Judicialização do Transporte Aéreo: uma aplicação de aprendizado de máquina

Judicialization of Air Transport: an application of machine learning

Judicialización del Transporte Aéreo: una aplicación del aprendizaje automático

Como citar:

Oliveira, Thaís F L. & Celestino, Victor R. R. (2025). Judicialização do transporte aéreo: uma aplicação de aprendizado de máquina. *Revista Gestão & Tecnologia*, vol 25, nº 3, p: 149-177

Thaís Ferreira Lopes Oliveira, Mestranda em Administração (Finanças e Métodos Quantitativos) no PPGA/UnB
https://orcid.org/0009-0000-7180-9614

Victor Rafael Rezende Celestino, Professor Adjunto no Departamento de Administração da Universidade de Brasília (UnB).
https://orcid.org/0000-0001-5913-2997

"Os autores declaram não haver qualquer conflito de interesse de natureza pessoal ou corporativa, em relação ao tema, processo e resultado da pesquisa".

Editor Científico: José Edson Lara
Organização Comitê Científico
Double Blind Review pelo SEER/OJS
Recebido em 17/12/2024
Aprovado em 20/06/2025

Resumo

Objetivo: Padronizar a classificação dos motivos e causas dos processos judiciais contra companhias aéreas e analisar se essa padronização impacta no desempenho de modelos de aprendizado de máquina para prever os valores de indenizações judiciais.

Metodologia/procedimentos metodológicos: Foi aplicada uma metodologia semelhante à análise de conteúdo e revisões sistemáticas de literatura para reclassificar os motivos e causas dos processos judiciais. Em seguida, os modelos de aprendizado de máquina propostos por Torres, Guterres, and Celestino (2023) foram reaplicados com dados de processos judiciais de duas companhias aéreas nacionais, entre 2019 e 2022, abrangendo 16 estados brasileiros.

Originalidade/Relevância: O estudo combina dois temas previamente analisados separadamente por outros autores: a padronização das causas dos processos judiciais no setor aéreo e a previsão de indenizações.

Principais resultados: Os resultados obtidos demonstraram melhorias significativas nos três modelos avaliados, com acurácia, precisão, *recall* e *F1-score* superiores a 93%, e uma AUC-ROC acima de 98%. O modelo *Random Forest* se destacou como o mais eficaz em todas as métricas analisadas, especialmente na classificação precisa da categoria de indenizações "Baixo".

Contribuições teóricas/metodológicas: O estudo propõe a padronização da classificação das causas de processos judiciais no transporte aéreo e aplica com sucesso modelos de aprendizado de máquina para prever indenizações, oferecendo uma nova abordagem para a previsão de despesas no setor.

Palavras-chave: judicialização; transporte aéreo; aprendizado de máquina.

Abstract

Objective: To standardize the classification of reasons and causes of legal cases against airlines and analyze whether this standardization impacts the performance of machine learning models in predicting the values of judicial compensations.

Methodology/Procedures: A methodology similar to content analysis and systematic literature reviews was applied to reclassify the reasons and causes of legal cases. Then, the machine learning models proposed by Torres et al. (2023) were reapplied with data from legal cases of two national airlines between 2019 and 2022, covering 16 Brazilian states.

Originality/Relevance: The study combines two topics previously analyzed separately by other authors: the standardization of the causes of legal cases in the air transport sector and the prediction of compensations.

Main Results: The results showed significant improvements in all three evaluated models, with accuracy, precision, recall, and F1-score above 93%, and an AUC-ROC above 98%. The Random Forest model stood out as the most effective in all metrics analyzed, especially in the precise classification of the "Low" compensation category.

Theoretical/Methodological Contributions: The study proposes the standardization of the classification of legal case causes in air transport and successfully applies machine learning models to predict compensations, offering a new approach to predicting expenses in the sector.

Keywords: judicialization; air transport; machine learning.

Resumen
Objetivo: Estandarizar la clasificación de los motivos y causas de los procesos judiciales contra las aerolíneas y analizar si esta estandarización impacta en el desempeño de los modelos de aprendizaje automático para predecir los valores de las indemnizaciones judiciales.
Metodología/procedimientos metodológicos: Se aplicó una metodología similar al análisis de contenido y revisiones sistemáticas de literatura para reclasificar los motivos y causas de los procesos judiciales. Posteriormente, los modelos de aprendizaje automático propuestos por Torres et al. (2023) fueron reaplicados con datos de procesos judiciales de dos aerolíneas nacionales, entre 2019 y 2022, abarcando 16 estados brasileños.
Originalidad/Relevancia: El estudio combina dos temas previamente analizados por otros autores de manera separada: la estandarización de las causas de los procesos judiciales en el sector aéreo y la predicción de indemnizaciones.
Principales resultados: Los resultados mostraron mejoras significativas en los tres modelos evaluados, con precisión, exactitud, recall y F1-score superiores al 93%, y una AUC-ROC superior al 98%. El modelo Random Forest se destacó como el más eficaz en todas las métricas analizadas, especialmente en la clasificación precisa de la categoría de indemnizaciones "Bajo".
Contribuciones teóricas/metodológicas: El estudio propone la estandarización de la clasificación de las causas de los procesos judiciales en el transporte aéreo y aplica con éxito modelos de aprendizaje automático para predecir indemnizaciones, ofreciendo un nuevo enfoque para la predicción de gastos en el sector.
Palabras clave: judicialización; transporte aéreo; aprendizaje automático.

Introdução

O Brasil se destaca por um alto número de litígios envolvendo empresas de transporte aéreo. Em 2021, dados do Instituto Brasileiro de Direito Aeronáutico (IBAER) indicaram que 98,5% das ações cíveis no mundo contra companhias aéreas estavam concentradas no país (CNJ, 2021). Esse cenário gera altos custos relacionados a despesas legais e a compensações financeiras devido a indenizações, o que afeta os resultados financeiros das empresas do setor, gerando implicações na viabilidade da operação, no aumento do preço das passagens, na disponibilidade de malha aérea para determinados locais, além de criar um ambiente de insegurança jurídica.

Devido a essas implicações, compreender as causas desse fenômeno, assim como os fatores mais relevantes para a definição do valor da indenização judicial são de extrema importância. Um trabalho anterior de Torres et al. (2023) iniciou justamente esse estudo. Os autores compararam três modelos de aprendizado de máquina (*Random Forest*, *Support Vector*

Machines e *Naive Bayes*) a fim de identificar aquele com melhor desempenho na previsão do valor de indenização final pago ao consumidor. Os resultados obtidos mostraram um melhor poder preditivo para o *Random Forest*, porém os autores apresentaram uma limitação relacionada à diferença na classificação dos atributos de motivos e causas dos processos judiciais por cada companhia aérea, o que pode ter influenciado no desempenho dos algoritmos, visto que eventos semelhantes podem ter sido interpretados de forma diferente na criação do banco de dados.

Essa diversidade de classificações ocorre devido ao fato dos processos serem iniciados em diferentes instâncias da justiça e envolverem empresas distintas. Trabalhos anteriores estudaram a aplicação de técnicas de aprendizado de máquina para agrupar julgamentos jurídicos. No entanto, os resultados de Sabo, Dal Pont, Wilton, Rover, and Hübner (2022), por exemplo, demonstraram uma necessidade do envolvimento de especialistas para a definição dos rótulos, devido a complexidade dos dados textuais legais. "Multi-view overlapping clustering for the identification of the subject matter of legal judgments" (2023) também enfrentaram algumas limitações do contexto jurídico, como o fato de que os documentos legais tendem a estar relacionados a múltiplos assuntos e possuem uma semântica complexa. Assim, o agrupamento de casos com o mesmo objeto pode contribuir para aprimorar o processo de análise nesse contexto.

Nesse sentido, neste trabalho investigamos se o desempenho dos modelos de aprendizado de máquina desenvolvidos por Torres et al. (2023) seria alterado com a padronização da classificação dos atributos de motivos e causas informados pelas companhias aéreas. Para isso, usamos um conjunto de dados com informações de processos judiciais de duas companhias aéreas nacionais, no período de 2019 a 2022, considerando apenas 16 estados brasileiros. Aplicamos uma técnica similar a utilizada na análise de conteúdo e em revisões sistemáticas de literatura, definindo critérios de inclusão e exclusão para reclassificar os motivos e as causas nas novas categorias. Indicadores de precisão geral e área sob a curva ROC foram empregados como métricas de desempenho para comparar os modelos. Os resultados obtidos demonstraram que o desempenho dos três modelos apresentaram melhora, com destaque para o *Random Forest*.

O restante do artigo prossegue da seguinte forma: na seção 2, apresentamos uma contextualização a respeito da judicialização do transporte aéreo no Brasil e da aplicação do aprendizado de máquina para problemas de judicialização; na seção 3, apresentamos os dados e o método, com a descrição da base de dados, as etapas de pré-processamento e as técnicas de aprendizado de máquina utilizadas nesta pesquisa; na seção 4, mostramos os resultados obtidos por cada modelo e, por fim, na seção 5, apresentamos as considerações finais e sugestões de pesquisas futuras.

Referencial teórico

Judicialização Do Transporte Aéreo No Brasil

Inicialmente, a judicialização do transporte aéreo brasileiro poderia ser associada a um serviço de baixa qualidade. No entanto, as companhias aéreas brasileiras são reconhecidas internacionalmente por sua eficiência operacional, especialmente em aspectos como pontualidade e regularidade, inclusive em comparação com nações como os Estados Unidos, que possuem mercados aéreos mais desenvolvidos e são equivalentes ao mercado brasileiro em termos de extensão geográfica (ABEAR, 2023). Essa realidade evidencia a eficácia operacional do sistema de transporte aéreo brasileiro, o que ressalta ainda mais a incógnita sobre os motivos para tantos processos judiciais contra essas empresas.

Segundo a ABEAR (2023), um dos possíveis motivos é o estímulo que a legislação brasileira realiza para a litigância. Atualmente, existem diversas legislações gerais e específicas que são aplicáveis aos conflitos consumeristas no transporte aéreo. No âmbito geral, aplica-se a Constituição Federal de 1988 e o Código de Defesa do Consumidor (Lei nº 8.078/1990). No âmbito específico, aplica-se o Código Brasileiro de Aeronáutica (Lei nº 7.565, de 19 de dezembro de 1986, modificada pela Lei 14034/2020), o Código Civil (Lei nº 10.406, de 10 de janeiro de 2002 e modificações), a Convenção de Montreal (promulgada pela Decreto nº 5.910, de 27 de setembro de 2006) e a Resolução nº 400 da ANAC, de 13 de dezembro de 2016. Apesar disso, é frequente ocorrer decisões judiciais que não seguem a legislação específica,

especialmente nos tribunais de primeira instância, que muitas vezes não estão familiarizados com os tratados internacionais firmados pelo Brasil e seguem apenas a Constituição Federal.

Adicionalmente, é comum ver decisões judiciais e administrativas que ampliam as responsabilidades das empresas aéreas para além do que está estabelecido nas normas que regem o setor (Scaranello, Klarmann, & de Barros Silva, 2022). Também ocorre das decisões concederem indenizações por danos morais em casos já regulamentados, como, por exemplo, em situações fora do controle das companhias, como atrasos de voos causados pelo gerenciamento do tráfego aéreo (ABEAR, 2023). Tais acontecimentos estimulam a excessiva judicialização das companhias aéreas, pois motiva cada vez mais consumidores a buscarem alguma forma de reparação perante o Poder Judiciário (Scaranello et al., 2022), o que resulta em aumento de custos e geram um ambiente de incerteza legal no setor aéreo.

De acordo com os Painéis de Indicadores do Transporte Aéreo da ANAC dos anos de 2019 a 2022, os gastos com assistência a passageiros, indenizações extrajudiciais e condenações judiciais representaram 1,6%, 3,1%, 2,1% e 1,5% do total de custos e despesas dos serviços aéreos de cada ano, respectivamente. Esses percentuais equivalem aproximadamente aos valores representados na Figura 1.

As despesas decorrentes de indenizações em processos judiciais ou administrativos, observadas na Figura 1, são bem próximas do gasto com as tarifas aeroportuárias[1], sendo que até o superou em 2020. Ressalta-se que as tarifas aeroportuárias brasileiras são consideradas umas das mais altas do mundo (Scaranello et al., 2022).

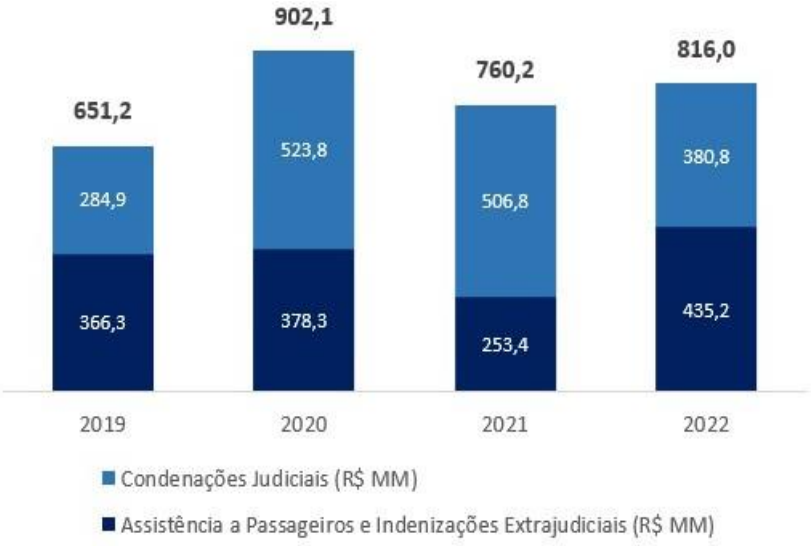
Além do impacto financeiro que as companhias aéreas enfrentam, as despesas elevadas com indenizações em processos judiciais ou administrativos também prejudicam os consumidores, pois esses custos são repassados para os preços das passagens. Além disso, o aumento no volume de processos já está afetando a disponibilidade de voos das empresas aéreas. Devido ao grande número de ações judiciais no Norte do país, as companhias aéreas têm

[1] As tarifas aeroportuárias representaram o percentual de 2,8% do total de custos e despesas dos serviços aéreos em 2019 e 2020, 2,4% em 2021 e 2,3% em 2022.

reduzido ou até eliminado voos entre cidades da região, que já é a menos atendida pelo transporte aéreo no Brasil (Ferraz, 2024).

Figura 1

Gastos com assistência a passageiros, indenizações extrajudiciais e condenações judiciais.



Aplicação Do Aprendizado De Máquina Nas Questões De Judicialização

O aprendizado de máquina (*machine learning*, em inglês) é um subcampo da inteligência artificial que estuda a capacidade de resolver problemas por meio de grandes volumes de dados. Seu objetivo é treinar algoritmos para identificar regras ou parâmetros que estabeleçam uma relação entre os dados de entrada (chamados de atributos preditivos) e os dados de saída (conhecidos como atributos alvo). Isso viabiliza a execução de diversas tarefas, como classificação, previsão e agrupamento de dados (Lenz, Neuman, Santarelli, & Salvador, 2020).

Alguns trabalhos anteriores aplicaram aprendizado de máquina em questões de judicialização e de transporte aéreo. Um problema pesquisado foi justamente a aplicação de abordagens de agrupamento ou *clustering*, com o objetivo de agrupar n indivíduos em grupos homogêneos (*clusters*), a fim de possibilitar uma análise com maior clareza do conjunto de

dados (Sicsú, Samartini, & Barth, 2023). Raghuveer et al. (2012) utilizaram *Latent Dirichlet Allocation* (LDA) para propor uma agrupamento com base em tópicos. Raghav, Balakrishna Reddy, Balakista Reddy, and Krishna Reddy (2015) aplicaram K-means para testar três agrupamentos utilizando citações e links sobre outros julgamentos. Zhang and Zhou (2019) usaram um algoritmo de vetor de parágrafo (Doc2Vec) para atualizar os resultados do agrupamento de documentos jurídicos, comparando a similaridade do texto sem reimplementar o agrupamento.

Alguns trabalhos também estudaram o tema considerando o contexto brasilerio. Sabo et al. (2022) buscaram agrupar julgamentos judiciais do Juizado Especial Cível localizado na Universidade Federal de Santa Catarina (JEC/UFSC), nas quais consumidores reivindicavam indenizações morais e materiais às companhias aéreas por falhas no serviço. Os autores aplicaram quatro algoritmos de *clustering*: hierárquico e lingo (*clustering* suave), *K-means* e propagação de afinidade (*clustering* rígido). Como resultado, os autores identificaram que o *clustering* hierárquico demonstrou ser a abordagem mais vantajosa, porém a tarefa de fornecer os rótulos cabia exclusivamente ao especialista. Por outro lado, o algoritmo lingo demonstrou a capacidade de gerar rótulos, permitindo ao especialista apenas identificar aqueles que não deveriam ser uma variável do julgamento. De forma geral, os autores concluíram que todas as abordagens testadas permitiram identificar padrões, porém com o apoio de especialistas jurídicos, demonstrando uma limitação das técnicas em explicar um evento jurídico.

Lima and Costa (2022) realizaram uma avaliação de seis diferentes abordagens para o agrupamento, também utilizando bases de dados de documentos jurídicos brasileiros. Eles testaram os algoritmos *K-means*, Mini *Batch K-Means* e HDBSCAN em diferentes hiperparâmetros e em conjunto ou não com o Mapa de Kohonen como técnica de pré-clustering. Além disso, o trabalho também propôs uma nova estrutura orientada para o Processamento de Linguagem Natural (PLN) para a avaliação específica de clusters textuais. Os resultados obtidos demonstraram que o K-means e o Mini Batch K-Means são as melhores escolhas, e o uso do mapa de Kohonen pode aumentar o desempenho geral do *clustering*.

Por outro lado, Torres et al. (2023) compararam o desempenho de três modelos de aprendizado de máquina (*Random Forest*, SVM e *Naive Bayes*) e da Regressão Logística Multinomial a fim de identificar as melhores técnicas para prever os valores indenizados. O estudo foi baseado em dados de ações judiciais ajuizadas em cidades brasileiras, no período de 2016 a 2021. Entre os modelos analisados, o *Random Forest* teve o melhor desempenho em relação ao modelos de aprendizagem de máquina, com valores semelhantes à Regressão Logística Multinomial, que se mostrou importante na classificação de conjuntos de dados categóricos. Contudo, apesar da melhor precisão, o tempo de processamento do *Random Forest* foi muito elevado, o que pode ser uma desvantagem se o poder computacional for limitado. Assim, a Regressão Logística Multinomial apresentou a melhor relação custo-benefício.

Metodologia

Bases De Dados

Utilizamos duas bases de dados disponibilizadas pelas companhias aéreas Gol Linhas Aéreas e Latam Airlines para o projeto de pesquisa "Diagnóstico sobre a Judicialização do Transporte Aéreo no Brasil: Uma aplicação de Aprendizado de Máquina"[2]. As bases continham informações dos processos judiciais movidos contra estas empresas no período de 2019 a 2022, considerando apenas 16 estados brasileiros (AM, BA, CE, DF, ES, GO, MG, MS, MT, PE, PR, RJ, RO, SC, SP e TO). Inicialmente, a base possuía 126.051 registros de processos judiciais.

Pré-processamento

Após a coleta dos dados, efetuamos uma etapa de pré-processamento para preparar os dados para a aplicação dos modelos. Nessa etapa, realizamos a padronização das categorias de

[2] Projeto de pesquisa iniciado em 05 de dezembro de 2023, com vigência até 05 de dezembro de 2024, decorrente de uma parceria da Associação Brasileira das Empresas Aéreas (ABEAR), da Associação Internacional de Transportes Aéreos (IATA), da Associação Latino-Americana e do Caribe de Transporte Aéreo (ALTA), da Junta de Representantes das Companhias Aéreas Internacionais do Brasil (JURCAIB), da Associação dos Magistrados Brasileiros (AMB) e da Universidade de Brasília (UnB).

motivos e causas e realizamos a limpeza e a organização dos dados, que consistiu na seleção das mesmas variáveis utilizadas por Torres et al. (2023), assim como a aplicação das categorias estabelecidas pelos autores. Após isso, todos os valores ausentes foram removidos, restando 98.114 registros de processos judiciais. Por fim, fiz a preparação dos dados para os modelos

Padronização Das Categorias De Motivos E Causas

Para padronizar os motivos e as causas, empregamos uma técnica de agrupamento por palavras-chave, desenvolvida pelo pesquisadores do projeto de pesquisa mencionado anteriormente. Essa técnica segue os critérios de inclusão e exclusão similares aos utilizados na análise de conteúdo e em revisões sistemáticas da literatura, com o objetivo de reunir os processos judiciais sobre o mesmo assunto. As novas categorias de motivos e causas apresentadas a seguir foram definidas pelos pesquisadores do projeto, assim como as palavras-chave para incluir os processos em cada grupo.

Em relação aos motivos, reclassificamos os processos em somente quatro categorias: Problemas Operacionais, Bagagem, Contrato e Outros, em que ficam aqueles que não se enquadram nas anteriores. Para fazer essa classificação, utilizamos as colunas "Objeto principal" e "Subobjeto principal", informadas inicialmente pelas empresas, em conjunto. Dessa forma, realizamos um trabalho de análise de inclusão e exclusão de palavras-chave descritas nas colunas "Objeto principal" e "Subobjeto principal", utilizando uma matriz no Excel. Por exemplo, caso o texto do objeto e subobjeto apresentassem a palavra-chave "acidente", o processo judicial receberia a classe "problemas operacionais". Ressaltamos que o processo judicial pode ter recebido mais de uma classe, tendo em vista que é comum os processos tratarem de diversos assuntos. A Tabela 1 apresenta a lista de palavras-chave definidas pelos pesquisadores do projeto de pesquisa e que consideramos para fazer as classificações de cada classe de motivo.

Tabela 1
Lista de palavras-chave do agrupamento de motivos

Classes	Palavras-chave
Problemas Operacionais	'acidente', 'aeroporto', 'aeronave', 'alteração', 'assento', 'assiatance', 'assistência', 'atendimento', 'atr.', 'atraso', 'avião', 'cadeira de rodas', 'cambio', 'canc.', 'cancel', 'catering', 'check in', 'conex', 'demora', 'denied', 'delay', 'desembarque', 'dowgrade', 'downgrade', 'embarque', 'espaço', 'falha', 'flight', 'fretamento', 'incidente', 'malha', 'manutenção', 'menor desacompanhado', 'overbooking', 'dano a passageiro', 'poltrona', 'pontualidade', 'problemas operacionais', 'refeição', 'remarcação', 'segurança', 'service', 'serviço', 'schedule', 'sobreventa', 'solicitação especial', 'tripulação', 'voo'
Bagagem	'anima', 'bagagem', 'baggage', 'carg', 'equipaje', 'extravio'
Contrato	'bilhete', 'boleto', 'cadastro', 'call center', 'clube', 'cobrança', 'contrato', 'corona', 'covid', 'diverg', 'document', 'economy', 'fee', 'fidelidade', 'fraud', 'flyer', 'milhas', 'multa', 'multiplus', 'show', 'pass', 'plusgrade', 'políticas', 'pontuação', 'prescrição', 'producto', 'programa', 'promoç', 'propaganda', 'reembolso', 'regra', 'reimbursement', 'requisitos de viaje', 'reserva', 'seguro', 'site', 'tarifa', 'taxa', 'upgrade', 'prorrogação de voucher', 'viag'
Outros	'Problemas Decorrentes de Ação/Omissão do Autor', 'DIVERSO', 'DIVERSOS', 'Ação De Indenização por Danos Materiais', 'Ação De Indenização por Danos Morais', 'Ação de Indenização por Danos Morais e Materiais', 'Ação De Obrigação De Fazer', 'Ação de Obrigação de Fazer com Indenização por Danos Morais', 'Ação De Reparação De Danos', 'Ação De Restituição', 'Ação Indenizatória', 'Acordo', 'AJUSTAR', 'Carta Precatória Cível', 'Comparência Tardia', 'Conhecimento/Inicial', 'Cumprimento de Sentença', 'Cumprimento Provisório de Sentença', 'Execução', 'Investigação', 'Irregularidade', 'Outros Feitos Não Especificados', 'por questões médicas', 'Procedimento Do Juizado Especial Cível', 'Reclamação Procon', 'Recursal', 'Reparação De Danos', 'Ressarcimento - Valor Pago C/c Danos Morais', 'Restituição - Valores C/c Reparação De Danos Morais'

Realizamos a mesma análise para os dados de causa, sendo definidas sete categorias: Aérea, Cliente, Terceiros, Casos Fortuitos, Força Maior, Não Informado e Outros. Para fazer essa classificação, utilizei as colunas de "Objeto principal", "Subobjeto principal" e "Causa", informadas inicialmente pelas empresas, em conjunto. A Tabela 2 apresenta a lista de palavras-chave definidas pelos pesquisadores do projeto de pesquisa e que considerei para fazer as classificações de cada classe de causa. Após essas duas análises de inclusão e exclusão, criei quatro colunas referentes às novas classificações de motivos e sete sobre as causas, em que o 1

indica que o processo faz parte daquele grupo e o 0 que ele não faz parte

Tabela 2
Lista de palavras-chave do agrupamento de causas

Classes	Palavras-chave
Aérea	'culpa tam', 'abastecimento', 'aeronave', 'ajuda técnica', 'bagagem', 'cancelamento indevido', 'carga', 'cargo', 'CCO', 'Clube TudoAzul', 'cobrança equivocada', 'conexão', 'divergência de informações', 'empresa', 'entrega', 'extravio', 'falha', 'impedimento', 'malha', 'manutenção', 'oferta não cumprida', 'operacional', 'overbooking', 'overload', 'pontuação não creditada', 'planejamento', 'reembolso', 'regra', 'reserva', 'segurança', 'serviço', 'transporte terrestre', 'tripula', 'voo'
Cliente	'aborrecido', 'Ausência de responsabilidade', 'autor', 'bagagem de mão','Bagagem entregue dentro do prazo legal', 'comprovação', 'cliente', 'compra não reconhecida', 'direito de arrependimento', 'documentação', 'exclusiva', 'incorformidade', 'Não Verificada Ocorrencia', 'no show', 'pax', 'pontos expirados', 'prescrição'
Terceiros	'aeroport', 'agência', 'externo', 'fraude', 'ilícito', 'imigração', 'terceiro', 'tráfego', 'torre de controle'
Casos Fortuitos	'fortuito', 'greve', 'indisciplinado'
Força Maior	'condições climáticas', 'corona', 'covid', 'força maior', 'force majeure', 'tempo', 'meteorologia', 'weather'
Outros	'Passageiro aborrecido'

Limpeza E Organização Dos Dados

Após a padronização das categorias de motivos e causas, realizamos a limpeza e organização dos dados. Primeiramente, selecionamos as mesmas variáveis utilizadas por Torres et al. (2023), sendo elas:

- *Companhia aérea:* por fornecer entendimento sobre a frequência de reclamações que cada empresa recebe dos passageiros, o que pode influenciar a decisão final dependendo das

reinvidicações;

- *Ano e temporada:* por permitirem que as companhias aéreas entendam a proporção de ações judiciais ao longo do período e avaliem se o comportamento dos valores de indenização muda devido a algum fator externo, como a pandemia da COVID-19. No caso da temporada, também é possível identificar se há mudanças nos meses de férias no Brasil, tendo em vista que ocorrem mais viagens por ser alta temporada;

- *Região:* por ser uma variável essencial para compreender o comportamento das ações judiciais em todo o Brasil, onde há uma vasta extensão territorial com regiões tanto altamente desenvolvidas quanto subdesenvolvidas, assim conhecer as áreas com as maiores concentrações de reclamações auxilia as companhias aéreas a agirem de maneira eficaz nessas localidades;

- *Dano moral e material:* por fornecerem informações sobre o tipo de compensação que os passageiros recebem, seja com maior incidência de danos patrimoniais, que visam compensar perdas materiais, ou danos morais, que são mais proeminentes nas decisões judiciais;

- *Motivos e causas:* por fornecerem dados sobre os principais motivos das reclamações dos clientes e a origem dos problemas que eles mencionaram;

- *Decisão:* por mostrar se a reclamação do passageiro foi considerada após o encerramento do processo;

- *Valor pago:* por representar o valor de indenização pago pelas companhias aéreas aos seus clientes em razão de reclamações apresentadas no processo.

Com as variáveis selecionadas, fizemos algumas etapas para padronizar as categorias das variáveis. Em relação a época, a partir dos meses da data de distribuição dos processos, definimos alta temporada para os meses de férias (janeiro, fevereiro, julho e dezembro) e baixa

temporada para os demais. No caso da região, reduzimos as informações de Unidade Federativa nas cinco regiões brasileiras (Centro-Oeste, Nordeste, Sul, Norte e Sudeste). Para dano moral, dano material e valor pago de indenização, categorizamos os dados em três intervalos, sendo eles:

Baixo <= R$1000,00; R$1000,00 < Médio <= R$5000,00; e Alto > R$5000,00.

Ressaltamos que as categorias da variável decisão apresentam categorias distintas em relação ao estudo de Torres et al. (2023), tendo em vista que elas foram apresentadas de uma forma diferente nas bases de dados utilizadas neste trabalho. Assim como as variáveis de motivo e causa, que foram padronizadas anteriormente em quatro e sete classes, respectivamente. A Tabela 3 apresenta um resumo final das variáveis utilizadas, suas definições e categorias.

Tabela 3
Variáveis

Variável	Definição	Categorias
Companhia aérea	Companhias aéreas analisadas	Gol Latam
Ano	Ano do início do processo judicial	2019 2020 2021 2022
Época	Época em que a ação foi iniciada	Alta Baixa
Região	Região em que a ação foi iniciada	Centro-Oeste Nordeste Sul Norte Sudeste
Valor Dano Moral	Compensação por dano moral	Baixo Médio Alto
Valor Dano	Compensação por dano material	Baixo

Variável	Definição	Categorias
Material		Médio Alto
Valor Pago	Total de compensação paga ao cliente	Baixo Médio Alto
Decisão	Julgamento do processo	Decisão favorável à aérea (improcedência) Condenação Extinção sem julgamento do mérito Acordo Não informado
Motivo	Motivo apresentado pelo passageiro para entrar com uma ação judicial contra a companhia aérea	Problemas operacionais Bagagem Contrato Outros
Causa	Fator informado pela empresa que levou ao problema relatado pelo cliente	Aérea Cliente Terceiros Fortuitos Força maior Não informado Outros

Preparação Dos Dados Para Os Modelos

Para preparar os dados para os modelos, primeiramente dividimos os dados de entrada (X) e os dados de saída (y), excluindo a coluna de valor pago do conjunto de dados principal para formar o X, e mantendo-a no y. Em seguida, tendo em vista que o conjunto de dados era formado por variáveis categóricas, aplicamos *one-hot enconding* nas variáveis do X e *label encoding* no

y. O *one-hot enconding* cria uma nova coluna para cada valor único de uma coluna existente, indicando com 1 quando o valor é afirmativo e com 0 quando negativo, ele é recomendado quando se trata de variáveis nominais. Por outro lado, o *label encoding* designa um valor inteiro (0, 1, 2, ...) para cada valor único da coluna existente, sendo recomendando quando se trata de variáveis ordinais.

Por fim, dividimos a base em treino e teste usando a função *train_test_split* do *scikit-learn*, com os seguintes parâmetros: test_size=0.1, stratify=y, random_state=42. Isso quer dizer que:

- *test_size=0.1:* 10% dos dados foram reservados para o conjunto de teste e 90% para o de treino.

- *stratify=y:* a divisão foi estratificada com base na variável y, ou seja, as proporções relativas de diferentes classes em y são mantidas nos conjuntos de treino e teste;

- *random_state=42:* garante a reprodutibilidade da divisão, ou seja, que a divisão será sempre a mesma se o código for executado várias vezes.

Algoritmos De Aprendizado De Máquina

Após a etapa de pré-processamento, executamosos três modelos de aprendizado de máquina, utilizando os algoritmos e melhores hiperparâmetros identificados por Torres et al. (2023), sendo eles:

- Naive Bayes Classifier (NB): algoritmo *MultinomialNB* com alpha igual a 8;

- Support Vector Machine (SVM): algoritmo *LinearSVC* com valor de penalidade (penalty) igual a 12, valor de perda (*loss*) igual a *hinge* e valor de regularização (C) igual a 2;

- *Random Forest Classifier (RF):* algoritmo *RandomForestClassifier* com o número de árvores (*n_estimators*) igual 1.000, medida de qualidade (*criterion*) igual a *gini*, profundidade máxima da árvore (*max_depth*) igual a 12 e número mínimo de divisões de um nó (*min_samples_split*) igual a 4.

Tendo em vista que os dados deste trabalho são multiclasse, foi necessária uma modificação interna nos algoritmos que permite dividir a classificação multiclasse em várias classificações binárias. Dessa forma, aplicamos a abordagem *one vc one* ("ovo") que treina um classificador binário para cada par de classes possíveis no conjunto de dados. Isso foi utilizado para obter a área sob a curva ROC de cada modelo.

Naive Bayes Classifier

O *Naive Bayes Classifier* (NB) é um algoritmo supervisionado fundamentado no raciocínio bayesiano. Ele é utilizado em tarefas de aprendizado em que cada instância *x* é descrita por um conjunto de atributos (Ex.: companhia aérea, ano, época, região, valor do dano moral e material etc.) e em que a classe alvo pode assumir qualquer valor de um conjunto *V* (Ex.: baixo, médio e alto). Dessa forma, o NB classifica uma instância de dados ao atribuir o valor mais provável de *y* para o alvo, considerando os valores de atributos (Mitchell & Mitchell, 1997).

O algoritmo Naive Bayes opera em várias etapas para realizar essa classificação de instâncias. Primeiramente, calcula as probabilidades *a priori* de cada classe com base nos dados de treinamento. Em seguida, estima as probabilidades condicionais de cada valor de atributo para cada classe. Utilizando o Teorema de Bayes e assumindo independência condicional entre os atributos, calcula a probabilidade *a posteriori* de cada classe para a instância em teste. A classe com a maior probabilidade *a posteriori* é então escolhida como a predição para a instância.

Support Vector Machine

A *Support Vector Machine* (SVM) é uma técnica que ajuda na seleção de um classificador adequado para um conjunto de dados específico (Faceli, Lorena, Gama, Almeida, & Carvalho, 2021). Assim, seu objetivo é encontrar uma fronteira que separe as possíveis classes no conjunto de dados e classificar o maior número possível de exemplos, maximizando a distância da fronteira aos pontos mais próximos desta (Géron, 2019).

Além disso, como o algoritmo que aplicamos foi o *LinearSVC* e ele não possui um método para calcular as probabilidades previstas de cada classe, utilizamos um calibrador de modelos para classificações multiclasse chamado *CalibratedClassifierCV*. Com isso, foi possível obter o valor da área sob a curva ROC.

Random Forest Classifier

O *Random Forest Classifier* (RF) é um algoritmo *ensemble*, isto é, que combina o resultado de múltiplos modelos em busca de produzir um melhor modelo preditivo. Assim, ele combina várias árvores de decisão, treinadas com atributos e conjunto de dados distintos, e utiliza o método *bagging* para unir o resultado de todas as árvores geradas e identificar a classe final por meio de votação (Breiman, 2001). Ressalta-se que árvores de decisão são um tipo de estrutura de dados que possui nós internos, ramos e nós folhas. Em cada nó interno ocorre a verificação de um dos atributos. Os ramos representam os possíveis resultados da verificação. E os nós folhas representam as classes a serem atribuídas (Rocha, 2020).

De forma geral, as principais etapas do algoritmo são: i) um conjunto de n amostras é criado utilizando *bootstraping*, isto é, são selecionadas amostras aleatórias com reposição (ou seja, a mesma linha pode ser escolhida mais de uma vez); ii) uma árvore de decisão é construída por amostra, resultando na predição de uma classe; iii) as classes determinadas por cada árvore são computadas como um voto, sendo que a classe com maior quantidade de votos se torna a classe predita do modelo (de Freitas, 2018; de Oliveira Torres, 2022).

Avaliação Das Métricas De Desempenho Dos Modelos

Para avaliar o desempenho de cada modelo utilizamos as métricas avaliadas por Torres et al. (2023) para fins de comparação, sendo elas: matriz de confusão, área sob a curva ROC, acurácia e tempo de processamento do modelo em minutos. Adicionalmente, também calculamos três métricas comumente usadas para problemas de classificação multiclasse: precisão, *recall* e F1-*score* (Rajendran, Srinivas, & Grimshaw, 2021).

A matriz de confusão é uma representação tabular do desempenho de um modelo de classificação no conjunto de teste, mostrando o número de previsões corretas e incorretas para cada categoria (Ahmad, Yousaf, Yousaf, & Ahmad, 2020). A Tabela 4 apresenta um exemplo dessa matriz, em que as linhas representam as categorias reais e as colunas as previstas,

resultando em quatro parâmetros:

- *verdadeiro positivo (TP):* casos em que o modelo previu corretamente a classe positiva (verdadeira);
- *falso positivo (FP):* casos em que o modelo previu incorretamente a classe positiva (falso);
- *falso negativo (FN):* casos em que o modelo previu incorretamente a classe negativa (falso);
- *verdadeiro negativo (TN):* casos em que o modelo previu corretamente a classe negativa (verdadeiro).

Tabela 4
Matriz de Confusão

	Previsto Positivo	**Previsto Negativo**
Real Positivo	TP	FN
Real Negativo	FP	TN

A área sob a curva ROC (AUC-ROC) mostra se o classificador funciona melhor do que uma escolha aleatória. Para isso, as taxas de verdadeiros-positivo e falsos-positivo são relacionadas, o que permite saber se o classificador consegue diferenciar bem as classes (Murphy, 2012). Quanto maior a AUC-ROC, melhor o modelo é em distinguir as classes.

A acurácia indica uma performance geral do modelo. Ela é definida como a taxa total de acerto do algoritmo, ou seja, calcula dentre todas as classificações, quantas ele classificou corretamente, considerando tanto as verdadeiras como as falsas (Ahmad et al., 2020). Já a precisão mostra a proporção de verdadeiros positivos em relação a todos os classificados como positivos.

Por outro lado, o *recall* mostra quantas observações positivas foram classificadas corretamente em relação ao total de observações positivas. Por fim, o F1-*score* é uma média harmônica entre a precisão e o *recall*, sendo utilizada para comprar dois modelos quando estas métricas são importantes (Sicsú et al., 2023). Todas essas medidas variam de 0 a 1, sendo que quando maior o valor, melhor o desempenho de classificação. A Tabela 5 apresenta as fórmulas de cada métrica.

Tabela 5
Fórmulas das métricas de avaliação de classificação.

Métrica	**Fórmula**
Acurácia	$\frac{TP + TN}{TP + TN + FP + FN}$
Precisão	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$
F1-score	$2 \text{ x } \frac{Precisão\ x\ Recall}{Precisão + Recall}$

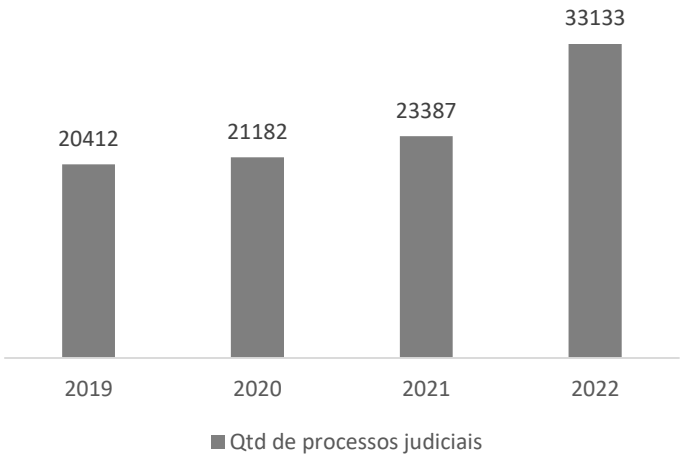
Para executar todas as etapas descritas utilizamos as bibliotecas *Pandas*, *Numpy* e *Scikit-learn* do Python. O computador que utilizamos possui as seguintes características: DELL, 32 GB de RAM, processador 13th Gen Intel(R) Core(TM) i9-13900H 2.60 GHz.

Apresentação e discussão dos resultados

Visão Geral Dos Dados

Ao todo analisamos 98.114 registros de processos judiciais, sendo que 77,13% eram da Latam e 22,87% eram da Gol. A Figura 2 apresenta a distribuição desses processos por ano.

Figura 2
Quantidade de processos judiciais por ano.



Observa-se na Figura 2 que a quantidade de processos foi crescendo ao longo dos anos. Em relação a época, 66,37% deles ocorreu em baixa temporada, enquanto 33,63% ocorreu em alta temporada. A maioria desses processos estão concentrados na região Sudeste (40,21%), seguida pelo Sul (17,86%), Norte (16,82%), Centro-Oeste (12,93%) e Nordeste (12,18%). Em relação a decisão, 49,99% resultou em condenação, 25,38% em acordo, 16,05% em decisão favorável à aérea (improcedência), 8,38% não foi informado e 0,20% em extinção sem julgamento do mérito. A Tabela 6 apresenta as frequências relativas dos valores de dano moral, valor material e valor pago.

Tabela 6
Frequências relativas dos valores de dano moral, valor material e valor pago

Valor	Dano moral	Dano material	Valor pago
Baixo	36,03%	36,03%	47,19%
Médio	38,73%	38,73%	29,61%
Alto	25,24%	25,24%	23,20%

Observa-se na Tabela 6, que tanto para o dano moral como para o dano material a classe

com maior incidência foi a de valor médio, isto é, entre maior que R$1000,00 e menor ou igual a R$5000,00. Já em relação a valor pago de indenização, a maior frequência é de valores baixos, ou seja, aqueles que são menores ou iguais a R$1000,00. A Tabela 7 apresenta a quantidade de processos por motivo e por causa. Ressaltamos que a soma dos processos ultrapassa o total de registro devido ao fato de um processo poder pertencer a mais de uma classe, quando trata de mais de um assunto.

Tabela 7
Quantidade de processos por motivo e causa

Motivo	**Qtd de processos**
Problemas operacionais	72327
Bagagem	11583
Contrato	21143
Outros	807
Causa	**Qtd de processos**
Aérea	51003
Cliente	14974
Terceiros	10113
Fortuitos	0
Força maior	14170
Não informado	7748
Outros	106

Observa-se na Tabela 7 que o principal motivo é problemas operacionais (68,32%) e a principal causa é a própria aérea (51,98%).

Resultados Dos Modelos

A Tabela 8 apresenta os resultados das métricas globais de desempenho de cada modelo, sendo que os melhores valores estão em negrito. Ressaltamos que para a precisão, o *recall* e o *F1-score* foi utilizado o parâmetro *average='weighted'*, tendo em vista o problema multiclasse. Esse parâmetro calcula a média dessas métricas, ponderando cada classe pela sua distribuição no conjunto de dados. Isso significa que classes com mais amostras têm um impacto maior

na métrica final do que classes com menos amostras, refletindo melhor o desempenho geral do modelo em um problema multiclasse.

Tabela 8
Desempenho global dos modelos de aprendizado de máquina

Métricas	**NB**	**SVM Linear**	**RF**
Acurácia	0,9315	0,9321	**0,9369**
Precisão	0,9344	0,9359	**0,9392**
Recall	0,9315	0,9321	**0,9369**
F1-score	0,9310	0,9314	**0,9365**
AUC-ROC	0,9824	0,9818	**0,9890**
Tempo de processamento (min)	**0,6456**	1,6223	1,7618

Observa-se na Tabela 8 que todos os modelos apresentaram resultados muito próximos, apenas com diferenças decimais. De forma geral, todos eles apresentaram uma acurácia de 93%, o que significa que 93% das previsões feitas pelo modelo estão corretas em relação ao total de previsões feitas. Em outras palavras, o modelo está acertando a classificação da maioria das instâncias. Em relação a precisão dessas predições, ou seja, a proporção de predições positivas corretas em relação ao total de predições positivas, o valor também foi de 93%. O *recall* foi de 93%, o que indica que os modelos estão capturando corretamente 93% das instâncias positivas existentes no conjunto de dados. Isto é relevante para problemas em que é crucial identificar corretamente todas as instâncias de uma classe. E o *F1-score*, que combina precisão e *recall* em uma única métrica, também foi de 93%, o que sugere um bom equilíbrio entre essas duas métricas, indicando que os modelos estão acertando tanto na identificação correta das instâncias positivas quanto na minimização dos falsos positivos. Por fim, a AUC-ROC indica quão bem o modelo consegue distinguir as diferentes classes e todos os modelos apresentaram um valor acima de 98%.

Em suma, o RF se destacou com os melhores resultados para todas as métricas entre os três modelos. Em relação aos tempos de processamento, o NB foi o que processou em menor

tempo devido à sua simplicidade em comparação com as outras técnicas, seguido pelo SVM e o RF, o que está de acordo com o trabalho de Torres et al. (2023) e outros autores como Ting, Ip, Tsang, et al. (2011), Tsangaratos and Ilia (2016) e Lei et al. (2017). A Tabela 9 apresenta a comparação dos desempenhos obtidos por Torres et al. (2023) e neste trabalho.

Tabela 9
Comparação do desempenho dos modelos de aprendizado de máquina

Métricas	**Torres et al. (2023)**			**Este trabalho**			**Diferença de desempenho**		
	NB	**SVM Linear**	**RF**	**NB**	**SVM Linear**	**RF**	**NB**	**SVM Linear**	**RF**
Acurácia	0,766	0,809	0,831	0,9315	0,9321	0,9369	21,61%	15,22%	12,74%
AUC-ROC	0,914	0,879	0,957	0,9824	0,9818	0,989	7,48%	11,70%	3,34%
Tempo de processamento (min)	0,58	32,19	256,13	0,602	0,6109	0,679	11,31%	-94,96%	-99,31%

Este trabalho teve como objetivo superar a limitação relacionada à diferença na classificação dos atributos de motivos e causas dos processos judiciais por cada companhia aérea, enfrentada por Torres et al. (2023). Dessa forma, foi feita a padronização da classificação dos atributos de motivos e causas e executado os mesmos modelos a fim de observar o desempenho após essa etapa. Observa-se na Tabela 9 que, de fato, houve uma melhora significativa em todos eles. O NB teve 21,61% de aumento de desempenho na acurácia e 7,48% na AUC-ROC. O SVM Linear teve um aumento de 15,22% na acurácia e 11,70% na AUC-ROC. E o RF teve 12,74% de aumento na acurácia e 3,34% na AUC-ROC. Em relação aos tempos de processamento, as reduções significativas no tempo do SVM Linear e do RF estão relacionadas ao fato de que nesse trabalho não teve a aplicação de *GridSearchCV*, tendo em vista que utilizamos os mesmos hiperparâmetros selecionados pelos autores.

Além das métricas globais apresentadas acima, também foram calculadas as matrizes de confusão de cada modelo. Por meio delas, é possível avaliar o desempenho de um classificador por meio da análise de quais classes foram classificadas corretamente. As Tabelas 10, 11 e 12 apresentam as matrizes de confusão de cada modelo

Tabela 10
Matriz de Confusão - Naive Bayes

	Predito: Alto	**Predito: Baixo**	**Predito: Médio**
Real: Alto	1961	134	181
Real: Baixo	0	4580	51
Real: Médio	0	306	2599

Tabela 11
Matriz de Confusão - SVM Linear

	Predito: Alto	**Predito: Baixo**	**Predito: Médio**
Real: Alto	1961	150	165
Real: Baixo	0	4628	3
Real: Médio	0	348	2557

Tabela 12
Matriz de Confusão - Random Forest

	Predito: Alto	**Predito: Baixo**	**Predito: Médio**
Real: Alto	1961	110	205
Real: Baixo	1	4566	64
Real: Médio	0	239	2666

A partir dos valores apresentados nas Figuraa 10, 11 e 12 foi possível calcular as métricas para cada classe. Em relação a precisão de cada classe:

- *Alto:* tanto o NB como o SVM Linear ficaram com 100%, enquanto o RF ficou com 99,95%;
- *Baixo:* o melhor foi o RF com 92,90%, seguido pelo NB com 91,24% e o SVM Linear com 90,28%;

- *Médio:* o melhor foi o SVM Linear com 93,83%, seguido pelo NB com 91,81% e o RF com 90,83%.

Em relação ao *recall* de cada classe:

- *Alto:* os três modelos ficaram com 86,16%;
- *Baixo:* o melhor foi o SVM Linear com 99,94%, seguido pelo NB com 98,90% e o RF com 98,60%;
- *Médio:* o melhor foi o RF com 91,77%, seguido pelo NB com 89,47% e o SVM Linear com 88,02%.

Em relação ao *F1-score* de cada classe:

- *Alto:* tanto o NB como o SVM Linear ficaram com 92,57%, enquanto o RF ficou com 92,54%;
- *Baixo:* o melhor foi o RF com 95,66%, seguido pelo NB com 94,91% e o SVM Linear com 94,87%;
- *Médio:* o melhor foi o RF com 91,30%, seguido pelo SVM Linear com 90,83% e o SVM Linear com 90,62%.

Em geral, a classe "Baixo" foi a melhor classificada pelos três modelos, com destaque para o RF, que obteve a maior precisão (92,90%), *recall* (98,60%) e *F1-score* (95,66%) nesta categoria.

Considerações finais

Os desafios enfrentados pelas companhias aéreas brasileiras em relação aos litígios judiciais são significativos, refletindo-se em altos custos legais e indenizações. Assim, a aplicação de técnicas de aprendizado de máquina neste contexto são úteis na previsão de valores de in- denizações, porém para que os modelos tenham uma boa performance é necessário se atentar para a complexidade dos dados jurídicos.

Os resultados obtidos revelaram que a variação na classificação dos motivos e causas dos processos judiciais por diferentes empresas impacta o desempenho dos modelos de aprendizado de máquina na previsão de indenizações. Além disso, a padronização desses atributos através

da metodologia baseada em análise de conteúdo e revisões sistemáticas de literatura mostrou-se eficaz para melhorar a precisão dos modelos. Especificamente, o modelo *Random Forest* se destacou como o mais eficaz neste trabalho, assim como no estudo anterior de Torres et al. (2023), o que reforça que ele pode ser útil para o planejamento estratégico das companhias aéreas diante de litígios judiciais.

Além disso, ressaltamos que estudos anteriores enfatizaram a necessidade de especialistas para a definição precisa dos rótulos e categorias, dado o caráter multifacetado e a semântica complexa dos documentos legais. Portanto, a abordagem metodológica adotada neste trabalho oferece um caminho promissor para futuras pesquisas que visem aprimorar ainda mais a precisão e a aplicabilidade prática dos modelos de aprendizado de máquina no contexto da judicialização do transporte aéreo. Assim, em trabalhos futuros propomos o uso dessa metodolodia aliada a aplicação de outras técnicas de aprendizado de máquina, como redes neurais.

Por fim, o progresso neste campo não apenas contribui para a mitigação de riscos legais e financeiros enfrentados pelas companhias aéreas, mas também promove uma maior segurança jurídica no setor, potencialmente impactando positivamente a operação das empresas, a competitividade do mercado e, consequentemente, o bem-estar dos consumidores e *stakeholders* envolvidos.

Referências

ABEAR. (2023). *Panorama 2022 – o setor aéreo em dados e análises.*

Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2020). Fake news detection using machine learning ensemble methods. *Complexity, 2020*(1), 8885861. https://doi.org/10.1155/2020/8885861

ANAC. (2021a). *Painel de indicadores do transporte aéreo 2019.* ANAC. (2021b). *Painel de indicadores do transporte aéreo 2020.* ANAC. (2022). *Painel de indicadores do transporte aéreo 2021.* ANAC. (2023). *Painel de indicadores do transporte aéreo 2022.*

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5–32. https://doi.org/10.1023/A:1010933404324

CNJ. (2021). *Cartilha do transporte aéreo.*

de Freitas, C. C. G. (2018). *Demanda por seguro de automóvel no rio de janeiro* [Dissertação de Mestrado em Economia]. Rio de Janeiro, Rio de Janeiro.

de Oliveira Torres, G. (2022). *Judicialização no transporte aéreo brasileiro: uma análise por meio de aprendizado de máquina e de regressão logística multinomial* [Dissertação de Mestrado em Engenharia de Infraestrutura Aeronáutica]. São José dos Campos, São Paulo.

Faceli, K., Lorena, A. C., Gama, J., Almeida, T. A. d., & Carvalho, A. C. P. d. L. F. d. (2021). Inteligência artificial: uma abordagem de aprendizado de máquina.

Ferraz, L. (2024). *Em resposta a processos judiciais, aéreas cortam oferta de voos no norte do brasil.*

Géron, A. (2019). *Mãos à obra: Aprendizado de máquina com scikit-learn & tensorflow*. Alta Books.

Lei, M., Ge, J., Li, Z., Li, C., Zhou, Y., Zhou, X., & Luo, B. (2017). Automatically classify chinese judgment documents utilizing machine learning algorithms. In *Database systems for advanced applications: Dasfaa 2017 international workshops: Bdms, bdqm, secop, and dmmooc, suzhou, china, march 27-30, 2017, proceedings 22* (pp. 3–17).

Lenz, M., Neuman, F., Santarelli, R., & Salvador, D. (2020). *Fundamentos de aprendizagem de máquina*.

Lima, J. P., & Costa, J. A. (2022). Comparing clustering techniques on brazilian legal document datasets. In *International conference on hybrid artificial intelligence systems* (pp. 98– 110).

Mitchell, T. M., & Mitchell, T. M. (1997). *Machine learning* (Vol. 1) (No. 9). McGraw-hill New York.

Multi-view overlapping clustering for the identification of the subject matter of legal judgments. (2023). *Information Sciences*, *638*, 118956. doi: https://doi.org/10 .1016/j.ins.2023.118956

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

Raghav, K., Balakrishna Reddy, P., Balakista Reddy, V., & Krishna Reddy, P. (2015). Text and citations based cluster analysis of legal judgments. In *Mining intelligence and knowledge exploration: Third international conference, mike 2015, hyderabad, india, december 9- 11, 2015, proceedings 3* (pp. 449–459).

Raghuveer, K., et al. (2012). Legal documents clustering using latent dirichlet allocation. *IAES Int. J. Artif. Intell*, *2*(1), 34–37.

Rajendran, S., Srinivas, S., & Grimshaw, T. (2021). Predicting demand for air taxi urban aviation services using machine learning algorithms. *Journal of Air Transport Management,* 92, 102043. doi: https://doi.org/10.1016/j.jairtraman.2021.102043

Rocha, A. C. P. (2020). *Mineração de textos para classificação de processos judiciais trabalhistas* [Dissertação de Mestrado Profissional em Computação Aplicada]. Brasília, Distrito Federal.

Sabo, I. C., Dal Pont, T. R., Wilton, P. E. V., Rover, A. J., & Hübner, J. F. (2022). Clustering of brazilian legal judgments about failures in air transport service: an evaluation of different approaches. *Artificial Intelligence and Law*, *30*(1), 21–57. doi: https://doi.org/10.1007/ s10506-021-09287-3

Scaranello, B. B., Klarmann, J., & de Barros Silva, P. (2022). *Os impactos da cultura da judicialização no setor aéreo.*

Sicsú, A. L., Samartini, A., & Barth, N. L. (2023). *Técnicas de aprendizado de máquina*. Editora Blucher.

Ting, S., Ip, W., Tsang, A. H., et al. (2011). Is naive bayes a good classifier for document classification. *International Journal of Software Engineering and Its Applications*, *5*(3), 37–46.

Torres, G. d. O., Guterres, M. X., & Celestino, V. R. R. (2023). Legal actions in brazilian air transport: A machine learning and multinomial logistic regression analysis. *Frontiers in Future Transportation*, *4*.

Tsangaratos, P., & Ilia, I. (2016). Comparison of a logistic regression and naïve bayes classifier in landslide susceptibility assessments: The influence of models complexity and training dataset size. *Catena*, *145*, 164–179.

Zhang, H., & Zhou, L. (2019). Similarity judgment of civil aviation regulations based on doc2vec deep learning algorithm. In *2019 12th international congress on image and signal processing, biomedical engineering and informatics (cisp-bmei)* (pp. 1–8).