

FUNDAÇÃO GETULIO VARGAS
ESCOLA DE DIREITO DO RIO DE JANEIRO
GRADUAÇÃO EM DIREITO

ALESSANDRA CORRÊA CID

**Construção de um modelo preditivo para decisões do STF em Ações Diretas de
Inconstitucionalidade**

Rio de Janeiro

2018

FUNDAÇÃO GETULIO VARGAS
ESCOLA DE DIREITO DO RIO DE JANEIRO
GRADUAÇÃO EM DIREITO

ALESSANDRA CORRÊA CID

**Construção de um modelo preditivo para decisões do STF em Ações Diretas de
Inconstitucionalidade**

Trabalho de Conclusão de Curso, sob orientação do professor Leandro Molhano e co-orientação do professor Renato Rocha, apresentado à FGV DIREITO RIO como requisito parcial para obtenção do grau de bacharel em Direito.

Rio de Janeiro

2018

FUNDAÇÃO GETULIO VARGAS
ESCOLA DE DIREITO DO RIO DE JANEIRO
GRADUAÇÃO EM DIREITO

Construção de um modelo preditivo para decisões do STF em Ações Diretas de
Inconstitucionalidade

ALESSANDRA CORRÊA CID

Trabalho de Conclusão de Curso apresentado à FGV
DIREITO RIO como requisito parcial para obtenção
do grau de bacharel em Direito.

Comissão Examinadora:

Prof. Leandro Molhano

Prof. Renato Rocha

Prof. Diego Werneck

Prof. Thomaz Pereira

Assinaturas:

Nota Final: _____

Rio de Janeiro, 19 de dezembro de 2018.

RESUMO

Este trabalho objetiva iniciar a construção de um modelo de previsão de decisões judiciais utilizando um algoritmo de *machine learning* para prever decisões do Supremo Tribunal Federal (STF). A ideia que guia este trabalho é a de que testando-se e avaliando-se elementos do processo decisório judicial que são importantes para a realização de uma análise preditiva da decisão, é possível delimitar como e quais fatores podem estar influenciando no processo decisório. Para tanto, inicialmente, analisa-se a literatura que estuda a influência de fatores cognitivos, econômicos e institucionais nas decisões. Num segundo momento, desenvolve-se um modelo preditivo de decisões do Supremo Tribunal Federal em Ações Diretas de Inconstitucionalidade, utilizando um algoritmo de *random forest*. O algoritmo foi aplicado duas vezes. Na primeira, considerou-se os seguintes elementos como *features* do modelo: (i) relatórios das decisões; (ii) ano em que a decisão foi publicada; (iii) ministro relator; e (iv) requerente. Na segunda aplicação foi considerado, além do mencionado anteriormente, o voto do ministro relator. Na primeira aplicação, o algoritmo acertou o resultado de 85% dos casos previstos e, na segunda aplicação, acertou 90% dos resultados. A partir disso, foi analisada a importância de cada *feature* para a realização da análise preditiva. Esse exame permitiu discutir e delimitar como e quais fatores psicológicos, econômicos e institucionais podem estar influenciando no processo decisório judicial.

Palavras-chave: análise preditiva; *machine learning*; *random forest*; Supremo Tribunal Federal; ações diretas de inconstitucionalidade; processo decisório judicial.

ABSTRACT

The goal of this paper is to initiate the construction of a predictive model of decisions of the Brazilian Supreme Court (Supremo Tribunal Federal) in “*Ações Diretas de Inconstitucionalidade*” using a machine learning algorithm. We consider that by testing and evaluating the elements of the judicial decision-making process that are relevant to perform a predictive analysis of the decisions, it is possible to narrow down how and which factors can influence the decision making process. To do so, first, we analyzed the literature that studies the influence of cognitive, economic and institutional factors on decisions. Then, we began the construction of the predictive model using a random forest algorithm. The algorithm was applied twice. The first time, the algorithm was input with the following features: (i) the report of the case; (ii) the year in which the decision was published; (iii) the judge rapporteur; and (iv) the claimant. The second time, in addition to these features, the vote of the judge rapporteur was also considered. In the first application, the algorithm predicted 85% of the cases correctly and, in the second time, 90% of the cases. From these results, the importance of each feature for the prediction was analyzed. This process allowed us to discuss and narrow down how and which cognitive, economic and institutional factors could be influencing the judicial decision-making process.

Keywords: predictive analysis; machine learning; random forest; Supremo Tribunal Federal; judicial decision-making process.

ÍNDICE DE GRÁFICOS

Gráfico 1 - Resultado da primeira aplicação do algoritmo	33
Gráfico 2 - Resultado da segunda aplicação do algoritmo	36

ÍNDICE DE TABELAS

Tabela 1 - Relação entre fatores, efeitos no processo decisório e features	24
Tabela 2 - Decisões inicialmente encontradas	26
Tabela 3 - Vinte features mais importantes na primeira aplicação do algoritmo	34
Tabela 4 - Vinte features menos importantes na primeira aplicação do algoritmo	35
Tabela 5 - Vinte features mais importantes na segunda aplicação do algoritmo	37
Tabela 6 - Vinte features menos importantes na segunda aplicação do algoritmo	38

SUMÁRIO

1	INTRODUÇÃO	6
2	FATORES CAPAZES DE INFLUIR EM DECISÕES JUDICIAIS	9
2.1	VIESES COGNITIVOS	9
2.2	FATORES ECONÔMICOS	12
2.3	FATORES INSTITUCIONAIS.....	14
3	ESTUDOS SEMELHANTES	18
4	CONSTRUÇÃO DE UM MODELO PREDITIVO	22
4.1	METODOLOGIA.....	22
4.1.1	<i>Escolha das ações e das features</i>	22
4.1.2	<i>Construção da base de dados</i>	25
4.1.3	<i>Preparação dos dados para processamento pelo algoritmo</i>	29
4.1.4	<i>Aplicação do algoritmo</i>	32
4.2	RESULTADOS	33
4.3	LIMITAÇÕES	38
5	DISCUSSÃO DOS RESULTADOS	39
6	CONCLUSÃO	43
7	REFERÊNCIAS BIBLIOGRÁFICAS	45

1 INTRODUÇÃO

As decisões tomadas por juízes dos mais diversos níveis hierárquicos provocam profundo impacto na sociedade. Contudo, apesar da relevância dessas decisões, não é tarefa trivial avaliá-las como boas ou ruins¹. A alternativa mais comum consiste em verificar aspectos jurídicos objetivos, como quais fatos e dispositivos foram analisados na decisão, mas ainda resta uma parte significativa do direito que está sujeita à interpretação do magistrado e que, para ser avaliada, depende de como cada indivíduo percebe o direito vigente². Isso acontece porque a lei não fornece uma resposta exata para todos os casos³. É necessário, então, que o magistrado escolha um método interpretativo para decidir como e quais leis aplicar ao caso em questão. Nesse exercício, há espaço para que um juiz seja influenciado por diversos fatores externos ao direito.

Um dos fatores que podem influenciar magistrados são os chamados vieses cognitivos. Estes consistem em “atalhos” que auxiliam os indivíduos a tomarem decisões mais rápidas. Apesar de serem essenciais para auxiliar os seres humanos a tomarem decisões no dia a dia, vieses cognitivos podem ser nocivos caso impactem indivíduos, especialmente magistrados, durante o seu exercício profissional sem que eles percebam. Alguns exemplos de vieses são o da ancoragem⁴, o *framing*⁵ e o *hindsight bias*⁶ que podem levar, respectivamente, juízes a (i) basear determinado valor de sua decisão em um número com qual tiveram contato, não necessariamente ligado ao processo, (ii) a enxergar a sua decisão somente por um ponto de vista de perda ou ganho e (iii) a superestimar a sua capacidade de prever que um evento aconteceria.

De acordo com Richard Posner, magistrados também podem ser influenciados em suas decisões por possibilidades de promoção na carreira, salário, ganhos de reputação e tempo dedicado ao trabalho e ao lazer⁷. Nesse sentido, juízes seriam, como quaisquer outros seres humanos, maximizadores da sua função utilidade, reagindo aos incentivos da carreira no judiciário.

¹ POSNER, Richard A., **How Judges Think**, [s.l.]: Harvard University Press, 2010. p.3.

² *Ibid.*

³ SCHAUER, Frederick, **Thinking like a lawyer**, [s.l.]: Harvard University Press, 2009.

⁴ GUTHRIE, Chris; RACHLINSKI, Jeffrey J; WISTRICH, Andrew J, Inside the judicial mind, **Cornell L. Rev.**, v. 86, p. 777, 2000.

⁵ *Ibid.*

⁶ *Ibid.*

⁷ POSNER, Richard A, What Do Judges Maximize?(The Same Thing Everybody Else Does), 3 SUP. CT, **Econ. Rev.**, v. 1, p. 3-4, 1993.

Por fim, também é possível considerar a influência de questões institucionais na tomada de decisão de juízes. Essa influência pode se originar (i) das regras institucionais do judiciário, que limitam o comportamento dos juízes ou (ii) da relação dos magistrados com outros atores políticos do país. Em relação ao primeiro sentido, é possível mencionar três modelos sobre comportamento judicial baseados no judiciário dos Estados Unidos: o “legal”, o “atitudinal” e o “estratégico”. Ribeiro e Arguelhes⁸ analisam como eles poderiam ser aplicados ao Brasil e mostram que fatores institucionais do STF podem afetar hipóteses desses modelos sobre como os magistrados se comportam. No último sentido, magistrados e, em especial, os ministros do Supremo Tribunal Federal, seriam provocados a realizar uma opção por exercer ou não um poder político, ou seja, a atuar ou não no campo político junto a outros atores. Arguelhes⁹ mostra que é possível identificar momentos em que ministros do Supremo Tribunal Federal optaram por até mesmo reduzir o poder político do Tribunal, apesar de condições externas que, consideradas isoladamente, poderiam indicar elevada atuação política da Corte.

Assim, é possível verificar que diversos fatores exógenos, além do próprio direito, são capazes de influir na tomada de decisão judicial. A identificação de quais são esses fatores é importante para a melhor compreensão da sociedade em relação ao processo decisório judicial. A construção de um modelo preditivo pode contribuir para ampliar essa compreensão na medida em que permite enumerar elementos das decisões capazes de explicar com maior precisão os seus resultados. Nesse sentido, ao indicar quais *elementos (features)* testados na análise preditiva são relevantes para determinar o resultado das decisões judiciais (*e.g.*, partes envolvidas, data, relator, tema, entre outros), o modelo preditivo permite delimitar os *fatores* do processo judicial (cognitivo, econômico ou institucional) que podem estar influenciando a tomada de decisão – possibilitando, assim, que essas influências sejam melhor estudadas¹⁰.

⁸ RIBEIRO, Leandro Molhano; ARGUELHES, Diego Werneck, Preferências, Estratégias e Motivações: Pressupostos institucionais de teorias sobre comportamento judicial e sua transposição para o caso brasileiro, **Revista Direito e Práxis**, v. 4, n. 7, p. 85–121, 2013.

⁹ ARGUELHES, Diego Werneck, Poder não é querer: preferências restritivas e redesenho institucional no Supremo Tribunal Federal pós-democratização, **Universitas Jus**, v. 25, n. 1, 2014.

¹⁰ A distinção terminológica utilizada neste parágrafo será adotada ao longo do artigo para fins didáticos. Considera-se que *elementos* ou *features* são os aspectos concretos e imediatos que formam os contextos em que as decisões são tomadas (como o relator, a data, o resultado da decisão, os Ministros que votaram, etc.). Já os *fatores* são os tipos de influências não imediatamente observáveis que podem explicar as decisões judiciais (vieses cognitivos específicos, expectativa de promoção dos juízes, interações institucionais, etc.). Os elementos de um certo contexto decisório podem determinar quais fatores melhor explicam as decisões judiciais, havendo entre eles, assim, uma relação de causa e efeito.

Como exemplo, é possível considerar um modelo preditivo que identifique que o elemento mais relevante para a realização de uma previsão seja qual foi o juiz envolvido na decisão do caso. A partir desse dado, é possível realizar estudos mais aprofundados, analisando os possíveis fatores que podem estar levando esse elemento a ser considerado como o mais relevante. Uma possível explicação para esse resultado seria a incidência de vieses cognitivos. Outra explicação viável seria a de que é possível associar a determinados juízes um padrão de votação orientado a uma atuação política do Judiciário, tornando previsíveis as suas decisões. Da mesma forma, caso o elemento mais relevante para uma análise preditiva sejam as partes do caso, seria possível considerar que os juízes estivessem sendo mais parciais para determinados requerentes. Contudo, o resultado poderia indicar também que determinadas partes possuem advogados melhores e, por isso, é comum que os pedidos delas sejam deferidos. Em qualquer caso, a análise preditiva tem o importante papel de delimitar quais são os possíveis elementos do processo decisório judicial que merecem ser mais bem estudados para que se compreenda a influência de diferentes fatores no processo decisório judicial.

Nesse sentido, a ideia que norteia este trabalho é a de que, testando-se e avaliando-se elementos do processo decisório que são importantes para a realização de uma análise preditiva da decisão judicial, é possível delimitar como e quais fatores podem estar influenciando no processo decisório. O objetivo deste trabalho é, portanto, de iniciar a construção de um modelo preditivo das decisões do Supremo Tribunal Federal (STF) utilizando um algoritmo de *machine learning*.

Outros trabalhos já utilizaram abordagens semelhantes. Nos Estados Unidos, Katz, Bommarito e Blackman¹¹ utilizaram um algoritmo de *random forest*¹² para prever as decisões da Suprema Corte dos Estados Unidos e como cada *justice* votaria nos casos. Stern e Zhu¹³ também obtiveram sucesso utilizando um algoritmo de *Support Vector Machine* (SVM) para prever as

¹¹ KATZ, Daniel Martin; BOMMARITO, Michael J.; BLACKMAN, Josh, A general approach for predicting the behavior of the Supreme Court of the United States, **PLOS ONE**, v. 12, n. 4, p. e0174698, 2017.

¹² Conforme será descrito na sessão 4.1.4, *random forest* é um algoritmo de *machine learning* composto por diferentes árvores de decisão. Essas árvores são construídas a partir da base de dados de treino do algoritmo. Ao receber a base de dados para teste, cada uma dessas árvores estima a probabilidade do caso em questão ser julgado como procedente ou improcedente.

¹³ STERN, HENRI, Zhu, Michael, Predicting Supreme Court Decisions Using Supervised Learning. Disponível em: <<http://docplayer.net/78135288-Predicting-supreme-court-decisions-using-supervised-learning.html>>, acesso em 6 nov., 2018.

decisões da Suprema Corte americana. Na Europa, Aletras et al¹⁴ utilizaram um algoritmo de SVM para prever o resultado de decisões da Corte Europeia de Direitos Humanos.

Para este trabalho, optou-se por iniciar a construção de um modelo preditivo utilizando um algoritmo de *random forest* para prever decisões do Supremo Tribunal Federal em Ações Diretas de Inconstitucionalidade. A opção pela utilização de um algoritmo de *machine learning* se deu por esse ser um método que permite realizar a previsão não só por dados objetivos sobre o processo decisório, mas também por textos envolvidos no julgamento, como o relatório da decisão. Optou-se pelo estudo do STF pois as suas decisões são de extrema relevância, sendo o Tribunal mais importante do país e possuindo a palavra final sobre a compatibilidade de atos normativos com a Constituição de 1988. Além disso, conforme indica Posner¹⁵, é relevante estudar o comportamento judicial a partir da atuação de ministros de uma Suprema Corte pois estes magistrados possuem mais liberdade para decidir do que outros, já que possuem a palavra final dentro da hierarquia do judiciário.

Este trabalho está dividido em outras cinco partes. A segunda parte trata, de forma mais aprofundada, dos aspectos que podem influenciar o processo decisório de juízes. A terceira descreve os estudos semelhantes a este já realizados. Em seguida, a quarta parte trata da metodologia do trabalho, das limitações da pesquisa e dos resultados. A quinta parte consiste em uma discussão sobre os resultados da pesquisa. Por fim, a sexta conclui o trabalho.

2 FATORES CAPAZES DE INFLUIR EM DECISÕES JUDICIAIS

2.1 VIESES COGNITIVOS

Como visto acima, ao apontar elementos de decisões que são relevantes para determinar o resultado de um caso, este trabalho pode contribuir para o estudo sobre fatores que influem o processo decisório de tribunais, ao exemplo dos vieses cognitivos. Há extensa literatura sobre a influência de vieses cognitivos nas decisões de juízes, o que leva a crer que este é um aspecto muito importante para a compreensão sobre como são realizadas decisões judiciais.

¹⁴ ALETRAS, Nikolaos *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.

¹⁵ POSNER, **How Judges Think**. P.14, *op. cit.*

A respeito desses vieses, Daniel Kahneman sugere que a forma de pensar dos indivíduos pode ser dividida em dois sistemas: o sistema 1 e o sistema 2¹⁶. Com o sistema 1, indivíduos sofreriam maior influência de vieses cognitivos e, assim, tomariam decisões mais rápidas. Esses vieses atuariam como atalhos que permitiriam que um indivíduo avançasse por etapas de seu julgamento sem considerar todos os elementos envolvidos. Com o sistema 2, por outro lado, indivíduos tomariam decisões de forma mais deliberada, dedicando mais tempo e atenção às etapas de sua decisão. Ambos os sistemas são importantes para que um indivíduo realize suas atividades diárias. Contudo, o sistema 1 pode se apresentar como um problema caso indivíduos não notem que estão tomando decisões importantes com base nele.

Estudos como esse assumem especial relevo na análise das decisões dos juízes, devido ao impacto delas na vida das pessoas. Um primeiro exemplo de viés que pode incidir na decisão judicial é o chamado viés da ancoragem, mencionado na introdução. Ele indica que juízes, ao determinar valores numéricos em suas decisões, como multas ou indenizações, poderiam ser influenciados por outros números com os quais eles tiveram contato em um período próximo do julgamento, mesmo que esses valores não tivessem relação com o julgamento em si. Dessa forma, o juiz, não intencionalmente, determinaria um valor no julgamento que seria próximo ao valor que ele teve contato¹⁷. No Brasil, Leal e Ribeiro¹⁸ verificaram a possibilidade de o valor de indenização pedido pelas partes em juizados especiais provocarem um viés de ancoragem nos juízes. Contudo, não encontraram uma correlação forte o suficiente que indicasse a influência do viés da ancoragem na decisão.

O viés de “*framing*”, por sua vez, afeta a maneira como juízes classificam o resultado de um caso. Ele faria com que magistrados percebessem o resultado de sua decisão como uma situação binária de perdas ou ganhos¹⁹. Essas perdas ou ganhos seriam determinados pelo *status quo* e afetariam a forma como magistrados consideram riscos²⁰, já que, de acordo com a teoria, indivíduos são propensos a riscos em situações de baixa probabilidade de ganho e alta probabilidade de perda, ao mesmo tempo em que são avessos ao risco em situações de alta probabilidade de ganho e baixa

¹⁶ KAHNEMAN, Daniel, **Thinking, fast and slow**, [s.l.]: Farrar, Straus and Giroux New York, 2011.

¹⁷ GUTHRIE; RACHLINSKI; WISTRICH, Inside the judicial mind, *op. cit.*

¹⁸ LEAL, Fernando; RIBEIRO, Leandro Molhano, Heurística de ancoragem e fixação de danos morais em Juizados Especiais Cíveis no Rio de Janeiro: uma nova análise, **Revista Brasileira de Políticas Públicas**, v. 8, n. 2, 2018.

¹⁹ GUTHRIE; RACHLINSKI; WISTRICH, Inside the judicial mind, *op. cit.*

²⁰ GUTHRIE, Chris, Prospect theory, risk preference, and the law, **Nw. UL Rev.**, v. 97, p. 1115, 2002.

probabilidade de perda²¹. Esse viés afetaria o modo como o juiz observa o caso em questão e até mesmo como o juiz aconselha as partes a fechar um acordo²².

Um juiz também pode superestimar a sua própria capacidade de prever o acontecimento de eventos que já aconteceram. É o chamado “*hindsight bias*”. Magistrados sofreriam esse viés pois utilizariam a informação nova, de que um evento de fato aconteceu, para determinar como estimariam a probabilidade do acontecimento desse evento no passado²³.

Até mesmo o humor dos magistrados poderia influenciar as suas decisões (“*mood congruency effect*”)²⁴. Nesse sentido, Danziger et al²⁵ mostram que juízes podem tomar decisões mais favoráveis à prisioneiros que pedem liberdade condicional quando realizam a decisão em um momento distante de uma refeição.

A literatura também menciona o fenômeno de “*cultural cognition*”, no qual os julgamentos de magistrados seriam afetados por características de sua personalidade²⁶. Elementos como posição política e religião poderiam levar juízes a perceber os mesmos fatos de forma diferente. É possível também que juízes sejam afetados por preconceitos propagados na sociedade e internalizados em sua personalidade. Mesmo que de forma inconsciente, juízes podem alterar a forma como tratam outros magistrados em órgãos colegiados, as partes e os advogados de um caso. É o caso do preconceito racial²⁷, por exemplo, que pode fazer com que magistrados profiram decisões mais rígidas para indivíduos negros do que brancos, mesmo sem essa intenção explícita. É possível mencionar também o preconceito de gênero. Arguelhes, Gomes e Nogueira²⁸ observaram que, no Supremo Tribunal Federal, “relatoras femininas tendem a atrair 1.2 vezes mais

²¹ KAHNEMAN, **Thinking, fast and slow**, *op. cit.*

²² GUTHRIE; RACHLINSKI; WISTRICH, Inside the judicial mind, *op. cit.*

²³ *Ibid.*

²⁴ FEIGENSON, Neal, Emotional influences on judgments of legal blame: How they happen, whether they should, and what to do about it, *in: Emotion and the Law*, [s.l.]: Springer, 2009, p. 45–96.

²⁵ DANZIGER, Shai; LEVAV, Jonathan; AVNAIM-PESSO, Liora, Extraneous factors in judicial decisions, **Proceedings of the National Academy of Sciences**, v. 108, n. 17, p. 6889–6892, 2011.

²⁶ BENFORADO, Adam, **Unfair: The new science of criminal injustice**, [s.l.]: Broadway Books, 2015.

²⁷ RACHLINSKI, Jeffrey J; GUTHRIE, Chris; WISTRICH, Andrew J, Inside the bankruptcy judge’s mind, **BUL Rev.**, v. 86, p. 1227, 2006.

²⁸ ARGUELHES, Diego Werneck; GOMES, Juliana Cesário Alvim; NOGUEIRA, Rafaela, Gênero e comportamento judicial no Supremo Tribunal Federal: os ministros confiam menos em relatoras mulheres?, **Revista Brasileira de Políticas Públicas**, v. 8, n. 2, 2018. P. 24.

votos vencidos do que os relatores masculinos”. No tribunal pleno, esse valor aumentaria para 1.37 vezes²⁹.

É possível notar, portanto, que vieses cognitivos podem afetar amplamente o modo como juízes decidem, especialmente à luz de elementos específicos do processo judicial. No entanto, o entendimento sobre como esses vieses afetam os julgamentos na prática é muito incipiente no Brasil. Iniciar a construção de um modelo preditivo capaz de mapear os elementos relevantes para a previsão de decisões do STF é, portanto, um ponto de partida importante.

2.2 FATORES ECONÔMICOS

Além dos estudos sobre vieses cognitivos no comportamento de juízes, também é possível observar trabalhos que modelam o comportamento de juízes como seres racionais, agindo de acordo com os incentivos que lhes são apresentados. Richard Posner³⁰ propõe um modelo como esse. De acordo com o autor, o sistema judicial³¹ busca insular, na maior extensão possível, os juízes de incentivos econômicos. Para desenvolver o seu modelo, então, o autor busca responder à questão: juízes atuam de forma racional ou o sistema impede que eles atuem de forma a responder a qualquer incentivo? Para isso, Posner parte da premissa de que juízes não exercem seu trabalho pois buscam “mudar o mundo” e perseguir visões iluminadas, tratando-os como “pessoas comuns”. O autor argumenta que o comportamento de juízes pode ser visto, em parte, como a composição de três atores racionais maximizadores: a organização sem fins lucrativos, o eleitor e o espectador teatral.

Para comparar as organizações sem fins lucrativos aos juízes, Posner se baseia na teoria de Henry Hansmann³². De acordo com o autor, a sociedade como um todo (principal) não consegue

²⁹ Os pesquisadores também buscaram identificar se os ministros tenderiam a realizar mais pedidos de vista em processos nos quais as relatoras eram mulheres, contudo, não obtiveram um resultado estatisticamente significativo.

³⁰ POSNER, What Do Judges Maximize?(The Same Thing Everybody Else Does), 3 SUP. CT.

³¹ O autor refere-se ao sistema judicial Americano, mas também é possível notar a mesma situação no Brasil, já que, entre outros, juízes brasileiros possuem estabilidade, recebem um salário alto, se comparado com o salário mínimo do país, e não podem ter esse salário reduzido.

³² Posner descreve a teoria de Henry Hansmann da seguinte forma: as organizações sem fins lucrativos existem para melhorar o alinhamento de incentivos quando o principal não consegue observar o resultado do “produto”. Por exemplo, um doador de dinheiro para uma ONG que alimenta crianças na África não conseguirá observar (ou o fará a um custo muito alto) se as crianças estão de fato recebendo os alimentos. Uma empresa privada teria incentivos para diminuir ao máximo os custos, inclusive deixando de alimentar as crianças, pois seus acionistas são reclamantes

medir o resultado esperado da ação dos juizes (agentes), que seria a “justiça”. Por isso, recorreria a uma alternativa que não recebe ganhos econômicos por realizar a atividade, como é o caso do juiz (cuja remuneração é invariável) e da organização sem fins lucrativos em outras situações. No caso do eleitor, Posner afirma que o juiz, assim como um eleitor, obtém prazer em votar como uma forma de expressar seus pontos de vista. Além disso, o voto também representaria um poder que gera deferência e respeito, assim como acontece com as decisões dos magistrados. Por fim, em relação aos espectadores de peças dramáticas, Posner mostra que, assim como estes (que obtêm prazer de imergir no drama e discutir com seus colegas experientes no assunto), os juizes participam do ato de julgar não apenas com suas preferências pessoais, mas com suas bagagens culturais e acadêmicas, a fim de dar uma opinião desinteressada.

Contudo, Posner afirma que essas três analogias não são suficientes para descrever o comportamento de juizes e que ainda restam incentivos para outros tipos ganhos, como menos trabalho, mais prestígio e ver suas decisões serem mantidas nos tribunais superiores. Ele busca, então, desenvolver um modelo que inclua também esses incentivos, e propõe a seguinte função utilidade: $U = U(t_j, t_l, I, R, O)$, onde t_j é o tempo dedicado a julgar, t_l é o tempo dedicado ao lazer, I é a renda³³, R é a reputação e O são as outras fontes de utilidade discutidas no artigo, como prestígio, popularidade e importância de evitar reversões de suas decisões. Além disso, considerando apenas variáveis mais responsivas ao tempo dedicado à atividade de julgar (reputação, por exemplo, não costuma ter grande correlação com o tempo gasto com a redação dos votos), o autor propõe também a seguinte função: $U = U(I_f, I_v(t_v), t_j, t_l)$, onde I_f é o salário fixo e I_v é o salário variável dependente do tempo (t_v) dedicado aos trabalhos paralelos.

Confirmando alguns dos fatores indicados por Posner em seu modelo é possível citar o trabalho de Sisk, Heise e Morris³⁴. Esse estudo se baseou no período em que estava vigente, nos Estados Unidos, um guia de elaboração de sentenças criminais pautado nas infrações e nas características dos crimes e dos criminosos, o “Federal Sentencing Guidelines”. Durante um período de quase um ano, centenas de juizes federais decidiram sobre a constitucionalidade desse guia. Dado que a questão jurídica era essencialmente a mesma, os autores verificaram se, e como,

residuais dos lucros. Nesse sentido, ver: HANSMANN, Henry. “The role of nonprofit enterprises”. *Yale Law Journal*, vol. 89, p. 835, 1980.

³³ O que inclui aquela renda percebida de trabalhos paralelos, como magistério.

³⁴ SISK, Gregory C; HEISE, Michael; MORRISS, Andrew P, Charting the influences on the judicial mind: An empirical study of judicial reasoning, *NYUL rev.*, v. 73, p. 1377, 1998.

influências externas afetavam os julgamentos. A pesquisa confirmou a influência de fatores sociais nas decisões, como a experiência profissional prévia dos juízes, mas também verificou empiricamente algumas das hipóteses de Posner. Encontrou, por exemplo, forte correlação entre o quão perto os juízes estão de serem promovidos e a propensão dos mesmos a declarar a constitucionalidade do guia de sentenças. Os pesquisadores também afirmam ter observado indícios de que há prazer intrínseco dos magistrados em escrever seus votos.

O modelo de Posner, e a pesquisa confirmando alguns de seus resultados, exercem grande contribuição por apresentar juízes como pessoas de carne e osso, que pautam a sua atuação em incentivos e não somente em um desejo de mudar o mundo com a sua atuação. Se o direito não possui respostas claras para todos os casos e juízes são guiados por incentivos como quaisquer outros indivíduos, essas pesquisas reforçam a necessidade de ser estudada ainda mais a fundo a forma como juízes decidem.

É importante notar também que, ao contrário do que possa parecer de início, o modelo de Posner não é totalmente incompatível com as teorias indicadas sobre a influência de vieses cognitivos. Juízes podem ser guiados, racionalmente, por incentivos de renda, reputação, lazer e prestígio, ao mesmo tempo em que são influenciados em suas decisões, sem perceber, por vieses cognitivos que afetam a forma como esses atores entendem os processos que estão decidindo e o direito vigente.

2.3 FATORES INSTITUCIONAIS

Além dos incentivos econômicos propostos por Posner, pode-se tratar, no Brasil, de fatores institucionais. Nesse sentido, é possível considerar que a atuação de juízes pode ser pautada: (i) por regras institucionais de funcionamento do Judiciário, que podem limitar ou não o comportamento de juízes; e (ii) pela relação entre o Judiciário e outros poderes da República, de forma que magistrados podem ser levados a escolher entre exercer um poder político ou não³⁵.

Para analisar aspectos institucionais de funcionamento do Supremo Tribunal Federal, Ribeiro e Arguelles partem de três principais modelos explicativos de comportamento judicial

³⁵ Neste trabalho a posição ideológica de um magistrado não está sendo considerada como um fator institucional de influência no processo decisório. Considera-se que a posição política de um magistrado pode alterar a sua forma de ver o mundo, atuando como um viés cognitivo em sua tomada de decisão.

identificados na literatura norte-americana: o modelo “legal”, o “atitudinal” e o “estratégico”³⁶. O modelo legal³⁷ descreve a tomada de decisão judicial como um exercício exclusivo de aplicação do direito vigente ao caso em questão, sem que ocorra a influência de outros fatores, como a opinião pessoal do juiz.

O desenvolvimento do modelo atitudinal, descrito pelos autores com base na proposta de Segal e Spaeth³⁸, seria uma resposta ao modelo legal. De acordo com o modelo atitudinal, o direito teria “necessariamente um caráter vago e variável”³⁹, sendo “inerentemente incerto e indefinido em seus conceitos, regras e princípios”⁴⁰. Dessa forma, juízes teriam um grau de liberdade durante a sua tomada de decisão. Essa liberdade seria utilizada por magistrados para decidir casos de modo a colher os “maiores benefícios, tendo em vista seus próprios interesses”, as chamadas “*policies*”⁴¹. Nesse contexto, apesar de serem limitados pelo contexto judicial em que estivessem julgando e pelos fatos do caso, juízes utilizariam o direito somente como uma forma de racionalização da sua tentativa de maximização de interesses pessoais.

Por fim, o modelo estratégico “assume a premissa de que os juízes são motivados pela maximização de suas preferências por *policies*, mas inclui no modelo restrições a esse processo maximizador”⁴². Algumas dessas restrições seriam (i) a ação de outros atores, que poderiam, por exemplo, retaliar contra o tribunal, e (ii) a influência de um próprio juiz dentro do tribunal em relação a outro, restringindo o comportamento do seu colega.

Esses três modelos de comportamento judicial, apesar de muito conhecidos, também sofreram críticas. Dentre essas críticas, é importante destacar as relacionadas ao papel do direito nas decisões judiciais. Bayley e Maltzman⁴³ mostram que uma decisão judicial que coincide com

³⁶ RIBEIRO; ARGUELHES, Preferências, Estratégias e Motivações: Pressupostos institucionais de teorias sobre comportamento judicial e sua transposição para o caso brasileiro, *op. cit.*

³⁷ Para descrever esse modelo, os autores se baseiam na descrição fornecida por Segal e Spaeth em: SEGAL, Jeffrey; SPAETH, Harold. (2002), *The Supreme Court and the Attitudinal Model Revisited*. Cambridge, Cambridge University Press. P. 48.

³⁸ SEGAL, Jeffrey A; SPAETH, Harold J, **The Supreme Court and the attitudinal model revisited**, [s.l.]: Cambridge University Press, 2002.

³⁹ RIBEIRO; ARGUELHES, Preferências, Estratégias e Motivações: Pressupostos institucionais de teorias sobre comportamento judicial e sua transposição para o caso brasileiro, *op. cit.*, p. 13.

⁴⁰ *Ibid.* P. 10

⁴¹ *Ibid.* P. 11.

⁴² *Ibid.* P. 16.

⁴³ BAILEY, Michael A; MALTZMAN, Forrest, Does legal doctrine matter? Unpacking law and policy preferences on the US Supreme Court, **American Political Science Review**, v. 102, n. 3, p. 369–384, 2008.

uma opinião política não está necessariamente seguindo essa opinião política, podendo ter chegado a uma resposta semelhante por meio do direito. Os pesquisadores também mostram empiricamente que elementos do direito influenciam as decisões de ministros da Suprema Corte americana, refutando a ideia de que o direito é utilizado somente como forma de racionalização das preferências pessoais dos ministros. Kritzer e Richards⁴⁴ também ressaltam a importância do direito na tomada de decisão dos ministros da Suprema Corte americana, sugerindo a existência de “regimes jurisprudenciais”. Nesses regimes, determinados casos julgados pela Corte estabeleceriam parâmetros e fatores relevantes de tomada de decisão para os casos seguintes a serem julgados.

Com base nesses três modelos norte-americanos e em críticas já realizadas a eles, Ribeiro e Arguelhes indicam lições importantes para a construção de modelos explicativos do comportamento de ministros no Supremo Tribunal Federal. De acordo com os autores, é importante definir a medida em que as decisões dos ministros são *sinceras*, expressando de fato a opinião dos tomadores de decisão, e não *estratégicas*, expressando uma resposta a outros atores políticos.

Os autores concluem que, no STF, os ministros não temeriam tentativas de reduzir o poder do Tribunal ou de alterar a sua composição, já que ações questionando essas medidas seriam eventualmente levadas ao próprio STF. Assim, os ministros teriam a palavra final sobre a compatibilidade, com o ordenamento jurídico brasileiro, de medidas que reduzissem o seu poder.

Ribeiro e Arguelhes também indicam que no STF “não há regras institucionalizadas que exijam ou encorajem a elaboração de um voto único por parte dos Ministros”⁴⁵, mas que ainda podem existir incentivos para os ministros “se comportarem estrategicamente quanto ao *resultado* (e não quanto à *fundamentação*) de seus votos”⁴⁶. Nesse sentido, um ministro poderia optar, por exemplo, por não ser a única opinião dissidente em um caso de muita atenção popular.

Os autores também ressaltam a importância de considerar se os casos decididos pelos ministros do STF são ou não difíceis, ou seja, se possuem ou não respostas jurídicas claras. De acordo com os autores, um modelo de comportamento dos ministros que considera todas as decisões proferidas no tribunal, ou seja, as decisões monocráticas, em turma e no Plenário, devem

⁴⁴ RICHARDS, Mark J; KRITZER, Herbert M. Jurisprudential regimes in Supreme Court decision making. **American Political Science Review**, v. 96, n. 2, p. 305–320, 2002.

⁴⁵ RIBEIRO; ARGUELHES, Preferências, Estratégias e Motivações: Pressupostos institucionais de teorias sobre comportamento judicial e sua transposição para o caso brasileiro, *op. cit.*, p. 27.

⁴⁶ *Ibid.* P. 28.

levar em conta que os ministros enfrentarão muitos casos fáceis entre todas essas ações. Ao julgar esse tipo de caso, um ministro não possuiria muita liberdade para alterar a forma de aplicar o direito, de modo que “estaremos nos movendo para longe de uma premissa institucional importante do modelo atitudinal, qual seja, a de que o ‘direito’ não exerce nenhuma pressão clara sobre a decisão dos Ministros”⁴⁷. Por outro lado, considerando apenas os casos do Plenário, seria possível considerar que os ministros encaram mais casos difíceis, em que o direito pode ter um papel mais relevante nas decisões.

Em artigo posterior, no qual examina a relação entre a atuação dos ministros do STF e outros atores políticos, Arguelhes⁴⁸ argumenta que o contexto político no qual o Tribunal é chamado a tomar decisões não são suficientes para determinar que a atuação dos ministros será política. De acordo com o autor a atuação política ou não do Tribunal seria resultado “das preferências majoritárias, em uma dada composição do tribunal, sobre quais são as formas adequadas e/ou oportunas de se pronunciar sobre questões políticas.”⁴⁹. Dessa forma, mesmo que os ministros estivessem inseridos em um contexto no qual seria esperada uma atuação amplamente política do Tribunal, seria possível que os ministros compartilhassem crenças sinceras de que o Tribunal não deveria agir dessa forma, guiando a atuação da Corte para fora do campo político.

Além desses trabalhos, também é relevante mencionar o de Virgílio Afonso da Silva⁵⁰ no tocante ao impacto dos ministros relatores no processo decisório do Supremo Tribunal Federal. Silva indica que o voto do relator é seguido na maior parte das decisões do STF, mas questiona a influência dos relatores nos casos mais polêmicos. De acordo com o autor, o fato de ministros já levarem seus votos prontos para os casos do Tribunal que recebem mais atenção do público “tem o potencial de relativizar o papel do relator como o ministro que pauta o debate, já que o debate é pautado pelos votos já escritos”⁵¹. Uma possível maneira de atenuar esse efeito seria com a divulgação do voto relator em momento anterior à sessão de julgamento.

⁴⁷ *Ibid.* P. 31.

⁴⁸ ARGUELHES, Poder não é querer: preferências restritivas e redesenho institucional no Supremo Tribunal Federal pós-democratização, *op. cit.*

⁴⁹ *Ibid.* P. 18.

⁵⁰ SILVA, Virgílio Afonso da, “UM VOTO QUALQUER”? O PAPEL DO MINISTRO RELATOR NA DELIBERAÇÃO NO SUPREMO TRIBUNAL FEDERAL, **REI - REVISTA ESTUDOS INSTITUCIONAIS**, v. 1, n. 1, p. 180–200, 2015.

⁵¹ *Ibid.* P. 12.

Os estudos descritos nesta seção mostram, portanto, a importância de se considerar a influência de fatores institucionais no comportamento de juízes. Esses fatores podem influenciar o “espaço de manobra” que juízes possuem para atuar e a relação entre o Judiciário e outros Poderes. Tanto esta seção quanto as anteriores identificaram fatores discutidos na literatura capazes de influenciar o processo decisório judicial. Mostrou-se que juízes (i) podem ser influenciados por vieses cognitivos durante a sua tomada de decisão; (ii) podem levar em conta aspectos pessoais econômicos, como tempo de férias e possibilidade de promoção; e (iii) podem atuar de acordo com as regras institucionais do próprio tribunal onde atuam e em resposta a outros atores políticos.

3 ESTUDOS SEMELHANTES

Conforme delimitado na introdução, o objetivo deste trabalho é contribuir para a identificação dos elementos cognitivos, econômicos e institucionais que influenciam o processo decisório judicial, por meio da construção de um modelo de previsão das decisões do STF em ADIs. É possível citar outras pesquisas, no exterior, que tiveram objetivos semelhantes⁵².

Katz, Bommarito e Blackman⁵³ utilizaram um algoritmo de “*Random Forest*” para prever as decisões da Suprema Corte dos Estados Unidos e como cada *justice* votaria nos casos. De acordo com os autores, o interesse em construir o modelo preditivo se deu devido a qualidade indeterminada de outros trabalhos preditivos da mesma Corte e ao fato de muitos trabalhos serem feitos de forma *ex post*, para determinar as causas que levaram a determinadas decisões, e não *ex ante*, isto é, para determinar o resultado de fato do caso. Para a pesquisa, os autores utilizaram os dados já catalogados na “*Supreme Court Database*”⁵⁴. Entre os dados escolhidos, as *features* do modelo, estavam o “*justice*”, “*term*”, corte de origem, requerente e “*issue*” discutido. O algoritmo obteve 70,2% de acerto no resultado dos casos utilizando esse método e 71,9% de acerto em relação

⁵² Optou-se por limitar a revisão da literatura estrangeira a casos em que a análise preditiva de decisões judiciais foi realizada por meio de técnicas de *machine learning*. Contudo, existem trabalhos que realizaram a construção de modelos preditivos de decisões judiciais utilizando outras técnicas. Como exemplo, é possível citar o Supreme Court Forecasting Project. Referência: **Supreme Court Forecasting 2002: Forecast**, disponível em: <<http://wusct.wustl.edu/forecast.php>>, acesso em: 27 fev. 2019.

⁵³ KATZ, Daniel Martin; BOMMARITO II, Michael J; BLACKMAN, Josh. A general approach for predicting the behavior of the Supreme Court of the United States. **PloS one**, v. 12, n. 4, p. e0174698, 2017.

⁵⁴ **The Supreme Court Database**, disponível em: <<http://supremecourtdatabase.org/>>, acesso em: 19 nov. 2018.

aos votos dos ministros. Stern e Zhu⁵⁵ também obtiveram sucesso utilizando dados da “*Supreme Court Database*”. Os autores utilizaram, para cada um dos *justices*, diferentes classificadores de *Support Vector Machine* (SVM). Com isso, foi realizada a previsão correta de 67% dos casos analisados.

Em “*Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective*”, Aletras et al⁵⁶ utilizaram classificadores de SVM para prever o resultado de decisões da Corte Europeia de Direitos Humanos. A criação da base de dados e a preparação desses dados foi realizada pelos próprios autores, a partir dos casos disponíveis na HUDOC⁵⁷, a plataforma que contém os casos da Corte. As *features* do modelo preditivo foram todas criadas a partir de sessões da própria decisão fornecida pelo Tribunal: “*procedure*”, “*circumstances*”, “*facts*”, “*relevant law*”, “*law*” e “*full case*”. Os autores obtiveram em média uma previsão correta em 75% dos casos. No trabalho também foi identificado que em determinados tipos de caso a categoria “*circumstances*”, em que são descritos os fatos do caso, era mais relevante para a previsão do resultado do julgamento do que a categoria “*law*”, em que eram indicados os argumentos legais do caso. De acordo com os autores, é possível que esses resultados indiquem que a tomada de decisão dos juízes da Corte é mais influenciada por aspectos do caso que não são relacionados ao direito.

Apesar de ainda não existirem trabalhos, como os citados acima, que aplicam técnicas de *machine learning* para a previsão de decisões do Supremo Tribunal Federal, é relevante mencionar os trabalhos que utilizaram a técnica estatística de *ideal point estimation* para prever decisões do STF. Simplificadamente, essa técnica consiste em identificar dimensões nas quais as opiniões dos ministros podem ser classificadas. Essas dimensões são determinadas pelo programa estatístico e, para identificar em que consistem elas, é necessário que os pesquisadores analisem os tipos de dados que foram classificados em cada uma. A classificação dos ministros em cada uma das dimensões é utilizada, então, para prever como o ministro irá votar em casos futuros.

⁵⁵ STERN, HENRI, Predicting Supreme Court Decisions Using Supervised Learning. Disponível em: <<http://docplayer.net/78135288-Predicting-supreme-court-decisions-using-supervised-learning.html>>, acesso em 6 nov., 2018.

⁵⁶ ALETRAS *et al*, Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective, *op. cit.*

⁵⁷ HUDOC - European Court of Human Rights, disponível em: <<https://hudoc.echr.coe.int/>>, acesso em: 20 nov. 2018.

Hartmann e Hudson⁵⁸ aplicaram o programa W-NOMINATE (*Weighted Nominal Three Step Estimation*) para estimar a posição de ministros do STF, utilizando os dados da base do projeto Supremo em Números da FGV⁵⁹. Os pesquisadores identificaram as seguintes dimensões da análise para os votos dos ministros: (i) direito penal, (ii) direito administrativo e tributário, (iii) federação e (iv) pensões públicas. Concluiu-se, então, que a opinião dos ministros se dividiria entre diferentes áreas em que possuiriam opiniões mais conservadoras ou não.

Mueller e Ferreira⁶⁰ utilizaram o mesmo programa para estimar o voto de cada ministro do STF em ADIs e no caso do Mensalão, além do parecer da PGR (Procuradoria Geral da República) e da AGU (Advocacia Geral da União) para esses mesmos casos. De acordo com os autores, a principal conclusão da pesquisa é de que as posições dos ministros podem ser determinadas por duas dimensões: (i) divergências em relação a políticas econômicas do Poder Executivo e (ii) divergências em relação ao papel dos estados e a sua autonomia para criar leis e estabelecer políticas públicas. Utilizando essas dimensões, os autores tiveram sucesso na previsão de 95% das ADIs analisadas.

Leoni e Ramos⁶¹ estimaram o voto de ministros do STF em ADIs e ADCs (Ações Declaratórias de Constitucionalidade). Para os casos envolvendo o Partido dos Trabalhadores (PT) e o Presidente da República, os autores também estimaram as suas posições. Utilizando apenas uma dimensão para a classificação dos ministros, a pesquisa resultou na previsão correta de 86% das decisões analisadas. De acordo com os autores, os resultados mostram que presidentes indicam ministros das mais diversas posições ideológicas, não se restringindo a ministros alinhados com a sua posição. Os mesmos também indicam que em grande parte dos casos, o governo, representado pela figura do Presidente da República, estava em lados opostos da dimensão em relação ao partido de oposição, o PT.

⁵⁸ HUDSON, Alexander; HARTMANN, Ivar Alberto, Can you bury ideology? An empirical analysis of the ideal points of the Ministers of Brazil's Supremo Tribunal Federal, **A&C-Revista de Direito Administrativo & Constitucional**, v. 17, n. 68, p. 43–59, 2017.

⁵⁹ **Supremo em Números**, disponível em: <<http://www.fgv.br/supremoemnumeros/>>, acesso em: 4 nov. 2018.

⁶⁰ MUELLER, Bernardo; FERREIRA, Pedro Fernando Almeida Nery, How Judges Think In The Brazilian Supreme Court: Estimating Ideal Points And Identifying Dimensions, *in*: , [s.l.]: ANPEC-Associação Nacional dos Centros de Pós-Graduação em Economia [Brazilian Association of Graduate Programs in Economics], 2014.

⁶¹ LEONI, Eduardo L.; RAMOS, Antonio P. Judicial preferences and judicial independence in new democracies: the case of the Brazilian Supreme Court. 2006. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.180.1470&rep=rep1&type=pdf>>

Silva⁶² também realizou análise semelhante, utilizando decisões liminares e definitivas do colegiado em ADIs, ADCs, ADPFs (Arguições de Descumprimento de Prefeito Fundamental) e ADOs (Ações Diretas de Inconstitucionalidade por Omissão). O autor concluiu que, quando há divergência entre os ministros, elas se relacionam à organização da federação e dos poderes em nível estadual. Além disso, o mesmo também observa na classificação dos ministros que “Gilmar tem procurado forçar o que o noticiário vem chamando de ‘maioria silenciosa’ (Rosa, Celso de Mello, Barroso, Fux, Carmen Lúcia e Lewandowski) a tomar posições mais próximas de suas preferências”⁶³.

Por fim, Lannes et al⁶⁴, realizam a *ideal point estimation* do voto de ministros do STF em casos não unânimes de ADIs. Em sentido contrário ao de Silva⁶⁵, os autores indicam que há, a partir do governo do presidente Lula, uma clara divergência entre os ministros indicados por Lula e os ministros indicados por outros presidentes.

Em conclusão, é possível dividir os trabalhos aqui descritos em dois grupos. O primeiro é composto por trabalhos realizados no exterior que utilizaram algoritmos de *machine learning* para prever as decisões de Cortes superiores. O segundo é formado por trabalhos que realizaram a previsão de decisões do STF, especialmente em ADIs, por meio da técnica de *ideal point estimation*. Esta pesquisa visa contribuir para esses dois grupos de trabalhos, a partir da utilização de um algoritmo de *random forest* para prever decisões do Supremo Tribunal Federal. Considera-se que a utilização do algoritmo de *machine learning* é uma evolução interessante em relação aos trabalhos estatísticos aqui mostrados pois permite que sejam utilizadas técnicas de processamento de linguagem natural para fornecer ao algoritmo não somente dados objetivos sobre o caso, mas também textos. Assim como os outros trabalhos, espera-se que a construção desse modelo preditivo contribua para a melhor compreensão sobre os fatores relevantes no processo decisório do STF e possíveis elementos influentes nas decisões dos ministros.

⁶² SILVA, Jeferson Mariano, MAPEANDO O SUPREMO: As posições dos ministros do STF na jurisdição constitucional (2012-2017) 1, **Novos Estudos**, n. 110, p. 34–54, 2018.

⁶³ *Ibid.* P. 16.

⁶⁴ LANNES, O; DESPOSATO, S; INGRAM, M, Judicial behavior in civil law systems: changing patterns on the Brazilian Supremo Tribunal Federal, *in*: **CICLO 2012 DO PROGRAMA DE SEMINÁRIOS CIEF-CERME-LAPCIPP-MESP**, 2, [s.l.: s.n.], 2012.

⁶⁵ É importante notar que o trabalho de Silva e o de Lannes et al trataram de períodos diferentes. O de Silva utilizou decisões de 2012 a 2017 enquanto o de Lannes et al utilizou decisões de 1989 a 2010. É possível que os resultados diferentes reflitam as mudanças na composição do STF.

4 CONSTRUÇÃO DE UM MODELO PREDITIVO

4.1 METODOLOGIA

4.1.1 Escolha das ações e das *features*

No STF, a principal via de controle constitucional consiste no controle concentrado de constitucionalidade, em que o Tribunal julga Ações Diretas de Inconstitucionalidade, Ações Declaratórias de Constitucionalidade (ADCs), Ações Diretas de Inconstitucionalidade por Omissão (ADOs) e Arguições de Descumprimento de Preceito Fundamental (ADPFs). Dentre essas ações, as ADIs são propostas em número mais expressivo no Tribunal. De acordo com o I Relatório do Supremo em Números⁶⁶, de 1998 a 2009 foram recebidas mais ADIs do que ADCs, ADOs e ADPFs no Supremo. Dessa forma, por funcionar como uma relevante porta de entrada para o controle de constitucionalidade de atos normativos no Supremo Tribunal Federal, as Ações Diretas de Inconstitucionalidade foram escolhidas como objeto da análise preditiva de decisões do STF nessa pesquisa.

Dentre as decisões em ADI, optou-se por somente utilizar decisões colegiadas, já que elas refletiriam a tomada de decisão do tribunal em conjunto e seriam mais representativas de casos difíceis, conforme argumentam Ribeiro e Arguelhes⁶⁷. Dentre as decisões colegiadas, somente as decisões sobre tutela de urgência e as decisões de mérito sobre constitucionalidade foram adicionadas à base de dados da pesquisa. Considerou-se que outras decisões, como aquelas oriundas de agravos regimentais e embargos de declaração, não refletiriam adequadamente a tomada de decisão do Tribunal em relação ao mérito das questões enfrentadas. Além disso, verificou-se que os relatórios dos agravos e embargos eram, em geral, curtos, o que não fornecia dados suficientes para o modelo.

⁶⁶ FALCÃO, Joaquim; CERDEIRA, Pablo de Camargo; ARGUELHES, Diego Werneck, I Relatório Supremo em Números- o Múltiplo Supremo, 2012.

⁶⁷ RIBEIRO; ARGUELHES, Preferências, Estratégias e Motivações: Pressupostos institucionais de teorias sobre comportamento judicial e sua transposição para o caso brasileiro, *op. cit.*

Cinco elementos (*features*) foram escolhidos como parâmetro para o algoritmo prever o resultado final da decisão do STF: (i) relatórios das decisões; (ii) ano em que a decisão foi publicada; (iii) ministro relator; (iv) requerente; e (v) voto do relator. Esses elementos foram escolhidos considerando possíveis efeitos no processo decisório. Os *relatórios das decisões* foram escolhidos considerando que a descrição dos argumentos das partes no caso poderia provocar um viés de *framing*. O grau de detalhamento e extensão do relatório também poderiam indicar um esforço feito pelo ministro relator para conhecer o caso, considerando questões de tempo e lazer. Além disso, o relatório também pode indicar informações relevantes sobre o direito em si envolvido no caso.

O ministro responsável por escrever o relatório, o *ministro relator* do caso, foi escolhido para que fosse analisada a relevância desse ator no processo decisório do STF, conforme debatido por Virgílio Afonso da Silva. É possível que em certos casos ministros exerçam uma menor influência nas decisões, porém o contrário também pode acontecer. Certos relatores podem ser mais propensos a determinados vieses cognitivos do que outros. Caso esses ministros reproduzam esses vieses frequentemente em seus relatórios, é possível que esse elemento contribua para o resultado do caso. Ademais, determinados ministros também podem ser mais propensos a refletir percepções econômicas de tempo e prestígio e percepções institucionais sobre o exercício de poder político pela Corte, em seus relatórios e votos.

O *ano da decisão*, por sua vez, pode estar relacionado à composição do tribunal. Assim, é possível considerar alterações na posição do Tribunal, em relação a, por exemplo, atividade política ou não, de acordo com as crenças dos ministros que compõe a Corte em determinado momento. Já o *voto do relator* foi escolhido pois acredita-se que eles possam provocar um viés de ancoragem nos outros ministros, por ser o primeiro a ser proferido. Além disso, uma alta correlação entre o resultado final das decisões e o voto do relator pode indicar que os ministros estão sobrecarregados, optando por analisar a fundo somente alguns casos. Nesse sentido, o voto do relator também poderia indicar uma influência institucional.

Por fim, o requerente foi escolhido considerando a sua relação com os casos trazidos. Determinados requerentes podem ir para a Corte frequentemente com casos semelhantes, o que geraria respostas semelhantes do Tribunal. Ademais, caso certos requerentes possuam alguma relação política com o Tribunal, é possível que o Tribunal reaja de forma política em relação aos mesmos, indicando uma influência institucional. Nesse contexto, é possível que a deferência profissional que ministros tenham em relação aos seus pares na Procuradoria-Geral da República

(também funcionários públicos da elite do Estado e supostamente “desinteressados”), produza, por exemplo, um provimento desproporcional a pedidos desse órgão em relação aos outros legitimados ativos.

A Tabela 1, a seguir, resume como as *features* escolhidas podem se relacionar com os fatores de influência no processo decisório:

Tabela 1 – Relação entre fatores, efeitos no processo decisório e *features*

Fator influente	Efeito no processo decisório	Elemento (<i>feature</i>) escolhido
Vieses cognitivos	<ul style="list-style-type: none"> - Ancoragem de resultados - Dificuldade de ver o caso a partir de diferentes perspectivas (<i>framing</i>) - Dificuldade de alinhar pensamentos com o de quem fez o ato normativo (<i>hindsight bias</i>) - Alteração no resultado devido ao contexto ou ao momento do dia (<i>mood congruency effect</i>) - Alteração na percepção de casos por características da personalidade do juiz (<i>cultural cognition</i>). 	<ul style="list-style-type: none"> - Relatório das decisões - Voto do relator - Ministro relator
Econômico	<ul style="list-style-type: none"> - Distribuição do tempo dedicado a julgar casos de acordo com o tempo de lazer - Tomada de decisão de acordo com considerações de prestígio, popularidade e promoção na carreira. 	<ul style="list-style-type: none"> - Relatório das decisões - Ministro relator
Institucional	<ul style="list-style-type: none"> - Atuação como resposta às regras do Tribunal - Votação estratégica de acordo com os votos dos outros ministros 	<ul style="list-style-type: none"> - Ano em que a decisão foi publicada - Requerente - Ministro relator

	<ul style="list-style-type: none"> - Maior “liberdade” para decisão em casos difíceis e menor em casos fáceis, nos quais o direito exerceria maior pressão. - Reação a outros atores políticos ou não, dependendo das crenças compartilhadas pelos ministros da Corte. 	<ul style="list-style-type: none"> - Voto do relator
--	--	---

Como se verá abaixo, em consideração a esses cinco elementos, decidiu-se aplicar o algoritmo duas vezes. Uma, considerando somente fatores anteriores à decisão do caso, o que não inclui o voto do relator. A segunda vez, considerando o voto do relator, para que seja possível observar se esse elemento causa alguma diferença no resultado.

4.1.2 Construção da base de dados

Para a construção da base de dados utilizada neste trabalho⁶⁸, foi realizada uma consulta por todas as ADIs constantes na base de dados Elasticsearch⁶⁹ e do Projeto Supremo em Números da FGV. Essa base contém o inteiro teor de milhares de processos do STF. A consulta resultou no inteiro teor de 11.030 decisões em ADIs. Essas decisões forem inseridas em uma *dataframe*⁷⁰ do Pandas⁷¹, um pacote da linguagem de programação Python. Dentre elas, foram filtradas 7.088 decisões que estavam classificadas como monocráticas, de forma que somente decisões colegiadas fossem utilizadas na pesquisa. Para identificar o resultado das decisões colegiadas restantes, foi realizada uma consulta na base de dados em formato SQL do Projeto Supremo em Números, que contém diversos dados sobre processos no STF. Essa base é única no país por conter dados que descrevem diversos elementos dos processos do Tribunal, como resultado do julgamento. Contudo, devido a uma divergência na identificação de processos entre as duas bases de dados, somente foi

⁶⁸ Após a apresentação deste trabalho, o código utilizado para construção da base de dados, preparação dos dados e aplicação do algoritmo será disponibilizado na plataforma Github.

⁶⁹ **Elasticsearch**, disponível em: <<https://www.elastic.co/products/elasticsearch>>, acesso em: 4 nov. 2018.

⁷⁰ O *dataframe* do Pandas funciona como uma espécie de tabela que pode ser manipulada de diferentes maneiras, facilitando a análise das informações.

⁷¹ **Python Data Analysis Library — pandas: Python Data Analysis Library**, disponível em: <<https://pandas.pydata.org/>>, acesso em: 4 nov. 2018.

possível obter o resultado de 323 decisões que estavam no Elasticsearch. Dentre essas decisões, a distribuição de resultados era a seguinte (Tabela 2):

Tabela 2 – Decisões inicialmente encontradas

Classificação da decisão	Total de decisões
Procedência total	133
Negativa de mérito	118
Procedência parcial	40
Negativa de admissão	15
Decisão sem análise de mérito	13
Admissão	4

Inicialmente, pretendia-se realizar uma separação aleatória de casos para serem utilizados na pesquisa. Contudo, devido ao baixo número de decisões encontradas entre o Elasticsearch e o SQL, optou-se por utilizar a maior parte das decisões encontradas. Dessas, foram retiradas da base as decisões de negativa de admissão, as sem análise de mérito e as unicamente de admissão, já que decisões de mérito seriam mais valiosas para compreender o funcionamento do Tribunal. Além disso, as decisões de procedência parcial e procedência total foram unificadas em uma só classificação, para que restasse somente uma classificação binária das decisões: procedência ou improcedência⁷². Assim, restaram 173 decisões de procedência e 118 de improcedência.

Em seguida, os relatórios das decisões restantes foram separados dos respectivos votos. Houve uma tentativa de separar esses relatórios de forma automatizada, porém, havia uma ausência de padrão na forma como os relatórios se iniciavam e terminavam, de forma que não foi possível automatizar esse processo. Dessa forma, optou-se por separar todos os relatórios manualmente. Foi selecionado todo o conteúdo de cada um dos relatórios⁷³, o que incluiu também informações contidas

⁷² Optou-se por unir decisões de procedência parcial com decisões de procedência total para que todas as ações em que houve qualquer tipo de procedência estivessem na mesma categoria. Apesar de criar um viés, essa escolha foi feita para permitir que houvesse uma classificação binária das decisões. Devido ao número reduzido de decisões na base de dados, não as classificar de forma binária iria tornar o modelo preditivo pouco eficaz.

⁷³ Por um problema na conversão das decisões no formato em que são disponibilizadas no site do STF para o formato em que são disponibilizadas na base Elasticsearch, a primeira e a última página dos relatórios estavam duplicadas em muitas das decisões analisadas. Ao selecionar o relatório, essas páginas duplicadas não foram incluídas.

no final das páginas de alguns processos sobre assinatura digital e número da folha⁷⁴. Nessa etapa, embargos e agravos também foram manualmente retirados da base. Assim, foram mantidas somente decisões de mérito e de medida cautelar em ADIs. Ao final, havia 160 decisões de procedência e 62 de improcedência. Como esse número estava muito desbalanceado, o que causaria um viés muito grande no modelo preditivo utilizado, optou-se por adicionar mais casos de improcedência à análise. Caso essa proporção não tivesse sido corrigida, o algoritmo poderia “aprender”, durante o seu treinamento, que a maioria das decisões era de procedência, optando, assim, por desconsiderar outros elementos relevantes do processo e prever a maior parte das decisões como procedente somente com base nesta informação⁷⁵.

Para corrigir essa proporção, criou-se uma função que sorteava decisões aleatórias dentre as decisões colegiadas em ADI identificadas na base de dados do Elasticsearch. Após sortear uma decisão, essa função verificava se a decisão já estava dentre as selecionadas da base ou se já havia sido sorteada antes e, em caso negativo para ambas as questões, retornava a decisão. Para cada decisão sorteada foi verificado se era uma decisão de improcedência. Em caso positivo, o relatório dessa decisão era separado. Esse procedimento foi realizado repetidamente até serem identificadas 100 decisões de improcedência. Essas decisões foram adicionadas à base junto às outras decisões identificadas anteriormente. Após essa etapa, o ano em que as decisões foram publicadas foi separado. A base Elasticsearch já havia indicado a data em que a decisão foi publicada, porém foi aplicada uma função que manteve somente o ano em que a decisão ocorreu. Considerou-se que como as datas são elementos muito específicos, existiriam muitos valores únicos que não trariam informações relevantes para o modelo. Assim, foi selecionado somente o ano de publicação.

Os dados relacionados aos outros três elementos foram coletados manualmente do inteiro teor das decisões⁷⁶. As informações foram preenchidas em uma planilha do Excel e foram

⁷⁴ Na base de dados Elasticsearch o inteiro teor dos processos é fornecido com um texto corrido. Com isso, estão incluídas nesse texto corrido todas as informações contidas no PDF do qual a decisão foi retirada do site do STF. Muitos dos PDFs contêm, no final da página uma informação em relação à assinatura digital do processo e ao número da folha. Assim, essas informações foram passadas para o texto da base Elasticsearch, mas não foram retiradas durante o processo de separação dos relatórios por aparecerem de forma muito repetida no meio do texto.

⁷⁵ Caso a base de dados utilizada na pesquisa contivesse todos as decisões em ADIs no STF não seria necessário realizar essa correção, já que a base refletiria a proporção de decisões realizada pelo tribunal. Contudo, como a base continha um número reduzido de casos, optou-se por manter uma proporção semelhante de decisões.

⁷⁶ Para obter os dados relativos aos outros parâmetros da pesquisa, tentou-se inicialmente buscá-los na base de dados SQL do Supremo em Números. Dentre eles, somente dois estavam na base: o ministro relator e o requerente. Contudo, houve novamente uma divergência na identificação de processos entre as duas bases de dados, de forma que foi possível obter informações em relação a poucos casos por meio da base do SQL. Em relação ao ministro relator havia, ainda, um problema adicional: a base do Supremo em Números se refere ao ministro atual do caso. Contudo, para um

posteriormente exportadas para o *dataframe*. Na planilha, os votos dos relatores foram classificados com os mesmos parâmetros de procedência e improcedência utilizados para classificar o resultado final das decisões. Para identificar os requerentes, inicialmente foram adicionados à planilha os nomes de todos os requerentes como estavam nas decisões. Devido ao tamanho da amostra, foi decidido generalizar a classificação dos requerentes de acordo com o rol do artigo 103 da Constituição⁷⁷ de legitimados ativos para impetrar ações diretas de inconstitucionalidade. Caso isso não fosse feito, haveria um número elevado de requerentes com pouca participação na amostra, o que não traria uma informação relevante para a análise do modelo⁷⁸. Assim, os requerentes foram classificados em: (i) “associação” (Art. 103, IX), (ii) “Procurador-Geral da República” (Art. 103, VI); (iii) “partido” (art. 103, VIII); (iv) “Conselho Federal da Ordem dos Advogados do Brasil” (Art. 103, VII); (v) “governador” (Art. 103, V); e (vi) “Mesa da Assembleia Legislativa” (Art. 103, IV). Outros legitimados ativos, como Presidente da República, não foram classificados por não estarem representados nos processos da amostra. A coleta desses dados permitiu também que a existência de agravos e embargos na amostra fosse checada novamente, de forma que foram retirados os casos encontrados que ainda se enquadravam nessas categorias.

Esse processo de obtenção de dados resultou em um *dataframe* com 313 decisões de mérito e de medida cautelar do Supremo Tribunal Federal em ações diretas de inconstitucionalidade. Dentre essas decisões, 158 tiveram como resultado a procedência do pedido e 155 a improcedência.

modelo preditivo, é muito mais relevante analisar o relator original do acórdão, porque se trata de uma informação prévia à decisão. Caso haja uma mudança de relator para o acórdão, por exemplo, essa mudança irá ocorrer após a decisão, de forma que não é desejável utilizá-la para construir um modelo preditivo desse mesmo julgamento.

⁷⁷ BRASIL. Constituição Federal da República Federativa do Brasil, Art. 103: “Podem propor a ação direta de inconstitucionalidade e a ação declaratória de constitucionalidade: I - o Presidente da República; II - a Mesa do Senado Federal; III - a Mesa da Câmara dos Deputados; IV - a Mesa de Assembleia Legislativa ou da Câmara Legislativa do Distrito Federal; V - o Governador de Estado ou do Distrito Federal; VI - o Procurador-Geral da República; VII - o Conselho Federal da Ordem dos Advogados do Brasil; VIII - partido político com representação no Congresso Nacional; IX - confederação sindical ou entidade de classe de âmbito nacional.”

⁷⁸ Na amostra somente haviam dois casos em que o processo apresentava mais de um requerente. Um desses casos (ADI 1649) foi apresentada por múltiplos partidos. Assim, o requerente foi classificado como um partido. Contudo, no caso da ADI 1163, os requerentes se enquadravam em duas categorias diferentes da classificação (Procurador Geral da República e Associação). Como era o único caso que apresentava essa questão, optou-se por retirar o processo da base.

4.1.3 Preparação dos dados para processamento pelo algoritmo

Algoritmos de *machine learning* não são capazes de receber, como dados de treinamento e teste, textos corridos ou colunas com valores referentes a diferentes parâmetros. Dessa forma, foi necessário preparar as colunas do *dataframe* referentes aos relatores, aos requerentes e aos relatórios para processamento pelo algoritmo.

Para processamento das informações relativas ao relator e ao requerente foi necessário transformar cada um dos valores únicos dessas duas categorias em uma nova coluna do *dataframe* que classificava as decisões com os valores 0 ou 1. Como exemplo, no caso da coluna referente ao ministro Gilmar Mendes, os processos dos quais ele foi relator foram marcados com 1 e o restante com 0. O mesmo foi feito para todos os outros relatores e requerentes. Essa transformação foi realizada utilizando o “OneHotEncoder”, função que faz parte do pacote Scikit-learn⁷⁹ do Python. Ao final desse processo, foram adicionadas 35 colunas ao *dataframe* (total de valores únicos nas colunas relator e requerente).

Para preparar os relatórios das decisões para processamento, o texto de cada um deles foi transformado em vetores. Para isso, primeiro, foi utilizada uma função⁸⁰ que (i) retirou *stopwords*⁸¹ contidas nos textos dos relatórios, (ii) retirou todos os números⁸² contidos no texto dos relatórios e (iii) transformou esses relatórios em uma lista, de forma que cada item da lista correspondia a um relatório. A partir dessa lista, foi criada uma série de vetores utilizando a função “Tf-idfVectorizer”⁸³ do pacote Scikit-learn do Python. Com essa função, o texto de cada relatório na lista foi dividido em

⁷⁹ PEDREGOSA, Fabian *et al*, Scikit-learn: Machine Learning in Python, **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

⁸⁰ Essa função foi criada pelo professor Renato Rocha Souza e apresentada na disciplina Introdução à Data Science, oferecida pelo mesmo na Escola de Matemática Aplicada da FGV no primeiro semestre de 2018. O código apresentado na disciplina está disponível em: <https://github.com/rsouza/FGV_Intro_DS/blob/master/notebooks/ML_SUP_Textual_Classification.ipynb>, acesso em: 4 nov. 2018.

⁸¹ *Stopwords* são palavras como “ou”, “na” e “de” que não contribuem para identificar o conteúdo do texto e, assim, costumam ser retiradas de textos para processamento de linguagem natural.

⁸² No momento em que as palavras foram separadas em unigramas e bigramas, os números contido nos relatórios, vindos de artigos ou outros valores informados pelos requerentes, perderam seu valor, já que era, somente números isolados. Assim, optou-se por retirar os números da análise.

⁸³ **sklearn.feature_extraction.text.TfidfVectorizer** — **scikit-learn 0.20.0 documentation**, disponível em: <http://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html>, acesso em: 4 nov. 2018.

unigramas e bigramas⁸⁴. Em seguida, foram criados vetores que representavam cada um desses termos, ou seja, cada um dos unigramas e bigramas. Cada vetor possui um valor referente a presença daquele termo em cada relatório da base. Esse valor é calculado da seguinte forma⁸⁵:

$$tf - idf(t, d) = tf(t, d) \times idf(t)$$

Nesse cálculo $tf(t, d)$ corresponde à frequência do termo no relatório em questão e $idf(t)$ corresponde à frequência do termo em todos os relatórios da base. O valor de $tf(t, d)$ é calculado de forma que⁸⁶: $tf(t, d) = 1 + \log(tf)$. Além disso, o valor de $idf(t)$ é calculado da seguinte forma, sendo n_d o número total de documentos e $df(d, t)$ o número total de documentos que contém o termo⁸⁷:

$$idf(t) = \log \frac{1 + n_d}{1 + df(d, t)} + 1$$

Com isso, caso um termo não conste no relatório em questão, o valor $tf - idf(t, d)$ correspondente a esse relatório será zero e, caso o termo conste no relatório em questão e em muitos outros relatórios, o seu valor $tf - idf(t, d)$ será reduzido. Esse cálculo, chamado *term frequency-inverse document frequency* (TF-IDF), é realizado dessa forma para que sejam identificados termos que constam em muitos textos da base de dados, já que, se um termo consta em muitos textos, ele não possui um valor muito grande para a análise do conteúdo de cada texto individualmente⁸⁸.

A função “Tf-idfVectorizer” foi aplicada com todos os seus parâmetros padrões, com exceção dos seguintes:

⁸⁴ Unigramas e bigramas são formas de representar termos. Unigramas consistem simplesmente em palavras isoladas. Bigramas consistem em duas palavras que são representadas em conjunto por aparecerem frequentemente na base de dados juntas. Dessa forma, a função Tf-idfVectorizer representa as palavras dos relatórios por unigramas, caso elas não apareçam com frequência acompanhadas de outras palavras, e por bigramas, caso elas palavras apareçam com frequência junto de outras palavras.

⁸⁵ **Feature extraction** — **scikit-learn 0.20.0 documentation**, disponível em: <http://scikit-learn.org/stable/modules/feature_extraction.html#text-feature-extraction>, acesso em: 4 nov. 2018.

⁸⁶ *Ibid.*

⁸⁷ *Ibid.*

⁸⁸ *Ibid.*

- (i) “Min_df”: esse parâmetro não inclui, na construção dos vetores, termos que possuam uma frequência menor que esse parâmetro⁸⁹, que, no caso, foi aplicado com valor 5.
- (ii) “Max_features”: esse parâmetro limita o número de termos dos textos a serem vetorizados, somente criando vetores dos termos com maior frequência⁹⁰. Esse parâmetro foi aplicado de forma que somente foram criados vetores com os 1000 termos que mais apareceram nos relatórios.
- (iii) “Sublinear_tf”: esse parâmetro indica a forma como o $tf(t, d)$ é calculado⁹¹. Esse parâmetro foi aplicado de forma que $tf(t, d)$ foi calculado da seguinte forma:
$$tf(t, d) = 1 + \log(tf)$$

A aplicação da função “Tf-idfVectorizer” retornou 1000 vetores, as chamadas *features* do modelo preditivo. Para reduzir esse número foi aplicada a função “VarianceThreshold” também do pacote Scikit-learn do Python. Essa função remove todas as *features* que tem uma variância menor que determinado parâmetro⁹². Com isso, é possível reduzir o número de *features* utilizadas no modelo preditivo, retirando as que, por possuir pouca variância, não contribuem muito para a distinção entre os relatórios. Para decidir o parâmetro de variância a ser utilizado, foram testados alguns valores. É importante notar que, devido ao tamanho reduzido da base de dados, o valor da variância teve que ser muito pequeno. Não foi possível encontrar *features* com variância maior que 0,002. Aplicando esse valor, restariam somente 3 vetores na base de dados. Assim, optou-se por aplicar o parâmetro de 0,0008 na função “VarianceThreshold”, de forma que foram mantidas 277 *features*. Estas foram adicionados ao *dataframe* originalmente utilizado.

Como o objetivo da pesquisa consistia em realizar duas aplicações do algoritmo, uma considerando o voto do relator como *feature* e outro não considerado, foi criado um novo *dataframe*, semelhante ao que estava sendo utilizando, porém sem a coluna correspondente aos votos dos relatores. Por fim, as decisões desses dois *dataframes* foram separadas de forma aleatória, utilizando

⁸⁹ *Ibid.*

⁹⁰ *Ibid.*

⁹¹ *Ibid.*

⁹² **Feature selection — scikit-learn 0.20.0 documentation**, disponível em: <http://scikit-learn.org/stable/modules/feature_selection.html#variance-threshold>, acesso em: 4 nov. 2018.

a função “train_test_split”⁹³ do pacote Scikit-learn do Python. Dessa forma, 30% das decisões de cada *dataframe* foram reservadas para teste e 70% para treinamento do algoritmo⁹⁴.

4.1.4 Aplicação do algoritmo

Para prever as decisões, foi aplicado um algoritmo de *random forest*. Simplificadamente, um algoritmo de *random forest* é composto por diferentes árvores de decisão que são construídas a partir da base de dados de treino do algoritmo. Ao receber a base de dados para teste, cada uma dessas árvores estima a probabilidade de o caso em questão ser julgado como procedente ou improcedente. É contabilizada, então, a média da probabilidade de cada resultado estimada por todas as árvores de decisão. A maior média é o resultado previsto pelo algoritmo para aquela decisão⁹⁵.

O algoritmo foi aplicado com todos os seus parâmetros padrões, com exceção dos seguintes:

- (i) “n_estimators”: corresponde ao número de árvores criadas⁹⁶. Esse parâmetro foi aplicado de forma que foram criadas 300 árvores de decisão.
- (ii) “n_jobs”: corresponde ao número de processadores que serão utilizados para aplicar o algoritmo⁹⁷. Esse parâmetro foi aplicado de forma que todos os processadores foram utilizados.
- (iii) “random_state”: caso esse parâmetro seja fixado em um número inteiro, garante-se que os resultados do algoritmo não serão alterados caso ele seja aplicado mais de uma vez sob os mesmos dados⁹⁸. Dessa forma, esse parâmetro foi fixado com o número inteiro 0.

⁹³ **sklearn.model_selection.train_test_split** — **scikit-learn 0.20.0 documentation**, disponível em: <http://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html>, acesso em: 4 nov. 2018.

⁹⁴ A divisão 70% para treinamento e 30% para teste é tipicamente utilizada pela comunidade científica para divisão de dados a serem processados por algoritmos.

⁹⁵ MÜLLER, Andreas C.; GUIDO, Sarah. **Introduction to Machine Learning with Python: A Guide for Data Scientists**. [s.l.]: O’Reilly Media, Inc., 2016. pp.84-86.

⁹⁶ **sklearn.ensemble.RandomForestClassifier** — **scikit-learn 0.20.0 documentation**, disponível em: <<http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>>, acesso em: 5 nov. 2018.

⁹⁷ **sklearn.ensemble.RandomForestClassifier** — **scikit-learn 0.20.0 documentation**, disponível em: <<http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>>, acesso em: 5 nov. 2018.

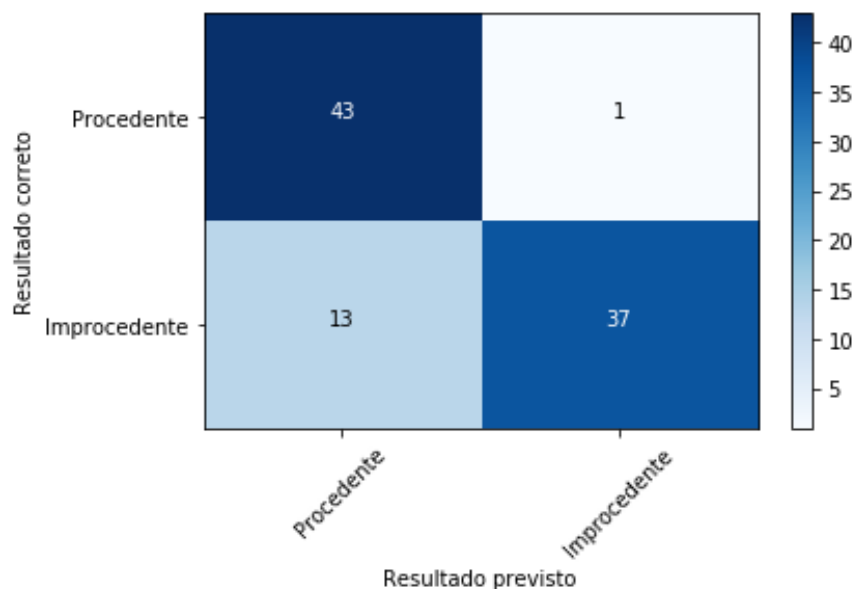
⁹⁸ MÜLLER, Andreas C.; GUIDO, Sarah. **Introduction to Machine Learning with Python: A Guide for Data Scientists**. [s.l.]: O’Reilly Media, Inc., 2016. pp.84-86.

4.2 RESULTADOS

O algoritmo foi aplicado duas vezes. Na primeira vez, não foi utilizado o voto do relator como parâmetro de treinamento e teste do algoritmo. Na segunda vez, esse parâmetro foi utilizado.

Na primeira aplicação, o algoritmo obteve sucesso em 85% dos resultados previstos. No Gráfico 1 abaixo é possível observar os casos procedentes e improcedentes previstos corretamente, além dos que foram previstos de maneira incorreta⁹⁹:

Gráfico 1 - Resultado da primeira aplicação do algoritmo



Conforme é possível observar, o algoritmo errou mais na previsão dos casos improcedentes, já que treze casos improcedentes foram classificados como procedentes, enquanto somente um caso procedente foi classificado como improcedente.

⁹⁹ Esse gráfico foi desenvolvido a partir de uma função criada pelo professor Renato Rocha Souza e apresentada na disciplina Introdução à Data Science, oferecida pelo mesmo na Escola de Matemática Aplicada da FGV no primeiro semestre de 2018. O código apresentado na disciplina está disponível em: <https://github.com/rsouza/FGV_Intro_DS/blob/master/notebooks/ML_SUP_Categorical_Classification.ipynb>, acesso em: 4 nov. 2018.

Considerando esse resultado, é possível também analisar outras métricas. O algoritmo obteve 77% de precisão¹⁰⁰ nos casos procedentes e 97% nos improcedentes. O valor de “recall”¹⁰¹ para os casos procedentes foi de 98% e para os improcedentes de 86%. Além disso, o “F1 score”¹⁰² foi de 86%, para os casos procedentes e 84% para os improcedentes.

Utilizando a função “feature_importances_” do pacote Scikit-learn do Python pode-se observar quais elementos foram mais importantes para o algoritmo determinar o resultado das decisões. Nesse caso, as vinte *features* mais importantes para o algoritmo foram as seguintes, conforme a Tabela 3:

Tabela 3 - Vinte *features* mais importantes na primeira aplicação do algoritmo

Colocação	Feature	Importância
1	Ano	0,10425
2	"documento assinado"	0,03024
3	"documento"	0,02336
4	"estadual"	0,01903
5	"df"	0,01801
6	"fls"	0,01746
7	Partido	0,01356
8	"federaldocumento"	0,01334
9	"partido"	0,01200
10	"federaldocumento assinado"	0,01073
11	"tribunal federaldocumento"	0,00959
12	"blicos"	0,00921
13	"lei estadual"	0,00904
14	"direito"	0,00830
15	"cias"	0,00772

¹⁰⁰ Precisão é uma medida calculada da seguinte forma: $P = \frac{T_p}{T_p + F_p}$. No caso de, por exemplo, decisões de procedência, T_p indica as decisões classificadas corretamente como de procedência e F_p indica as decisões classificadas incorretamente pelo algoritmo como de procedência, já que eram de improcedência. Fonte: **Precision-Recall — scikit-learn 0.20.2 documentation**, disponível em: <https://scikit-learn.org/stable/auto_examples/model_selection/plot_precision_recall.html>, acesso em: 28 fev. 2019.

¹⁰¹ “Recall” é uma medida calculada da seguinte forma: $R = \frac{T_p}{T_p + F_n}$. No caso de, por exemplo, decisões de procedência, T_p indica as decisões classificadas corretamente como de procedência e F_n indica as decisões classificadas incorretamente pelo algoritmo como de improcedência, já que eram de procedência. Fonte: *Ibid.*

¹⁰² “F1 score” é uma medida calculada da seguinte forma: $F1 = 2 \frac{P \times R}{P + R}$. Nesse cálculo, P indica o valor da precisão e R indica o valor do “recall”. Fonte: *Ibid.*

16	"executivo"	0,00767
17	"artigo"	0,00762
18	"estaduais"	0,00745
19	"ministro"	0,00731
20	"registro"	0,00715

O ano de publicação da decisão foi considerado como a *feature* de maior relevância para a realização da análise preditiva. Em seguida estão os partidos, requerentes considerados como sétima *feature* mais relevante, e expressões contidas nos relatórios. A palavra “partido” aparece entre essas expressões relevantes, em nona posição. Isso indica que para a realização da análise preditiva foi relevante considerar não só o partido como requerente, mas também a presença da palavra “partido” no relatório. Importante notar também que muitas das expressões consideradas como relevantes para o algoritmo são expressões como “documento assinado” e “tribunal federal documento” que constam no final da página de muitos processos e não tem relação com o caso em si. Algumas expressões estão abreviadas, mas podem ser inferidas, tais como “df” (Distrito Federal), “fls” (folhas) e “cias” (companhias).

Após essas 20 primeiras *features*, os próximos elementos, diferentes de expressões dos relatórios, a aparecer na classificação são a Procuradoria Geral da República, em 144º lugar, e o Min. Luis Roberto Barroso, em 171º. É possível observar que dentre as 20 *features* consideradas como *menos* relevantes para a análise preditiva estão diversos ministros relatores e, como requerente, a Mesa da Assembleia Legislativa, conforme indica a Tabela 4:

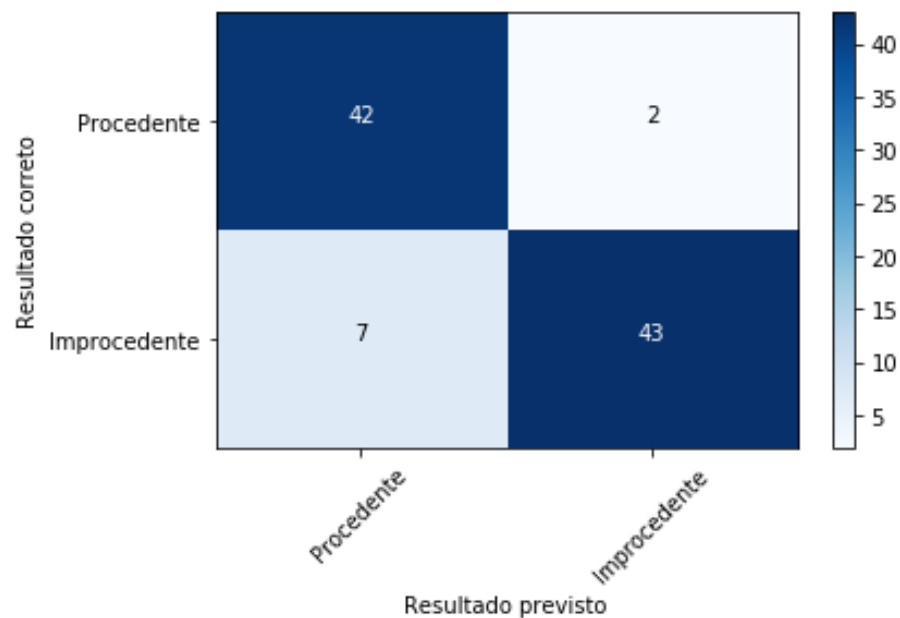
Tabela 4 - Vinte *features* menos importantes na primeira aplicação do algoritmo

Colocação	<i>Feature</i>	Importância
294	Min. Octavio Gallotti	0,00030
295	Min. Carlos Velloso	0,00030
296	"rond"	0,00029
297	"partidos"	0,00028
298	"estado goi"	0,00023
299	"rond nia"	0,00023
300	Min. Moreira Alves	0,00022
301	Min. Eros Grau	0,00020
302	Min. Ilmar Galvão	0,00016
303	Min. Alexandre De Moraes	0,00011

304	Min. Paulo Brossard	0,00009
305	Mesa Da Assembléia Legislativa	0,00008
306	"estado rond"	0,00006
307	Min. Nelson Jobim	0,00001
308	Min. Sydney Sanches	0,00000
309	Min. Cezar Peluso	0,00000
310	Min. Aldir Passarinho	0,00000
311	Min. Sepúlveda Pertence	0,00000
312	Min. Celso De Mello	0,00000
313	Min. Ellen Gracie	0,00000

Na segunda aplicação do algoritmo, considerando também o voto do relator no caso, o algoritmo obteve sucesso em 90% dos casos previstos. O Gráfico 2 abaixo indica a distribuição das previsões do algoritmo¹⁰³:

Gráfico 2 - Resultado da segunda aplicação do algoritmo



¹⁰³ Esse gráfico foi desenvolvido a partir de uma função criada pelo professor Renato Rocha Souza e apresentada na disciplina Introdução à Data Science, oferecida pelo mesmo na Escola de Matemática Aplicada da FGV no primeiro semestre de 2018. O código apresentado na disciplina está disponível em: <https://github.com/rsouza/FGV_Intro_DS/blob/master/notebooks/ML_SUP_Categorical_Classification.ipynb>, acesso em: 4 nov. 2018.

Conforme é possível observar, o algoritmo acertou mais casos utilizando a informação do voto do relator. Somente sete casos improcedentes foram estimados como procedentes, enquanto dois casos procedentes foram estimados como improcedentes.

Considerando esse resultado, o algoritmo obteve 86% de precisão nos casos procedentes e 96% nos improcedentes. O valor de “*recall*” para os casos procedentes foi de 95% e para os improcedentes de 86%. Além disso, o “*F1-score*” foi de 90%, para os casos procedentes e 91% para os improcedentes.

Utilizando a função “*feature_importances_*” do pacote Scikit-learn do Python é possível observar os 20 elementos mais importantes para o algoritmo realizar a previsão, conforme indica a Tabela 5:

Tabela 5 - Vinte *features* mais importantes na segunda aplicação do algoritmo

Colocação	<i>Feature</i>	Importância
1	Voto relator	0,14737
2	Ano	0,08619
3	"documento"	0,02229
4	"documento assinado"	0,01948
5	"fls"	0,01757
6	"df"	0,01508
7	"estadual"	0,01452
8	Partido	0,01189
9	"tribunal federaldocumento"	0,01011
10	"partido"	0,00980
11	"federaldocumento"	0,00897
12	"estaduais"	0,00862
13	"direito"	0,00754
14	"cf"	0,00751
15	"federaldocumento assinado"	0,00742
16	"lei estadual"	0,00740
17	"autor"	0,00718
18	"ia"	0,00657
19	"fl"	0,00655
20	"blicos"	0,00648

O voto do relator foi considerado como a *feature* mais importante para a realização da análise preditiva. O ano de publicação e os partidos mantiveram-se como elementos relevantes. Muitas das expressões consideradas como importantes na primeira aplicação também se destacaram na segunda aplicação. Contudo, é possível identificar também novas abreviações, como “cf”, que pode se referir à Constituição Federal.

Dentre as 20 *features* consideradas como *menos* relevantes para a análise preditiva estão diversos ministros relatores e o Conselho Federal da Ordem dos Advogados do Brasil, conforme indica a Tabela 6:

Tabela 6 - Vinte *features* menos importantes na segunda aplicação do algoritmo

Colocação	<i>Feature</i>	Importância
295	Min. Néri Da Silveira	0,00029
296	Min. Francisco Rezek	0,00028
297	Conselho Federal Da Ordem Dos Advogados Do Brasil	0,00028
298	"norte"	0,00025
299	"rond"	0,00023
300	"estado rond"	0,00021
301	Min. Carlos Velloso	0,00020
302	Min. Sepúlveda Pertence	0,00018
303	"rond nia"	0,00012
304	Min. Octavio Gallotti	0,00010
305	Min. Celso De Mello	0,00009
306	Min. Cezar Peluso	0,00006
307	Min. Eros Grau	0,00006
308	Min. Alexandre De Moraes	0,00005
309	Min. Moreira Alves	0,00000
310	Min. Sydney Sanches	0,00000
311	Min. Nelson Jobim	0,00000
312	Min. Paulo Brossard	0,00000
313	Min. Ellen Gracie	0,00000
314	Min. Aldir Passarinho	0,00000

4.3 LIMITAÇÕES

Como notado, este trabalho consiste em uma construção inicial de um modelo de previsão de decisões judiciais utilizando um algoritmo de *machine learning*. Existem aspectos na

metodologia utilizada nesta pesquisa que impõem limitações aos resultados do algoritmo descritos acima. Em primeiro lugar, o tamanho da base de dados, 313 decisões, é muito baixo em relação ao que costuma ser utilizado na área de *big data* para treinar e testar algoritmos preditivos. Esse problema ficou claro na análise da variância dos relatórios e pode estar afetando o resultado encontrado para as previsões.

O fato de a data utilizada como variável ser o ano em que a decisão foi publicada também é um limitador. É possível que a composição do Tribunal tenha mudado entre a data em que a decisão foi proferida e a data em que a decisão foi publicada. Além disso, a utilização do ano como parâmetro, em vez da data completa, pode também incluir diferentes composições do Tribunal.

Outro limitador do trabalho é o inteiro teor das decisões. Como é possível observar nas expressões dos relatórios que compõem as *features*, muitas palavras estão quebradas por espaços ou linhas de forma que elas não são identificadas por inteiro. Também poderia ser importante retirar dos relatórios informações que fazem parte das páginas dos processos, mas não do conteúdo dos relatórios, como as informações mencionadas em relação à assinatura digital do processo.

Para trabalhos futuros também é importante identificar outras *features* relevantes, como ministros presentes no julgamento e área do direito do qual a decisão faz parte¹⁰⁴. Parâmetros padrões das funções utilizadas também poderiam ser alterados no futuro para melhorar a previsão, além de serem testados outros algoritmos.

Muitos desses limitadores podem ser solucionados identificando formas de obter computacionalmente dados em relação às decisões. A obtenção manual de dados limita muito o que pode ser feito¹⁰⁵.

5 DISCUSSÃO DOS RESULTADOS

Apesar das limitações mencionadas, os resultados encontrados em ambas as aplicações do algoritmo não podem ser desprezados. Considerando os trabalhos estatísticos mencionados na

¹⁰⁴ Inicialmente tentou-se obter a área do direito correspondente às decisões utilizadas no trabalho da base SQL do Supremo em Números, porém houve um problema no cruzamento da identificação da base Elasticsearch e SQL de modo que muitas decisões não foram identificadas. Assim, optou-se por não utilizar essa informação.

¹⁰⁵ A base de dados SQL do Supremo em Número já contém muitas informações em relação aos processos e poderia ser utilizada. Ela poderia, assim, ser muito útil em trabalhos futuros, desde que fosse solucionado o problema relatado de relacionar a base de dados Elasticsearch com a base SQL.

seção de “Estudos Semelhantes”, a pesquisa obteve resultados próximos ao de Leoni e Ramos¹⁰⁶, que obtiveram 86% de acerto em sua previsão, e Mueller e Ferreira¹⁰⁷, que obtiveram 95% de acerto. É interessante também comparar os resultados dos modelos preditivos do Supremo Tribunal Federal com os realizados nos Estados Unidos. Enquanto os trabalhos de Katz et al e de Stern e Zhu obtiveram, respectivamente, 70,2% e 67,7% de acerto, os trabalhos relacionados ao STF tiveram uma taxa de acertos maior. Essa informação pode indicar que existem diferenças significativas entre o processo decisório de ambos os tribunais¹⁰⁸, conforme alertam Ribeiro e Arguelhes¹⁰⁹.

O número de acertos do algoritmo em ambas as aplicações, muito acima de 50%, indica que, comparativamente a uma pessoa que determinasse aleatoriamente o resultado das decisões, o algoritmo teve um grau de acerto alto. Além disso, considerado as outras métricas indicadas de precisão, “*recall*” e “*F1 score*”, o algoritmo obteve resultados entre 77% e 98%, o que também indica um desempenho relevante.

Nas aplicações do algoritmo, o ano em que a decisão foi publicada foi um dos fatores mais relevantes para a realização da análise preditiva¹¹⁰. Uma possível conclusão a que se pode chegar a partir desses dados é a de que quais são os ministros participantes do Tribunal em dado período é o elemento mais relevante para a análise, seja pela influência de fatores como as suas crenças pessoais ou pelo próprio modo como consideram o direito. O ano de publicação também pode indicar uma composição diferente em outros poderes, impactando o tipo de caso que chega no Tribunal ou a percepção dos membros do STF sobre o modo como devem atuar.

¹⁰⁶ LEONI; RAMOS, Judicial preferences and judicial independence in new democracies: the case of the Brazilian Supreme Court, *op. cit.*

¹⁰⁷ MUELLER; FERREIRA, How Judges Think In The Brazilian Supreme Court: Estimating Ideal Points And Identifying Dimensions, *op. cit.*

¹⁰⁸ É possível que uma diferença relevante seja o relator do processo. Nesta pesquisa, o relator cumpriu papel relevante entre as variáveis, sendo, por si só uma variável, e também sendo o responsável pelo relatório, outra variável do modelo. Contudo, nos Estados Unidos, o processo deliberativo da Suprema Corte assume a forma *per curiam*, diferentemente do STF (processo de voto *per seriatim*), de forma que os estudos lá realizados não contam com o voto do relator. Sobre essas diferenças, ver DA SILVA, V. A., Deciding without deliberating, **International Journal of Constitutional Law**, v. 11, n. 3, p. 557–584, 2013.

¹⁰⁹ RIBEIRO; ARGUELHES, Preferências, Estratégias e Motivações: Pressupostos institucionais de teorias sobre comportamento judicial e sua transposição para o caso brasileiro, *op. cit.*

¹¹⁰ Conforme relatado na metodologia, é possível que o ano seja um elemento relevante por se relacionar com a composição do Tribunal. Embora a escolha de um período de um ano possa trazer limitações à pesquisa, por ser possível haver composições diferentes no mesmo ano, o grau de relevância dessa variável para o modelo preditivo foi muito elevado em comparação as outras, indicando que essa variável deve estudada com mais profundidade.

É interessante comparar essas hipóteses com a colocação dos ministros relatores entre as *features* relevantes. Muitos ministros relatores se classificaram no final do *ranking*, com uma importância muito baixa para a análise. Contudo, observa-se que os últimos na colocação foram os ministros menos representados na tabela. Ministros como Dias Toffoli e Gilmar Mendes, com maior representação entre os casos, não estão entre as 20 *features* menos importantes para a análise. Dessa forma, nota-se que é necessário considerar não só a importância dos elementos, mas como ela se relaciona com a quantidade de dados na amostra¹¹¹.

Os partidos, como requerentes, e a presença da palavra “partido” nos relatórios também se mostraram relevantes para a análise, apesar de os partidos não serem os requerentes mais representados na base de dados¹¹². Essa observação pode indicar a influência de fatores institucionais nas decisões, caso as ações levadas por partidos tenham caráter político.

Para além disso, dentre as expressões relevantes para a realização da análise preditiva, pode-se observar “estadual”, “estaduais” e “lei estadual”. É possível que essas expressões sejam relevantes por estarem relacionadas a casos de conflito federativo e de debate em relação ao papel e autonomia dos estados. Os trabalhos de Hartmann e Hudson¹¹³, Mueller e Ferreira¹¹⁴ e Silva¹¹⁵ identificaram que essas eram importantes dimensões para determinar o voto de ministros e a relevância dessas expressões reforça a análise realizada pelos pesquisadores.

Também foram consideradas relevantes para a análise a presença no relatório de expressões que tipicamente se encontravam no final das páginas dos relatórios, como “documento assinado” e “tribunal federal documento”. O trecho abaixo exemplifica como são essas notas de pé de página¹¹⁶:

¹¹¹ O mesmo pode ser considerado para a Mesa da Assembleia Legislativa, que foi representada somente 4 casos na amostra, e o Conselho Federal da OAB, que foi representado em 16 casos. Ambos os requerentes foram considerados como de baixa importância para realização da análise preditiva, porém foram também pouco representados na amostra.

¹¹² Dos casos analisados, 49 possuíam partidos como requerentes. Os requerentes mais representados na amostra foram os governadores, com 109 casos.

¹¹³ HUDSON; HARTMANN, Can you bury ideology? An empirical analysis of the ideal points of the Ministers of Brazil's Supremo Tribunal Federal, *op. cit.*

¹¹⁴ MUELLER; FERREIRA, How Judges Think In The Brazilian Supreme Court: Estimating Ideal Points And Identifying Dimensions, *op. cit.*

¹¹⁵ SILVA, MAPEANDO O SUPREMO: As posições dos ministros do STF na jurisdição constitucional (2012-2017) 1.

¹¹⁶ STF, ADI 4617/DF, Relator Ministro Luiz Fux, julgado em 19/06/2013, *DJe* 12/02/2014.

“Supremo Tribunal Federal Documento assinado digitalmente conforme MP nº 2.200-2/2001 de 24/08/2001, que institui a Infraestrutura de Chaves Públicas Brasileira - ICP-Brasil. O documento pode ser acessado no endereço eletrônico <http://www.stf.jus.br/portal/autenticacao/> sob o número 4342060.\x0cSupremo Tribunal Federal\n\nInteiro Teor do Acórdão - Página 5 de 30\n\nRelatório\n\nADI 4617 / DF”

Apesar de não apresentarem nenhum conteúdo relevante sobre o caso em si, a presença de muitas dessas notas em um mesmo relatório pode indicar que esse relatório possui muitas páginas, já que cada uma dessas notas se encontra no final de uma página da decisão¹¹⁷. Dessa forma, é possível que tenha sido identificada uma relação entre o número de páginas dos relatórios e o resultado da decisão. Essa relação pode ser um resultado de incentivos econômicos dos ministros, conforme descrito por Posner¹¹⁸. Isso porque, o tempo dedicado para escrever o relatório e julgar a decisão seriam refletidos em seu resultado. É possível também que o tamanho dos relatórios reflita fatores institucionais. Conforme descrevem Ribeiro e Arguelhes¹¹⁹ o STF julga um grande número de processos, o que envolve casos fáceis e difíceis. Assim, o tribunal pode alocar mais tempo para escrever relatórios e votar casos difíceis, e menos para os casos fáceis. A presença das abreviações “fl”, de folha, e “fls”, de folhas, entre as 20 *features* mais relevantes reforça essa análise, na medida em que é possível que casos mais complexos, com um maior número de folhas, tenham mais páginas mencionadas no relatório.

Por fim, além dos elementos já mencionados, a segunda aplicação do algoritmo mostrou que o voto do relator é muito importante para a realização da análise preditiva. Existem diferentes fatores que podem estar contribuindo para esse resultado. É possível que o voto do relator possua um efeito de ancoragem nos outros ministros, os incentivando a votar da mesma forma. Além disso, o modo com o relator descreve o caso e vota também pode exercer uma influência nos outros ministros, através do viés de “*framing*”, de forma que os ministros tenham uma visão do caso

¹¹⁷ É relevante notar que essas notas de pé de página não estavam presentes em todos os relatórios analisados. Apesar de muitos relatórios possuírem essas notas, alguns não possuíam. É possível que isso tenha ocorrido devido ao modo como o inteiro teor dos acórdãos foi importado para a base de dados ou pelo fato de processos mais antigos não possuírem assinatura digital, não precisando, assim, dessa observação.

¹¹⁸ POSNER, What Do Judges Maximize?(The Same Thing Everybody Else Does), 3 SUP. CT, *op. cit.*.

¹¹⁹ RIBEIRO; ARGUELHES, Preferências, Estratégias e Motivações: Pressupostos institucionais de teorias sobre comportamento judicial e sua transposição para o caso brasileiro, *op. cit.*

semelhante a que foi posta pelo ministro relator. Também é possível que esse resultado reflita fatores institucionais. Se os ministros julgarem muitos casos, eles podem escolher determinados casos para os quais dedicar mais tempo, seguindo o voto do relator nos outros. Outra hipótese é que os relatores atuam de forma estratégica. Se o relator não sabe como os outros ministros vão votar, ele pode moderar o seu voto, incorporando considerações que ele acredita que os outros ministros concordem.

Uma última hipótese relativa à importância do voto do relator refere-se à possibilidade de que esse resultado retrate um consenso de opinião entre os ministros do Tribunal, de forma que a votação entre eles seja semelhante. Nesse sentido, o fato de o conteúdo dos relatórios e o voto do relator se mostrarem como *features* mais relevantes para a análise preditiva do que o ministro relator do caso pode indicar que a decisão final de um caso é mais influenciada pela opinião da Corte sobre o direito em questão, refletida no relatório e no voto do relator, do que por questões pessoais dos ministros. Caso contrário, o ministro relator por si só seria uma *feature* também muito relevante, já que cada ministro influenciaria a decisão de uma maneira diferente.

Em conclusão, os resultados da análise preditiva indicam que ministros do STF podem estar sofrendo a influência de fatores econômicos, psicológicos e institucionais. Sugeriu-se que a relevância do ano para a análise preditiva poderia estar relacionada à composição do Tribunal e à influência de fatores pessoais dos ministros responsáveis pela decisão. A relevância dos partidos, tanto como requerentes quanto como palavra mencionada no relatório, sugere a influência de fatores institucionais na tomada de decisão. Considerou-se que o número de páginas do relatório também poderia estar sendo relevante para a análise preditiva. Por fim, foram sugeridas diversas hipóteses psicológicas e institucionais que podem explicar a relevância do voto do relator para a realização da análise preditiva. Esses resultados, e suas explicações possíveis, oferecem um caminho para explorações mais aprofundadas em trabalhos futuros.

6 CONCLUSÃO

Este trabalho foi desenvolvido com base na ideia de que, testando-se e avaliando-se elementos do processo decisório que são importantes para a realização de uma análise preditiva da decisão, é possível delimitar como e quais fatores podem estar influenciando no processo decisório. Determinou-se, então, como objetivo do trabalho, iniciar a construção de um modelo de previsão

de decisões do Supremo Tribunal Federal em Ações Diretas de Inconstitucionalidade utilizando um algoritmo de *machine learning*.

A construção desse modelo preditivo considerou a influência, na tomada de decisão judicial, de três grupos de fatores: psicológico, econômico e institucional. Fatores psicológicos seriam definidos principalmente por vieses cognitivos, como ancoragem e *framing*. Fatores econômicos seriam relacionados a uma busca racional dos juízes por maximização da sua função utilidade, definida por elementos como tempo de lazer e prestígio. Por fim, fatores institucionais seriam relacionados às regras institucionais, que limitam o comportamento de juízes, e à relação de juízes com outros atores políticos.

Com base nesses fatores foram escolhidas cinco *features* para o modelo preditivo e, a partir delas, foi construída a base de dados e aplicado o algoritmo de *random forest*. Consideradas as limitações do trabalho, os resultados da previsão do algoritmo foram positivos. Na primeira vez em que foi aplicado, sem considerar o voto do relator, o algoritmo acertou 85% dos casos previstos. Na segunda aplicação, considerando o voto do relator, o algoritmo acertou 90% dos casos. Esses resultados se mostraram alinhados com outros trabalhos que utilizaram a técnica de *ideal point estimation* para prever o voto de ministros do STF.

Quatro *features* foram consideradas as mais relevantes para a realização da análise preditiva: (i) ano em que a decisão foi publicada; (ii) partidos, como requerentes e como palavra mencionada no relatório; (iii) palavras relacionadas ao número de páginas do relatório; e (iv) o voto do relator, na aplicação que considerou esse elemento. Confirmando a ideia do trabalho, o exame das *features* mais importantes para a realização da análise preditiva permitiu discutir e delimitar como e quais fatores psicológicos, econômicos e institucionais podem estar influenciando no processo decisório judicial. Isso foi feito especialmente com base em uma análise comparativa da importância de cada elemento identificado como relevante. Assim, foi possível verificar que alguns dos fatores que se imaginava serem relevantes com base em estudos anteriores, tais como a composição do Tribunal e o voto do relator, confirmaram-se relevantes para o modelo preditivo. Além disso, o fato de certas *features* não terem sido relevantes para a realização da análise preditiva indica que certos elementos podem não exercer tanta influência no processo decisório como se esperava.

É possível definir três principais contribuições realizadas por este trabalho. A primeira, e principal, é a contribuição para a compreensão sobre o processo decisório judicial, em especial do

STF. A análise preditiva indicou elementos do processo decisório que merecem ser estudados mais a fundo e permitiu delimitar fatores que podem estar influenciando juízes.

A segunda contribuição relaciona-se com a base de dados. Conforme mencionado, há uma dificuldade em realizar análises do STF considerando diferentes tipos de dados. Contudo, este trabalho contribuiu com uma base formada pelo relatório de 313 decisões e informações relativas ao relator, ano em que foram publicadas, requerente e voto do relator.

Por fim, a terceira contribuição é em relação à área do direito e da computação. Apesar de técnicas computacionais oferecerem grandes potenciais para pesquisas relacionadas ao direito, poucos estudos de fato aproveitam-se dessas técnicas. Conforme mencionado, apesar de já terem sido realizados estudos no exterior utilizando *machine learning* para prever decisões de Cortes Superiores, o mesmo ainda não havia sido realizado no Brasil. Espera-se que o início da construção desse modelo preditivo para o Supremo Tribunal Federal contribua para o desenvolvimento de outros trabalhos semelhantes no país.

7 REFERÊNCIAS BIBLIOGRÁFICAS

ALETRAS, Nikolaos; TSARAPATSANIS, Dimitrios; PREOȚIUC-PIETRO, Daniel; *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.

ARGUELHES, Diego Werneck. Poder não é querer: preferências restritivas e redesenho institucional no Supremo Tribunal Federal pós-democratização. **Universitas Jus**, v. 25, n. 1, 2014.

ARGUELHES, Diego Werneck; GOMES, Juliana Cesário Alvim; NOGUEIRA, Rafaela. GÊNERO E COMPORTAMENTO JUDICIAL NO SUPREMO TRIBUNAL FEDERAL: OS MINISTROS CONFIAM MENOS EM RELATORAS MULHERES? **Revista Brasileira de Políticas Públicas**, v. 8, n. 2, 2018.

BAILEY, Michael A; MALTZMAN, Forrest. Does legal doctrine matter? Unpacking law and policy preferences on the US Supreme Court. **American Political Science Review**, v. 102, n. 3, p. 369–384, 2008.

BENFORADO, Adam. **Unfair: The new science of criminal injustice**. [s.l.]: Broadway Books,

2015.

DA SILVA, V. A. Deciding without deliberating. **International Journal of Constitutional Law**, v. 11, n. 3, p. 557–584, 2013.

DANZIGER, Shai; LEVAV, Jonathan; AVNAIM-PESSO, Liora. Extraneous factors in judicial decisions. **Proceedings of the National Academy of Sciences**, v. 108, n. 17, p. 6889–6892, 2011.

FALCÃO, Joaquim; CERDEIRA, Pablo de Camargo; ARGUELHES, Diego Werneck. I Relatório Supremo em Números-o Múltiplo Supremo. 2012.

FEIGENSON, Neal. Emotional influences on judgments of legal blame: How they happen, whether they should, and what to do about it. *In*: **Emotion and the Law**. [s.l.]: Springer, 2009, p. 45–96.

GUTHRIE, Chris. Prospect theory, risk preference, and the law. **Nw. UL Rev.**, v. 97, p. 1115, 2002.

GUTHRIE, Chris; RACHLINSKI, Jeffrey J; WISTRICH, Andrew J. Inside the judicial mind. **Cornell L. Rev.**, v. 86, p. 777, 2000.

HUDSON, Alexander; HARTMANN, Ivar Alberto. Can you bury ideology? An empirical analysis of the ideal points of the Ministers of Brazil's Supremo Tribunal Federal. **A&C-Revista de Direito Administrativo & Constitucional**, v. 17, n. 68, p. 43–59, 2017.

KAHNEMAN, Daniel. **Thinking, fast and slow**. [s.l.]: Farrar, Straus and Giroux New York, 2011.

KATZ, Daniel Martin; BOMMARITO, Michael J.; BLACKMAN, Josh. A general approach for predicting the behavior of the Supreme Court of the United States. **PLOS ONE**, v. 12, n. 4, p. e0174698, 2017.

LANNES, O; DESPOSATO, S; INGRAM, M. Judicial behavior in civil law systems: changing patterns on the Brazilian Supremo Tribunal Federal. *In*: **CICLO 2012 DO PROGRAMA DE SEMINÁRIOS CIEF-CERME-LAPCIPP-MESP**, 2. [s.l.: s.n.], 2012.

LEAL, Fernando; RIBEIRO, Leandro Molhano. HEURÍSTICA DE ANCORAGEM E FIXAÇÃO DE DANOS MORAIS EM JUIZADOS ESPECIAIS CÍVEIS NO RIO DE JANEIRO: UMA NOVA ANÁLISE. **Revista Brasileira de Políticas Públicas**, v. 8, n. 2, 2018.

MUELLER, Bernardo; FERREIRA, Pedro Fernando Almeida Nery. How Judges Think In The Brazilian Supreme Court: Estimating Ideal Points And Identifying Dimensions. *In*: [s.l.]: ANPEC-

Associação Nacional dos Centros de Pós-Graduação em Economia [Brazilian Association of Graduate Programs in Economics], 2014.

PEDREGOSA, Fabian; VAROQUAUX, Gaël; GRAMFORT, Alexandre; *et al.* Scikit-learn: Machine Learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

POSNER, Richard A. **How Judges Think**. [s.l.]: Harvard University Press, 2010.

POSNER, Richard A. What Do Judges Maximize?(The Same Thing Everybody Else Does), 3 SUP. CT. **Econ. Rev**, v. 1, p. 3–4, 1993.

RACHLINSKI, Jeffrey J; GUTHRIE, Chris; WISTRICH, Andrew J. Inside the bankruptcy judge’s mind. **BUL Rev.**, v. 86, p. 1227, 2006.

RIBEIRO, Leandro Molhano; ARGUELHES, Diego Werneck. Preferências, Estratégias e Motivações: Pressupostos institucionais de teorias sobre comportamento judicial e sua transposição para o caso brasileiro. **Revista Direito e Práxis**, v. 4, n. 7, p. 85–121, 2013.

SCHAUER, Frederick. **Thinking like a lawyer**. [s.l.]: Harvard University Press, 2009.

SEGAL, Jeffrey A; SPAETH, Harold J. **The Supreme Court and the attitudinal model revisited**. [s.l.]: Cambridge University Press, 2002.

SILVA, Virgílio Afonso da. “UM VOTO QUALQUER”? O PAPEL DO MINISTRO RELATOR NA DELIBERAÇÃO NO SUPREMO TRIBUNAL FEDERAL. **REI - REVISTA ESTUDOS INSTITUCIONAIS**, v. 1, n. 1, p. 180–200, 2015.

SILVA, Jeferson Mariano. MAPEANDO O SUPREMO: As posições dos ministros do STF na jurisdição constitucional (2012-2017) 1. **Novos Estudos**, n. 110, p. 34–54, 2018.

SISK, Gregory C; HEISE, Michael; MORRISS, Andrew P. Charting the influences on the judicial mind: An empirical study of judicial reasoning. **NYUL rev.**, v. 73, p. 1377, 1998.

STERN, HENRI, Zhu, Michael. Predicting Supreme Court Decisions Using Supervised Learning.

1.13. Feature selection — scikit-learn 0.20.0 documentation. Disponível em: <http://scikit-learn.org/stable/modules/feature_selection.html#variance-threshold>. Acesso em: 4 nov. 2018.

3.2.4.3.1. sklearn.ensemble.RandomForestClassifier — scikit-learn 0.20.0 documentation.

Disponível em: <<http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>>. Acesso em: 5 nov. 2018.

4.2. Feature extraction — scikit-learn 0.20.0 documentation. Disponível em: <http://scikit-learn.org/stable/modules/feature_extraction.html#text-feature-extraction>. Acesso em: 4 nov. 2018.

Elasticsearch. Disponível em: <<https://www.elastic.co/products/elasticsearch>>. Acesso em: 4 nov. 2018.

HUDOC - European Court of Human Rights. Disponível em: <[https://hudoc.echr.coe.int/eng#{"documentcollectionid2":\["GRANDCHAMBER","CHAMBER"\]}](https://hudoc.echr.coe.int/eng#{)>. Acesso em: 20 nov. 2018.

Precision-Recall — scikit-learn 0.20.2 documentation. Disponível em: <https://scikit-learn.org/stable/auto_examples/model_selection/plot_precision_recall.html>. Acesso em: 28 fev. 2019.

Python Data Analysis Library — pandas: Python Data Analysis Library. Disponível em: <<https://pandas.pydata.org/>>. Acesso em: 4 nov. 2018.

sklearn.feature_extraction.text.TfidfVectorizer — scikit-learn 0.20.0 documentation. Disponível em: <http://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html>. Acesso em: 4 nov. 2018.

sklearn.model_selection.train_test_split — scikit-learn 0.20.0 documentation. Disponível em: <http://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html>. Acesso em: 4 nov. 2018.

Supreme Court Forecasting 2002 : Forecast. Disponível em: <<http://wusct.wustl.edu/forecast.php>>. Acesso em: 27 fev. 2019.

Supremo em Números. Disponível em: <<http://www.fgv.br/supremoemnumeros/>>. Acesso em: 4 nov. 2018.

The Supreme Court Database. Disponível em: <<http://supremecourtdatabase.org/>>. Acesso em:

19 nov. 2018.

