partir	3,11	partir	2,53
beneficio	2,61	beneficio	2,23
ajuizamento	2,39	valores	2,19
dano	2,29	remessa	2,15
remessa	2,10	dano	2,12
rua	2,02	incapacidade	2,06
correcao	1,93	ajuizamento	1,98
oscar	1,90	devera	1,87
devera	1,88	correcao	1,84
fl	1,87	correcao monetaria	1,79
incapacidade	1,84	procedente pedido	1,76
valores	1,82	rural	1,72
monetaria	1,79	ministros	1,65
citacao	1,77	rj telefone	1,63
rural	1,77	embargos	1,61
mora	1,72	modulacao	1,60
tnu	1,70	monetaria	1,60
ate	1,69	data ajuizamento	1,59

Elaboração dos autores.

Os textos da classe de parcialmente procedente, diferentemente dos de procedente e improcedente, apresentam vocabulário misto das duas partes. Esse é o comportamento esperado, já que essa classe é composta exatamente da união dos dois tipos de sentenças.

No entanto, podemos salientar que a palavra *parcialmente* se destaca muito (inclusive se comparada com as outras classes) em valores globais e relativos. Isso indica que pode haver, nesse caso, uma maior dependência do aparecimento desse termo de forma explicita no texto para que haja uma classificação correta dessa categoria.

TABELA 10
Importância semântica das sentenças de acordo, em ordem decrescente

Acordo			
Unigrama		Bigrama	
Palavra	Importância	Palavra	Importância
acordo	6,73	acordo	6,42
homologo	3,92	homologo	3,45
partes	3,01	partes	2,73
proposta	2,61	proposta	2,52
audiencia	2,10	transacao	1,99
transacao	2,08	audiencia	1,92
iii	1,98	iii cpc	1,86
dip	1,77	iii	1,69
rpv	1,74	dip	1,64
homologatoria	1,70	rpv	1,62
processo	1,65	termos	1,51
resolucao	1,57	intimada	1,34
termos	1,57	dias	1,34
extingo	1,52	processo	1,28
intimada	1,49	resolucao	1,26
juridicos	1,47	juridicos	1,26
dias	1,44	extinto processo	1,23
implantacao	1,42	expeca	1,20
extinto	1,39	intime	1,19
expeca	1,39	extingo	1,19

Elaboração dos autores.

A classificação de acordo apresenta termos muito distintos uns dos outros e um grau de importância equilibrado. Isso indica uma possível flexibilidade para a classificação dessa categoria. Isto é, mesmo quando um termo importante não aparece, ainda é possível

classificar essa sentença corretamente. Essa hipótese corrobora com o fato de o índice de classificação ser o maior para essa classe, tendo precisão, recall e F1-score em 99%.

Destaca-se também o aparecimento de siglas e termos, como *dip*, *rpv*, *iii*, e *iii cpc*. Esses termos, apesar de sem significância, quando analisados isoladamente são relevantes para a classificação desses dados nessas categorias. Os termos com *iii*, por exemplo, aparecem nas referências que são feitas quando o juiz declara o acordo.

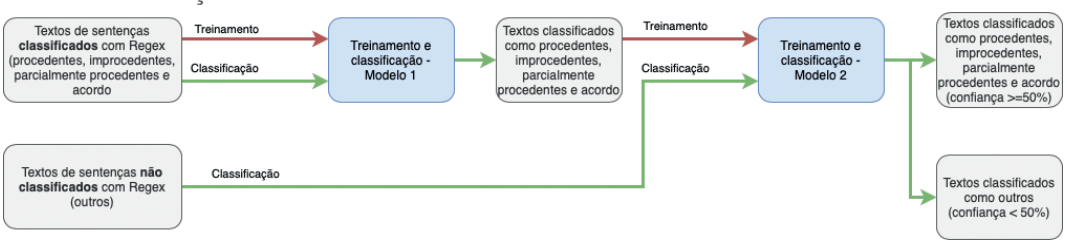
6.7 Classificação das sentenças classificadas como "outros"

A maior parte dos processos recaem sobre as quatro classes que definimos: procedente, improcedente, parcialmente procedente e acordo (aproximadamente 64%). Porém existem, além dessas, outras classes de menor representatividade que foram descartadas dada a insignificância estatística ou a inexistência de um fator determinístico para discriminação, as quais foram, então, classificadas no passo anterior como "outros".

Com o aprendizado estabelecido no passo anterior, um novo modelo é treinado usando os resultados obtidos apenas das quatro principais classes. Com essa estratégia, espera-se que os padrões textuais mais complexos que compõem cada classe possam ser identificados pela nova rodada do algorítmo.

Como essa segunda rodada de classificação forçará a classificação em uma das quatro categorias, utilizou-se um limiar de confiança para definir as sentenças que seriam classificadas ou não (permanecendo como outras). Definimos, intuitivamente, a chance aleatória pelo percentual de documentos por classe, o que não supera 40% por classe. Assim, toma-se como limiar de decisão o valor de 50% de confiança – valor padrão adotado quanto à teoria de decisão ingênua (Spitzer, 2013).

FIGURA 3
Reclassificação dos textos "outros"



Elaboração dos autores.
Obs.: Figura cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Os resultados das classificações obtidas após a segunda rodada do modelo são dispostas na tabela 11.

TABELA 11
Classificação total das sentenças

Classificação	Número de processos	Percentual (%)
Procedente	39	0,2
Improcedente	768	4,8
Parcialmente procedente	6243	38
Acordo	3462	21
Outros	6001	36
Total	**16513**	**100**

Elaboração dos autores.

É possível perceber que, de 36% das sentenças que não foram classificadas anteriormente, 64% passaram a ser classificadas em alguma categoria após essa análise. Obtém-se, portanto, uma taxa de efetividade total na classificação de 87%. O volume de classificações na categoria procedente em 0,2% indica que existe, possivelmente, uma estrutura consistente e clara para que se obtenha bons resultados ainda na etapa anterior, não tendo muito espaço para melhorias nessa etapa. A categoria improcedente apresenta comportamento similar à classe procedente, embora apresente valor superior igual a 4,8%.

A baixa classificação da categoria procedente pode indicar que: *i*) existe um padrão claro nessas sentenças, o que permite que sejam classificadas corretamente na primeira tentativa; e que *ii*) o algoritmo, mesmo com a análise de confiança, não teve melhoria significativa para essa classe.

Já para as classes de parcialmente procedente e acordo, o comportamento é diferente. As sentenças classificadas como parcialmente procedentes representam a maior parcela da classificação, com valor percentual em 38%. Isso equivale a um avanço significativo da classificação, especialmente se comparado com o volume total obtido na primeira etapa (classificado por Regex Trivial). O comparativo das tabelas 3 e 11 pode ser visto na tabela 12.

TABELA 12
Comparativo da reclassificação da categoria outros
(Em %)

Classificação	Classificação Regex Trivial	Reclassificação de outros
Procedente	23	0,2
Improcedente	28	5
Parcialmente procedente	4	38
Acordo	9	21
Outros	36	36
Total	**100**	**100**

Elaborado pelos autores.

Considerando os resultados, é possível inferir que – em ambiente de maior complexidade, como é o das sentenças de parcialmente procedente – considerar uma abordagem que leve em conta essa natureza pode trazer grandes benefícios. Assim, é possível afirmar que a classe de maior ganho para essa técnica é justamente a que obteve menor eficiência nas etapas anteriores, representando o maior ganho de valor nessa etapa de refinamento.

7 CONCLUSÃO

Este texto para discussão teve como objetivo apresentar uma aplicação de mineração de texto em grandes volumes de análise de classificação não supervisionada usando processamento de linguagem natural, no contexto jurídico do TRF2 com assunto de benefício de prestação continuada e aposentadoria rural, destacando alguns dos principais diferenciais do uso dessa tecnologia. De uma maneira geral, foi possível ensinar o computador a interpretar textos da área do direito, sem conhecimento prévio da teoria que embasava esse assunto, apontando apenas as classes que desejávamos que fossem classificadas.

Em especial, a classificação foi satisfatória para as categorias "procedente", "improcedente" e "acordo", enquanto que a classe "parcialmente procedente" obteve resultados menos atrativos – um resultado também esperado – haja vista a presença de elementos de subjetividade para definir sentenças dessa categoria, mesmo para juristas humanos.

Ademais, foram identificados termos e expressões com maior relevância para cada classe considerada, os quais podem ser utilizados como expressões regulares para as classificações dessas categoriais em outros trabalhos ou, ainda, no refinamento de classificações já existentes. Desse modo, o modelo proposto por este estudo pode subsidiar a obtenção de insumos para análises futuras baseadas em mineração de texto, explorando ainda mais o valor latente presente nos documentos judiciais dos tribunais brasileiros.

A inteligência artificial traz grande vantagem competitiva para as empresas privadas que a usam para analisar demandas, fraudes, gargalos, melhorias potenciais, análises de investimento, entre outros (Al-Augby, 2016; Kumar e Ravi, 2016). Com o emprego dessas ferramentas, a administração pública também pode se beneficiar dos potenciais ganhos de eficiência. Nesse sentido, esse trabalho buscou fornecer uma discussão inicial acerca do uso de técnicas de ciência de dados no contexto de análise de textos judiciais, definindo conceitos básicos de mineração de texto e apresentando um fluxo básico de tratamento e análise dos dados que possa ser replicado em futuros estudos correlatos.

No contexto de análises de processos judiciais, a mineração de texto como a apresentada neste estudo poderia ser diretamente aplicada, sem grande adaptação, à classificação de sentenças de outros tribunais, assuntos ou tipos de sentença. Como exemplo de aplicação em outros tipos de classe, pode-se citar as sentenças de execução. Estas, embora não se tenha atualmente um banco de dados consolidado com as informações necessárias, poderiam ser classificadas simplesmente como executadas ou não executadas, possibilitando a criação de um indicador de eficiência jurídica mais fidedigno, já que se é esperado atingir 100% de execução para uma eficiência teórica perfeita.

Como extensão dessa pesquisa, uma alternativa interessante é a aplicação do método em sentenças de fontes de diferentes origens, bem como a comparação de eficiência entre os resultados de classificação deste estudo. A expectativa é que os indicadores de *performance* de classificação sejam similares, mesmo que o modelo proposto seja replicado em textos cuja forma de escrita seja drasticamente diferente dos textos analisados por este trabalho. Outra análise em potencial é a classificação manual de amostra relevante das classificações que foram feitas, já que esta seria a única forma de garantir que o aprendizado da máquina é superior àquele feito na classificação trivial com expressão regular, embora análises não sistemáticas tenham sido realizadas e corroborem para essa hipótese ser verdadeira.

REFERÊNCIAS

ABU-MOSTAFA, Y. S.; MAGDON-ISMAIL, M.; LIN, H.-T. **Learning from data**. AML-Book, 2012.

AL-AUGBY, S. *et al*. Proposed investment decision support system for stock exchange using text mining method. *In*: AL-SADEQ INTERNATIONAL CONFERENCE ON MULTI-DISCIPLINARY IN IT AND COMMUNICATION SCIENCE AND APPLICATIONS (AIC-MITCSA), 2016, Baghdad, Iraq. **Anais…** New York: IEEE, 2016.

ALBUQUERQUE, P. H. *et al*. **Na era das máquinas, o emprego é de quem?** estimação da probabilidade de automação de ocupações no Brasil. Brasília: Ipea, 2019. (Texto para Discussão, n. 2457).

ALETRAS, N. *et al*. Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective. **PeerJ Computer Science**, v. 2, p. e93, 2016.

AQUINO, L.; COLARES, E. Acesso à justiça nos juizados especiais federais. **Boletim de Análise Político-Institucional**, Ipea, n. 3, p. 77-84, 2013.

ARMONAS COLOMBO, B.; BUCK, P.; MIANA BEZERRA, V. Challenges when using jurimetrics in Brazil – a survey of courts. **Future Internet**, v. 9, n. 4, p. 68, 2017.

BENGFORT, B.; BILBRO, R.; OJEDA, T. **Applied Text Analysis with Python**: enabling language-aware data products with machine learning. Sebastopol, CA: O'Reilly Media, 2018.

BENGIO, Y.; COURVILLE, A.; VINCENT, P. Representation learning: a review and new perspectives. **IEEE transactions on pattern analysis and machine intelligence**, v. 35, n. 8, p. 1798-1828, 2013.

BENGIO, Y.; GOODFELLOW, I.; COURVILLE, A. **Deep learning**. Cambridge, MA: The MIT Press, 2017.

BOSCH, M.; RIPANI, L. **El futuro del trabajo en América Latina y el Caribe**: ¿una gran oportunidad para la región? (versión interactiva). IDB, 2018.

CAMPOS, A. G. **Justiça do trabalho e produtividade no Brasil**: checando hipóteses dos anos 1990 e 2000. Rio de Janeiro: Ipea, 2017. (Texto para Discussão, n. 2330).

_____.**Resolução dos conflitos laborais no Brasil**: os papéis desempenhados pela Justiça do Trabalho. Rio de Janeiro: Ipea, 2018. (Texto para Discussão, n. 2362).

CAMPOS, A. G.; DI BENEDETTO, R. **Insumos para a regulamentação do Funget**: informações sobre execuções na justiça do trabalho. Rio de Janeiro: Ipea, 2015. (Texto para Discussão, n. 2140).

CASTRO, A. S. Modelos de decisão judicial e políticas públicas. **Radar: tecnologia, produção e comércio exterior**, Ipea, n. 22, p. 83, 2012.

_____. Sobre a produtividade dos serviços notariais e de registro no Brasil. **Boletim de Análise Político-Institucional**, Ipea, n. 5, p. 83-90, 2014.

_____. Uma Avaliação do impacto da reforma dos títulos executivos extrajudiciais (Lei n. 11.382/2006). **Boletim de Análise Político-Institucional**, Ipea, n. 8, p. 63-70, 2015.

CASTRO, A. S.; COELHO, D. S. C. **Indicadores básicos e desempenho da justiça estadual de primeiro grau no Brasil**. Brasília: Ipea, 2011. (Texto para Discussão, n. 1609).

CASTRO, A. S.; CUNHA, A. S. Dez anos de reformas na justiça: resultados e desafios. *In*: MONASTERIO, L.; NERI, M.; SOARES, S. (Eds.). **Brasil em Desenvolvimento 2014** – Estado, Planejamento e Políticas Públicas. Brasília: Ipea, p. 213-230, 2014.

CASTRO, A. S.; JESUS, L. A. **Judicialização dos pedidos de benefício de prestação continuada e aposentadoria rural** – TRF-1. Brasília, Ipea, 2018. (Nota Técnica Astec, n. 9).

CASTRO, A. S.; ROMEIRO, A. C.; CAVALCANTI, M. A. F. H. **Indicadores judiciais do mercado de crédito**: definição, metodologia e resultados. Brasília: Ipea, 2019. (Carta de Conjuntura, n. 43).

CAVALCANTE, P.; CAMÕES, M. Inovação no setor público: avanços e caminho a seguir no Brasil. *In*: CAVALCANTE, P. *et al*. **Inovação no setor público**: teoria, tendências e casos no Brasil. Brasília: Ipea; Enap, 2017.

CHOU, S.; HSING, T.-P. Text mining technique for Chinese written judgment of criminal case. *In*: PACIFIC-ASIA WORKSHOP INTELLIGENCE AND SECURITY INFORMATICS, 2010, Hyderabad, India. **Anais...** Switzerland: Springer, 2010.

CORREA, M. A. O. S.; LEAL, A. G. Identification of Overpricing in the Purchase of Medication by the Federal Government of Brazil, Using Text Mining and Clustering Based on Ontology. *In*: INTERNATIONAL CONFERENCE ON CLOUD AND BIG DATA COMPUTING, 2., 2018, Barcelona, Spain. **Anais...** New York: Association for Computing Machinery, 2018.

CUNHA, A. S. *et al*. **Custo unitário do processo de execução fiscal na Justiça Federal**: relatório de pesquisa. Brasília: Ipea; CNJ, 2011.

CUNHA, A. S.; MEDEIROS, B. A. SILVA, P. E. A. **Time and Cost in Brazilian Federal Courts**: tax foreclosure judicial proceedings. Brasília: Ipea, 2012. 10 lâminas.

CUNHA, A. S.; SILVA, P. E. A. Pesquisa empírica em direito. *In*: ENCONTRO DE PESQUISA EMPÍRICA EM DIREITO, 1., 2011, Ribeirão Preto, São Paulo. **Anais...** Brasília: Ipea, 2013.

CUNHA, A.; KLIN, I. V.; GOMES, O. A. **Custo e tempo do processo de execução fiscal promovido pela Procuradoria Geral da Fazenda Nacional (PGFN)**. Brasília: Ipea, 2011. (Comunicados do Ipea, n. 127).

DEMARZO, M. S. Internet das coisas: considerações acerca de consequências para o planejamento urbano e políticas públicas. **Brazilian Journal of Development**, v. 4, n. 7, p. 4209-4218, 2018.

FREY, C. B.; OSBORNE, M. A. **The future of employment**: How susceptible are jobs to computerisation? **Technological forecasting and social change**, v. 114, p. 254-280, 2017.

HADDAD, R. N. A motivação das decisões judiciais e a jurimetria: contribuições possíveis. *In*: Encontro Nacional do CONPEDI, 19., 2010, Fortaleza, Ceará. **Anais...** Florianópolis: Fundação Boiteux, v. 9, p. 10-11, 2010.

IPEA – INSTITUTO DE PESQUISA ECONÔMICA APLICADA. **Catálogo de Bases de Dados para Pesquisa** – Versão 1.4. Brasília: Ipea, 6 maio 2020. Disponível em: <https://bit.ly/3jW2ZGy>.

JESUS, L. A. *et al*. **Judicialização dos pedidos de benefício de prestação continuada e aposentadoria rural-Justiça Federal da 2a região**. Brasília: Ipea, 2018. (Nota técnica Astec, n. 8).

KATZ, S. Estimation of probabilities from sparse data for the language model component of a speech recognizer. **IEEE Transactions on Acoustics, Speech, and Signal Processing**, v. 35, n. 3, p. 400-401, mar. 1987.

KUMAR, B. S.; RAVI, V. A survey of the applications of text mining in financial domain. **Knowledge-Based Systems**, v. 114, p. 128-147, 2016.

LIU, Y.; ZHANG, M. **Neural network methods for natural language processing**. Williston, VT: Morgan & Claypool Publishers, 2018.

LOEVINGER, L. Jurimetrics – the next step forward. **Minnesota Law Review**, v. 33, 1948.

LUCINI, F. R. *et al*. Text mining approach to predict hospital admissions using early medical records from the emergency department. **International journal of medical informatics**, v. 100, p. 1-8, 2017.

LYONS, J. **Natural Language and universal grammar**: essays in linguistic theory. Cambridge: Cambridge University Press, 1991. v. 1

MASCHIO, P. *et al*. Um panorama acerca da mineração de dados educacionais no Brasil. *In*: Simpósio Brasileiro de Informática na Educação, 29., 2018, Fortaleza, Ceará. **Anais...** CBIE, p. 1936, 2018.

MENEZES, D.; BARROS, G. P. Breve análise sobre a jurimetria, os desafios para a sua implementação e as vantagens correspondentes. **Duc In Altum – Cadernos de Direito**, v. 9, n. 19, 2018.

MIKOLOV, T. *et al*. Efficient estimation of word representations in vector space. *In*: INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS, 1., 2013, Scottsdale, Arizona. **Anais…** 2013.

MILENA, M.; SERRA, P. Como utilizar elementos da estatística descritiva na jurimetria. **Revista Eletrônica do Curso de Direito das Faculdades OPET**, Curitiba, v. 4, n. 10, jun./dez., 2013.

MUKHERJEE, S.; SHAW, R. Big data–concepts, applications, challenges and future scope. **International Journal of Advanced Research in Computer and Communication Engineering**, v. 5, n. 2, p. 66–74, 2016.

NISHAVATHI, E.; JEYSHANKAR, R. Mapping the Science of Law: A Jurimetrics Analysis. **Library Philosophy and Practice**, Lincoln, p. 1–10, 2018.

OLIVEIRA, D. J. S.; BERMEJO, P. H. S. Mídias sociais e administração pública: análise do sentimento social perante a atuação do Governo Federal brasileiro. **Organizações & Sociedade**, v. 24, n. 82, p. 491-508, 2017.

PADULA, A. J. A. *et al*. **Segurança pública e inteligência artificial**: um estudo georreferenciado para o distrito federal. Brasília: Codeplan, 2017. (Texto para Discussão, n. 33).

PALAU, R. M.; MOENS, M.-F. Argumentation mining: the detection, classification and structure of arguments in text. *In*: INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW, 12., 2009, Barcelona, Spain. **Anais...** Association for Computing Machinery: New York, 2009

PINHEIRO, A. C. **Judiciário, reforma e economia**: a visão dos magistrados. Brasília: Ipea, 2003. (Texto para Discussão, n. 966).

_____. Reforma do Judiciário: uma nova fase. **Desafios do Desenvolvimento** – Ipea, v. 2, n. 6, p. 23, jan. 2005.

RAJA, U. *et al*. Text mining in healthcare: applications and opportunities. **J Healthc Inf Manag**, v. 22, n. 3, p. 52-6, 2008.

RUSKO, M. *et al*. Slovak automatic dictation system for judicial domain. *In*: LANGUAGE AND TECHNOLOGY CONFERENCE, 5., 2011, Poznań, Poland. **Anais...** Switzerland: Springer, 2011.

SCHWENGBER, S. B. **Mensurando a eficiência no sistema judiciário**: métodos paramétricos e não-paramétricos. 2006. Tese (Doutorado em Economia) – Faculdade de Economia, Administração, Contabilidade e Ciência da Informação e Documentação, Universidade de Brasília, Brasília, 2006.

SHALEV-SHWARTZ, S.; BEN-DAVID, S. **Understanding machine learning**: from theory to algorithms. Cambridge University Press, 2014.

SILVA, F. S. Justiça e território: estado da arte, abordagens possíveis e questões problemáticas a partir de uma meta-análise de estudos recentes. *In*: BOUERI, R; COSTA, M. A. (Eds.). **Brasil em Desenvolvimento** – Estado, Planejamento e Políticas Públicas. Brasília: Ipea, 2013. p. 173-196.

SPITZER, F. **Principles of random walk**. New York: Springer Science & Business Media, 2013.

VENDRUSCULO, L. G. *et al*. Aplicação da técnica de Text Mining e espacialização de informacões sócio-econômicas em sistemas objetivos de previsão de safra para a região da bacia hidrográfica do Pantanal. In: SIMPÓSIO DE GEOTECNOLOGIAS NO PANTANAL, 1., 2006, Campo Grande, MS. **Anais...** Embrapa Informática Agropecuária/INPE, 2006.

WANG, S.; MANNING, C. D. Baselines and Bigrams: Simple, Good Sentiment and Topic Classification. Annual Meeting of the Association for Computational Linguistics, 50., 2012, Jeju Island, South Korea. **Anais...** USA: Association for Computational Linguistics, 2012.

WHITE, M. Digital workplaces: vision and reality. **Business information review**, v. 29, n. 4, p. 205-214, 2012.

WYNER, A. *et al*. Approaches to text mining arguments from legal cases. *In*: FRANCESCONI, E. *et al*. (Eds.). **Semantic processing of legal texts**. Berlin: Springer, 2010. p. 60–79.

YUAN, W. *et al*. Linguistic feature analysis on judicial decisions based on keyword extraction and high-frequency word statistics – taking paper of sentence for example. *In*: **Recent Developments in Intelligent Computing, Communication and Devices**. Singapore: Springer, 2019. p. 1135-1142.

ZABALA, F.; SILVEIRA, F. Jurimetria: estatística aplicada ao Direito. **Revista Direito e Liberdade**, v. 16, n. 1, p. 87-103, 2014.

Ipea – Instituto de Pesquisa Econômica Aplicada

Assessoria de Imprensa e Comunicação

EDITORIAL

EDITORIAL

Coordenação
Reginaldo da Silva Domingos

Assistente de Coordenação
Rafael Augusto Ferreira Cardoso

Supervisão
Camilla de Miranda Mariath Gomes
Everson da Silva Moura

Revisão
Amanda Ramos Marques
Ana Clara Escórcio Xavier
Clícia Silveira Rodrigues
Idalina Barbara de Castro
Luiz Gustavo Campos de Araújo Souza
Olavo Mesquita de Carvalho
Regina Marta de Aguiar
Hellen Pereira de Oliveira Fonseca (estagiária)
Ingrid Verena Sampaio Cerqueira Sodré (estagiária)

Editoração
Aeromilson Trajano de Mesquita
Cristiano Ferreira de Araújo
Danilo Leite de Macedo Tavares
Herllyson da Silva Souza
Jeovah Herculano Szervinsk Junior
Leonardo Hideki Higa

Capa
Danielle de Oliveira Ayres
Flaviane Dias de Sant'ana

Projeto Gráfico
Renato Rodrigues Bueno

The manuscripts in languages other than Portuguese published herein have not been proofread.

Livraria Ipea
SBS – Quadra 1 – Bloco J – Ed. BNDES, Térreo
70076-900 – Brasília – DF
Tel.: (61) 2026-5336
Correio eletrônico: livraria@ipea.gov.br

Composto em adobe garamond pro 12/16 (texto)
Frutiger 67 bold condensed (títulos, gráficos e tabelas)
Brasília-DF

Missão do Ipea
Aprimorar as políticas públicas essenciais ao desenvolvimento brasileiro por meio da produção e disseminação de conhecimentos e da assessoria ao Estado nas suas decisões estratégicas.

ISSN 1415-4765

ipea Instituto de Pesquisa Econômica Aplicada

MINISTÉRIO DA
ECONOMIA

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL