UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS JURÍDICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM DIREITO

Isabela Cristina Sabo

A Machine Learning-based model for judgement results prediction and support in Brazilian Special Court's conciliation hearings

Florianópolis
2022

Isabela Cristina Sabo

A Machine Learning-based model for judgement results prediction and support in Brazilian Special Court's conciliation hearings

Tese submetida ao Programa de Pós-Graduação em Direito da Universidade Federal de Santa Catarina para a obtenção do título de doutora em Direito.
Orientador: Prof. Aires José Rover, Dr.
Coorientador: Prof. Jomi Fred Hübner, Dr.

Florianópolis
2022

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Sabo, Isabela Cristina
A Machine Learning-based model for judgement results prediction and support in Brazilian Special Court's conciliation hearings / Isabela Cristina Sabo ; orientador, Aires José Rover, coorientador, Jomi Fred Hübner, 2022.
150 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro de Ciências Jurídicas, Programa de Pós Graduação em Direito, Florianópolis, 2022.

Inclui referências.

1. Direito. 2. E-Judiciário. 3. Informática Jurídica. 4. Inteligência Artificial. 5. Meios Alternativos de Solução de Conflitos. I. Rover, Aires José . II. Hübner, Jomi Fred. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Direito. IV. Título.

Isabela Cristina Sabo

A Machine Learning-based model for judgement results prediction and support in Brazilian Special Court's conciliation hearings

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Giovanni Sartor, Dr.
Alma Mater Research Institute for Human-Centered Artificial Intelligence
Departamento de Direito
Universidade de Bolonha (Itália)

Prof. Fabiano Hartmann Peixoto, Dr.
Departamento de Direito
Universidade de Brasília

Prof. Marcelo Stemmer, Dr.
Departamento de Automação e Sistemas
Universidade Federal de Santa Catarina

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de doutora em Direito.

ICP-Edu
Documento assinado digitalmente
Claudio Macedo de Souza
Data: 27/09/2022 14:30:04-0300
CPF: ***.565.726-**
Verifique as assinaturas em https://v.ufsc.br

Cláudio Macedo de Souza
Coordenador do Programa

ICP-Edu
Documento assinado digitalmente
Aires Jose Rover
Data: 26/09/2022 18:54:35-0300
CPF: ***.815.549-**
Verifique as assinaturas em https://v.ufsc.br

Prof. Aires José Rover, Dr.
Orientador

Florianópolis, 2022.

To Feres, Fátima, Paulo, Paulinho, Angela, and Maria Clara,
with love.

ACKNOWLEDGEMENTS

The PhD degree achievement was only possible due to a wide support network, which I was privileged to count on during the last four years. You all made this journey more special.

Thanks to my supervisor, Professor Aires José Rover, for being the main encourager of this work; for having a non-traditional and innovative view of Law, which changed mine as well. Special thanks for his solid knowledge in Methodology, which always clarified my research path.

Thanks to my co-supervisor, Professor Jomi Fred Hübner, for kindly receiving me in the Automation and Systems department; for believing in this research and providing me the conditions to conduct it.

Thanks to my daily research (and coffee) partner, Thiago Raulino Dal Pont, for all his programming skills and hard work employed in this project; without his commitment, this thesis would not have happened.

Thanks to Judge Vânia Petermann, for trusting us; for facilitating both our access to the JEC/UFSC data and the application of the model in the conciliation hearings (even with all the bureaucracy this requires).

Thanks to Professors Fabiano Hartmann Peixoto, Marcelo Ricardo Stemmer, José Sérgio da Silva Cristovam, Paloma Maria Santos, Cláudio Macedo de Souza, José Renato Gaziero Cella, and Giovanni Sartor, for their participation in the three evaluation stages of this work (project, preliminary and final defence), as well as for the availability and willingness to review it with relevant contributions.

Thanks to all my colleagues from the EGOV/UFSC research group, through which I learned the importance of teamwork; for introducing me to interdisciplinary and empirical research; particularly, thanks for the friendship, experiences and insights shared in our weekly meetings.

Thanks to Professor Giovanni Sartor, for allowing me to do an internship at a renowned interdisciplinary department on Artificial Intelligence and Law[1]; for his cordiality and openness during our collaboration. Thanks as well to all the researches in the department for the classes and valuable knowledge shared; specially, thanks to my friends from “Mezzanino”, who made this experience a memorable one.

Thanks to National Council for Scientific and Technological Development (CNPq) and Coordination for the Improvement of Higher Education Personnel - Institutional Program for Internationalisation (CAPES/PrInt), for the doctoral and exchange scholarships.

Thanks to the Federal University of Santa Catarina (UFSC), for offering me its structure (public and free) and where I made wonderful friends, making it all worthwhile.

With affection, thanks to my family for their unconditional support and understanding in all my decisions. Special thanks to Feres Azevedo Salem, for being the important bridge in the Automation and Systems department and for introducing me to LaTeX; mainly, for being my life partner and having lovingly, closely and daily accompanied me in this adventure.

Finally, I thank God, because I believe that a higher power has put the right people in the right places and at the right times.

[1] Alma Mater Research Institute for Human-Centered Artificial Intelligence (CIRSFID – AI), University of Bologna, Italy.

*“We can’t solve problems by using the same kind
of thinking we used when we created them.”
(Albert Einstein)*

ABSTRACT

According to the latest "Justiça em Números" (JN) report, published by the National Justice Council (CNJ), the Brazilian Judiciary ended 2020 with 75.4 million lawsuits in progress, which are waiting for a definitive solution. Of these, approximately 3.8 million were new ones filed in the Special Civil Courts (JECs). Due to the high litigation rates, the CNJ has invested in some policies to improve procedural management. Two of these are (i) Alternative Dispute Resolution (ADR), e.g., mandatory conciliation hearings; and (ii) wide use of information technologies (E-Justice), especially those from Artificial Intelligence (AI) domain, e.g., Machine Learning (ML) and Natural Language Processing (NLP) techniques employed in legal texts. However, the development of intelligent solutions in Brazilian Justice faces some limitations, such as the lack of quality of the data produced. Considering this context, motivation, and challenges, our research problem and hypothesis are: Is it possible to apply a ML-based model to predict judgement results in the JECs, supporting the parties and improving the conciliation hearings? Yes, the predictions can empower the litigants to make their own decisions and increase the probability of an agreement. Hence, our objective is to demonstrate that a ML-based model constructed with a JEC database can be useful for the parties in the conciliation hearing, giving them some estimates of judgement results. From a systemic view and inductive approach, we address it through the following methodology and structure. Chapter 2: We employ systematic and narrative literature reviews to investigate knowledge gaps in terms of AI-based ADR/ODR systems. Chapter 3: We conduct a case study in the JEC/UFSC, starting with non-participant observations in local conciliation hearings to diagnose possible causes for the parties not reaching an agreement. Among them, we note that the conciliator does not suggest agreement options, since he or she is not provided with organised information about how the case will be decided. Then, we create a dataset composed of 1163 local judgements on the most recurrent subject of the hearings: consumers' claim for immaterial damage compensation regarding failures on air transport service. We perform different experiments with these legal texts by applying NLP and ML techniques focused on four tasks: (i) clustering to guide the extraction of attributes (judgement factors) and labels (judgement results); (ii) association rules to find relationships between them; (iii) classification to predict the verdict (categorical judgement result); and (iv) regression to predict the amount of immaterial damage compensation (numerical judgement result). Chapter 4: After achieving accurate results in the case study, our proposal is a ML-based model that includes a set of steps and techniques to appropriately prepare the data, find patterns in them, conduct the learning process and apply the output in the legal environment. In the end, we validate the proposed model in real cases through participant observations of the conciliation hearings, occasion on which we present to the parties the judgement possibilities and a voluntary survey questionnaire. The results predicted by the proposed model are well received and appreciated by the parties and their lawyers, and also get close to the real results. Chapter 5: We conclude that the Brazilian Judiciary and society benefit when litigation data is transformed into knowledge and provided to the parties as a way of encouraging self-composition and avoiding new lawsuits. We suggest, as future work, constructing an ODR system based on our model, whereby parties, lawyers, conciliators, and judges have an easy and open access to judgement factors and predictions.

Keywords: E-Justice. Online Dispute Resolution. Machine Learning. Natural Language Processing. Consumer Law. Immaterial Damage.

RESUMO

De acordo com o último relatório “Justiça em Números” (JN), publicado pelo Conselho Nacional de Justiça (CNJ), o Poder Judiciário brasileiro finalizou o ano de 2020 com 75.4 milhões de ações judiciais em tramitação, aguardando solução definitiva. Desses, aproximadamente 3.8 milhões tratam-se de novos processos propostos nos Juizados Especiais Cíveis (JECs). Devido aos altos índices de litigiosidade, o CNJ tem investido em algumas políticas para aprimorar a gestão processual. Duas delas são (i) soluções alternativas de conflito (ADR), e.g., a audiência de conciliação como fase obrigatória do processo judicial; e (ii) uso intensivo de tecnologias de informação (e-Judiciário), especialmente aquelas baseadas em Inteligência Artificial (AI), e.g., técnicas de Aprendizado de Máquina (ML) e Processamento de Linguagem Natural (NLP) aplicadas em textos jurídicos. Todavia, o desenvolvimento de soluções inteligentes na Justiça brasileira enfrenta algumas limitações, como a ausência de governança e estruturação dos dados produzidos. Considerando esse contexto, motivação e desafios, colocamos como problema e hipótese de pesquisa: É possível aplicar um modelo baseado em ML para prever resultados de julgamento nos JECs, dando apoio às partes e qualidade às audiências de conciliação? Sim, as predições podem empoderar os litigantes para tomar suas próprias decisões, aumentando a probabilidade de acordo. Assim, nosso objetivo é demonstrar que um modelo baseado em ML construído a partir de uma base de dados do JEC pode ser útil às partes nas audiências de conciliação, fornecendo-as estimativas sobre o resultado da sentença. A partir de uma visão sistêmica e do método de abordagem indutivo, nós enfrentamos o problema com a seguinte metodologia e estrutura. Capítulo 2: Nós realizamos revisões sistemática e narrativa de literatura para investigar lacunas de conhecimento sobre sistemas de ADR/ODR baseados em AI. Capítulo 3: Nós conduzimos um estudo de caso no JEC/UFSC, iniciando com observações não participativas das audiências de conciliação locais, a fim de diagnosticar possíveis razões pelas quais as partes não alcançam um acordo. Dentre elas, notamos que o conciliador sugere opções de acordo, uma vez que ele não é munido de informações organizadas sobre como será decidido o caso. Após isso, nós criamos uma base de dados composta por 1163 sentenças locais sobre a mais recorrente matéria das audiências: reclamações de consumidores relativas a falhas no serviço de transporte aéreo. Nós realizamos diferentes experimentos com essa base aplicando técnicas de ML e NLP focados em quatro tarefas: (i) clusterização para guiar a extração de atributos (fatores da sentença) e rótulos de classe (resultados da sentença); (ii) associação para encontrar relacionamentos entre eles; (iii) classificação para predizer o veredito da sentença (resultado categórico) e o valor da indenização por dano moral (resultado numérico). Capítulo 4: Após obter resultados acuráveis no estudo de caso, propomos um modelo baseado em ML que inclui uma série de passos e técnicas para preparar os dados de forma apropriada, para encontrar padrões neles, para conduzir o processo de aprendizado e para realizar uma aplicação no ambiente jurídico. Ao final, nós validamos o modelo proposto em casos reais por meio de observações participativas de audiências de conciliação, nas quais apresentamos às partes as possibilidades da sentença e um questionário voluntário e anônimo. Os resultados preditos pelo modelo proposto foram bem recepcionados e apreciados pelas partes (e respectivos advogados), e também se aproximaram dos resultados reais. Capítulo 5: Nós concluímos que o Poder Judiciário brasileiro e a sociedade se beneficiam quando dados de litígios são transformados em conhecimento e fornecidos às partes como forma de encorajar a autocomposição e evitar novas ações judiciais. Sugerimos, como trabalho futuro, a construção de um sistema de

ODR baseado no nosso modelo, por meio do qual as partes, advogados, conciliadores e juízes tenham fácil e aberto acesso aos fatores e predições sobre as sentenças.

Palavras-chave: E-Judiciário. Resolução de Conflitos Online. Aprendizado de Máquina. Processamento de Linguagem Natural. Direito do Consumidor. Dano Moral.

RESUMO EXPANDIDO

Introdução

O Poder Judiciário brasileiro enfrenta atualmente altos índices de litigiosidade e de congestionamento processual. O último relatório “Justiça em Números” (JN), publicação anual do Conselho Nacional de Justiça (CNJ), aponta números crescentes relativos ao ajuizamento de novas ações, das quais significativa parte ocorre nos Juizados Especiais Cíveis (JECs) (CNJ, 2021a). Estes órgãos são responsáveis por julgar causas de menor complexidade e por facilitar o acesso do cidadão à Justiça, através da isenção de custas processuais e dispensa de advogado (BRASIL, 1995). Para minimizar essa crise e aprimorar a gestão processual, o CNJ tem elaborado políticas envolvendo duas agendas: (i) soluções alternativas de conflito (ADR), e.g., incorporando a conciliação e a mediação como fase obrigatória nas ações judiciais; (ii) uso intensivo de tecnologias da informação (e-Justiça), e.g., criando aberturas e incentivos para promover soluções automatizadas baseadas em Inteligência Artificial (AI) (CNJ, 2010, 2021b). Nesta perspectiva, o emprego de técnicas de Processamento de Linguagem Natural (NLP) e Aprendizado de Máquina (ML) possibilitam a predição de decisões judiciais a partir da representação vetorial do texto jurídico processual (ASHLEY, 2022). O desenvolvimento de soluções automatizadas na Justiça brasileira enfrenta, porém, certas limitações, como a ausência de governança e estruturação dos dados produzidos (em grande maioria textuais). Além disso, o JN indica que, embora a incorporação das ADRs no processo judicial, os índices de conciliação ainda são baixos.

Considerando esse contexto, motivação e desafios, colocamos como *problema* e *hipótese* de pesquisa: É possível aplicar um modelo baseado em ML para prever resultados de julgamento nos JECs, dando apoio às partes e qualidade às audiências de conciliação? Sim, as predições podem empoderar os litigantes para tomar suas próprias decisões, aumentando a probabilidade de acordo e minimizando a cultura do litígio no Brasil.

Objetivos

O objetivo *geral* desta pesquisa é demonstrar que um modelo baseado em ML construído a partir de uma base de dados do JEC pode ser útil às partes nas audiências de conciliação, fornecendo-as estimativas sobre o resultado da sentença. Para atingi-lo, traçamos os seguintes objetivos *específicos*:

1. Elaborar revisões de literatura (sistemática e narrativa) para levantar as principais contribuições sobre o uso de AI em ambientes de ADR/ODR (consolidação do estado da arte).
2. Levantar dados no Juizado Especial Cível localizado na Universidade Federal de Santa Catarina (JEC/UFSC), especificamente:
   a) Dados quantitativos e qualitativos de audiências de conciliação (a matéria jurídica discutida, a disposição das partes em celebrar acordo, sugestões do conciliador, resultado da audiência, etc.);
   b) Textos de sentenças relativos à matéria jurídica de maior recorrência no JEC/UFSC.

3. Analisar os dados coletados no item 2.a para identificar as possíveis causas dos baixos índices de conciliação.
4. Aplicas técnicas de ML e NLP nos textos coletados no item 2.b (pré-processamento, representação, clusterização, classificação e regressão) para construir um modelo preditivo de sentenças.
5. Validar o modelo em audiências de conciliação do JEC/UFSC e com as partes nelas envolvidas.

Metodologia

A metodologia de pesquisa se divide em: (i) visão de mundo; (ii) método de abordagem; (iii) procedimentos e técnicas. A visão de mundo (como eu enxergo o problema de pesquisa) adotada é a *sistêmica/autopoiética* (MATURANA; VARELA, 2011; LUHMANN, 1983). O método de abordagem (como eu elaboro a pergunta a partir do problema de pesquisa, de um modo ainda não feito anteriormente) empregado é o *indutivo*. E, para cada objetivo específico, atribuímos diferentes *procedimentos* e *técnicas* (como eu soluciono o problema de pesquisa), conforme Tabela 1.

Table 1 – Procedimentos e técnicas de pesquisa

<table>
<tr><th>Objetivo específico</th><th>Procedimento</th><th>Técnicas</th></tr>
<tr><td>1</td><td>Busca na literatura (pesquisa descritiva)</td><td>Revisões sistemática e narrativa</td></tr>
<tr><td>2</td><td rowspan="4">Estudo de caso</td><td>Observação não-participante e coleta de dados in loco</td></tr>
<tr><td>3</td><td>Visualização de dados</td></tr>
<tr><td>4</td><td>Pré-processamento, representação, clusterização, associação, classificação e regressão de dados (técnicas de ML e NLP)</td></tr>
<tr><td>5</td><td>Observação participante e questionário</td></tr>
</table>

Para implementar as técnicas dos objetivos 3 e 4, utilizamos os softwares Orange 3 (DEMŠAR *et al.*, 2013) e Carrot[2] (OSIŃSKI; WEISS, D., 2019), ambos de código-fonte aberto, bem como a linguagem de programação Python em conjunto com a biblioteca Scikit-Learn, também de código-fonte aberto. Enfatizamos que esta é uma pesquisa *qualitativa*. Nas observações (não-participante e participante, técnicas dos objetivos 2 e 5) não será considerada a quantidade da amostra, e sim os elementos essenciais do ambiente observado (conciliadores, partes, juiz) (TRIVIÑOS, 1987).

Resultados e discussão

Busca na literatura/Pesquisa descritiva (objetivo 1)

Após conduzir as revisões sistemática e narrativa da literatura sobre o tema, verificamos trabalhos que relatam: (i) sistemas de ADR/ODR baseados em AI, em maioria, utilizando técnicas de Representação do Conhecimento (KR); (ii) uso de técnicas de ML para prever decisões e auxiliar no descongestionamento de tribunais, bem como para aprimorar a consistência dos julgados. Contudo, não encontramos trabalhos que relatam soluções/sistemas baseados em ML para prever decisões e, com isso, sugerir opções de acordo para incentivar a solução consensual do conflito (ADR/ODR). Logo, é esta lacuna do conhecimento que a pesquisa pretende explorar.

Estudo de caso (objetivos 2 a 5)

Observação não-participante das audiências de conciliação e coleta de sentenças no JEC/UFSC: Após a observação de 52 audiências, verificamos que um dos óbices ao acordo é a ausência de informações disponíveis aos envolvidos sobre casos anteriores e, como consequência, o conciliador não está apto a sugerir opções de acordo. Além disso, também computamos que a matéria jurídica mais recorrente dessa amostra é Direito do Consumidor, especificamente pedidos de danos morais em razão de falhas no serviço de transporte aéreo. Então, para a construção do modelo baseado em ML, coletamos no local 1163 sentenças publicadas entre Fevereiro de 2011 a Setembro de 2020, específicas sobre o tema.

Clusterização para guiar a extração de fatores e resultados das sentenças: Após a coleta das sentenças (dados textuais não estruturados), aplicamos técnicas de clusterização (aprendizado não supervisionado) para facilitar e reduzir viéses na extração de informações sobre os dados. Técnicas de *soft clustering* apresentaram melhor desempenho nesta tarefa, e.g., os algoritmos de Clusterização Hierárquica e Lingo. Como fatores das sentenças, encontramos, por exemplo, data do julgamento, juiz, tipo de juiz (substituto, titular ou voluntário), atraso de voo (e o intervalo de atraso), cancelamento de voo, extravio temporário de bagagem (e o intervalo de extravio), extravio definitivo de bagagem, exercício do direito de arrependimento, prática de *overbooking*, ocorrência de *no show*, entre outros. Como resultados da sentença, extraímos o veredito do juiz (procedência, procedência parcial, improcedência, extinção sem análise do mérito) e o valor da indenização por dano moral fixado pelo juiz.

Associação para descobrir relacionamentos entre os fatores e os resultados das sentenças: Após a extração de informações e estruturação dos dados, novamente utilizamos aprendizado não supervisionado para descobrir possíveis relações entre os fatores e os resultados da sentença. Aplicamos o algoritmo de associação *FP-Growth* para esta tarefa, e confirmamos a relação entre o intervalo de atraso e o valor da indenização por dano moral, isto é, quanto mais uma companhia aérea atrasa um voo do consumidor, mais alta é a indenização, e vice-versa. Também há relação entre o extravio definitivo de bagagem e a faixa maior de indenização, enquanto o extravio temporário com intervalo de 3 dias resultará em uma faixa média de indenização. Verificamos que a prática de *no show* também está associada à esta faixa. A indenização igual a zero (ausência de dano moral) está relacionada ao intervalo de atraso inferior a 4 horas e à culpa exclusiva do consumidor.

Classificação para prever o resultado categórico da sentença: Após utilizar aprendizado não supervisionado para estruturar os dados, recorremos ao aprendizado supervisionado, especificamente à tarefa de classificação, para fazer predições sobre o veredito da sentença (procedência, procedência parcial, improcedência, extinção sem análise do mérito) que é uma variável categórica. Após os experimentos, obtivemos um melhor desempenho adotando a classificação binária (junção das classes procedência e procedência parcial; e das classes improcedência e extinção sem análise do mérito), e com os algoritmos de aprendizado clássico, Regressão Logística e Redes Neurais, atingindo, dentre outras métricas de avaliação, uma acurácia de 92% e 91,4%, respectivamente.

Regressão para prever o resultado numérico da sentença: Recorremos à tarefa de regressão para fazer predições sobre o valor da indenização por dano moral, que é uma variável numérica. Após os experimentos, obtivemos um melhor desempenho com a abordagem *Ensemble Voting*, baseada nos algoritmos *Bagging*, *GBoost*, *XGBoost* e Redes Neurais, atingindo, dentre outras métricas de avaliação, um erro médio absoluto (MAE) de 915. Isso significa que o erro do modelo em relação ao valor predito da indenização é de aproximadamente R$ 1.000,00.

Observação participante das audiências de conciliação e questionário às partes envolvidas (validação do modelo): Considerando os resultados satisfatórios na experimentação, propomos um modelo baseado em ML. Nós o validamos em 13 novos casos do JEC/UFSC submetidos à audiência de conciliação, apresentando às partes envolvidas as predições dadas pelo modelo em cada processo judicial, com ênfase no resultado relativo ao valor da indenização por dano moral. Para auxiliar na compreensão das predições e servir como uma forma de explicação, também apresentamos os fatores relacionados às respectivas faixas de indenização (conforme observado na tarefa de associação). De modo geral, os resultados preditos pelo modelo foram bem recepcionados e apreciados pelas partes (e respectivos advogados), assim como se aproximaram dos resultados reais.

Considerações finais

O Poder Judiciário brasileiro e a sociedade se beneficiam quando dados de litígios são transformados em conhecimento e fornecidos às partes como forma de encorajar a autocomposição e evitar novas ações judiciais. Em relação à pergunta de pesquisa, podemos concluir afirmativamente que é possível aplicar um modelo baseado em ML para prever os resultados do julgamento nos JECs. Além disso, estas previsões aumentam a probabilidade de acordo, melhoram a qualidade das audiências de conciliação e são úteis para as partes litigantes. Com isto, confirmamos a hipótese e cumprimos o nosso objetivo geral.

As contribuições da tese podem ser definidas como acadêmicas e práticas. A contribuição acadêmica corresponde ao guia, o passo-a-passo que um pesquisador do Direito deve seguir para auxiliar no desenvolvimento de uma solução de AI para problemas jurídicos. Isto inclui: (i) que tipo de dados e como coletá-los (dados de texto de decisões judiciais, de preferência em formato TXT); (ii) como extrair informações dos dados e organizá-las (fatores e resultados de decisões judiciais); e (iii) quais técnicas aplicar para transformá-los em conhecimento (usando ferramentas que possam ser compreendidas pelos estudantes e pesquisadores do Direito - como Orange 3 e Carrot[2]). E ainda, (iv) como aplicar este conhecimento gerado sem violar a autonomia das partes e o processo legal. A contribuição

prática refere-se aos produtos resultantes da pesquisa: (i) uma base de dados tanto textuais (arquivos TXT) quanto estruturados (arquivo XLS) específicos sobre falhas de serviços de transporte aéreo (Direito do Consumidor), pertencentes ao JEC/UFSC; (ii) um modelo baseado em ML que fornece previsões precisas sobre os resultados das sentenças do JEC/UFSC, e que aborda quatro tarefas de ML: clusterização, associação, classificação e regressão. Em vias de trabalhos futuros, sugerimos a construção de um sistema de ODR baseado no modelo proposto, por meio do qual as partes, advogados, conciliadores e juízes tenham fácil e aberto acesso aos fatores e predições sobre as sentenças.

Todavia, é necessário observar que as técnicas de ML utilizadas para prever os resultados das sentenças são incapazes de explicá-los. Extraímos fatores e a buscamos encontrar uma relação entre eles e os resultados visando fornecer uma explicação para as partes. Portanto, a explicação foi dada, porém de modo limitado, através da nossa intervenção, e não automaticamente.

Palavras-chave: E-Judiciário. Resolução de Conflitos Online. Aprendizado de Máquina. Processamento de Linguagem Natural. Direito do Consumidor. Dano Moral.

LIST OF FIGURES

Figure 1 – Learning process (pipeline) 36
Figure 2 – Brazilian Judiciary organisation chart 45
Figure 3 – (Non-participant observation) Q1. Did all parties attend the conciliation hearing? 47
Figure 4 – (Non-participant observation) Q2. Did the parties have a lawyer? . . . 47
Figure 5 – (Non-participant observation) Q3a. What is the legal area? 48
Figure 6 – (Non-participant observation) Q3b. What is the Consumer Law area? . 48
Figure 7 – (Non-participant observation) Q4. Did the conciliator suggest any form of agreement? 49
Figure 8 – (Non-participant observation) Q5. Did the conciliator refer to previous similar cases to guide the parties? 49
Figure 9 – (Non-participant observation) Q6. Were the parties interesting to settle? 50
Figure 10 – (Non-participant observation) Q7. What was the outcome of the conciliation hearing? 50
Figure 11 – (SVM) Which is the best separation hyperplane? 59
Figure 12 – Logistic Regression versus Multiple Regression 59
Figure 13 – Random Forest 61
Figure 14 – Perceptron 62
Figure 15 – (Hierarchical clustering) Partial dendrogram in Orange 3 69
Figure 16 – (Hierarchical clustering) C1 Word cloud 70
Figure 17 – (Hierarchical clustering) C47 Word cloud 71
Figure 18 – (K-means) Two-dimensional representation of documents and their clusters in Orange 3 72
Figure 19 – (K-means) C1 Word cloud 73
Figure 20 – Lingo clustering in Carrot² 74
Figure 21 – Affinity Propagation clustering 75
Figure 22 – (Affinity Propagation) C29 Word cloud 76
Figure 23 – (Affinity Propagation) C52 Word cloud 77
Figure 24 – Results from 4-class classification 84
Figure 25 – Results from 2-class classification 85
Figure 26 – Results from regression baseline 90
Figure 27 – Results from regression full pipeline 91
Figure 28 – ML-based model proposed 96
Figure 29 – (Participant observation) Q1. What is your position in the case submitted for the conciliation hearing? 100
Figure 30 – (Participant observation) Q2. Did you reach an agreement during the conciliation hearing? 101

Figure 31 – (Participant observation) Q3. If the previous answer was “no”, what reason(s) would you attribute to the failure of the conciliation? 101
Figure 32 – (Participant observation) Q4. If the previous answer was “yes” or “partial”, what reason(s) would you attribute to the success of the conciliation? 102
Figure 33 – (Participant observation) Q5. Was the information presented by the conciliator about local judgements useful to you? 103
Figure 34 – (Participant observation) Q6. Were the values predicted of compensation for immaterial damage helpful to you? 103
Figure 35 – (Participant observation) Q7. Could you please measure your overall level of satisfaction with the activity provided? 104
Figure 36 – SLR path 127
Figure 37 – (SLR) Number of documents per year 133
Figure 38 – (SLR) Number of documents per author 133
Figure 39 – (SLR) Number of documents per university/country 134
Figure 40 – (SLR) Word cloud 135

LIST OF TABLES

Table 1 – Procedimentos e técnicas de pesquisa 12
Table 2 – Research procedures and techniques 27
Table 3 – Mediation and conciliation stages 32
Table 4 – ODR generations 33
Table 5 – Clustering setup parameters 68
Table 6 – (Hierarchical clustering) C1 Topic keywords 70
Table 7 – (Hierarchical clustering) C47 Topic keywords 71
Table 8 – (K-means) C1 Topic keywords 73
Table 9 – (Affinity Propagation) C29 Topic keywords 76
Table 10 – (Affinity Propagation) C52 Topic keywords 77
Table 11 – Clustering evaluation overview 77
Table 12 – Association rules ranges 81
Table 13 – Association setup parameters 81
Table 14 – Association rules selected 82
Table 15 – Classification setup parameters 83
Table 16 – Results from 4-class classification 84
Table 17 – Results from 2-class classification 84
Table 18 – Confusion matrix for Logistic Regression (4-class) 85
Table 19 – Confusion matrix for Naive Bayes (4-class) 86
Table 20 – Confusion matrix for Logistic Regression (2-class) 86
Table 21 – Confusion matrix for Naive Bayes (2-class) 86
Table 22 – Confusion matrix for SVM-RBF (2-class) 87
Table 23 – Regression setup parameters (baseline) 88
Table 24 – Regression setup parameters (full pipeline) 89
Table 25 – Results from regression baseline 90
Table 26 – Results from regression full pipeline 90
Table 27 – ML tasks vs. ML stages and techniques (overview) 93
Table 28 – (Participant observation) Q8. Would you like to leave a comment, critic or praise? 104
Table 29 – Predicted judgement results x Real judgement results 105
Table 30 – SLR search strategies 129
Table 31 – SLR selected papers 130

LIST OF ABBREVIATIONS AND ACRONYMS

ADAM	Adaptive Moment Estimation
ADR	Alternative Dispute Resolution
AI	Artificial Intelligence
B2B	Business-to-Business
BATNA	Best Alternative to a Negotiated Agreement
BERT	Bidirectional Encoder Representations from Transformers
BGB	German Civil Code
BOW	Bag of Words
CAPES	Coordination for the Improvement of Higher Education Personnel
CLAUDETTE	Automated Detector of Potentially Unfair Clauses in Online Terms of Service
CNJ	National Justice Council
CNN	Convolution Neural Networks
CNPq	National Council for Scientific and Technological Development
CT	Clustering Tendency
CUAD	Contract Understanding Atticus Dataset
DAI	Distributed Artificial Intelligence
DCENTR	Decentralised Construction Enabling Transparent Resolution
DL	Deep Learning
DM	Data Mining
EGOV	E-government, Digital Inclusion and Knowledge Society
eJRM	electronic Justice Relationship Management system
eJRM-IRS	electronic Justice Relationship Management Information Retrieval system
ETC	Extra Trees Classifier
FP-Growth	Frequent Pattern Growth
GBoost	Gradient Boosting
HCI	Human Computer Interaction
IDF	Inverse Document Frequency
IMODRE	Integrated Online Dispute Resolution Environment
IR	Information Retrieval
JEC	Special Civil Court
JEF	Federal Special Court
JN	“Justiça em Números”
JPES	Justice Pathway Expert System
kNN	k Nearest Neighbours
KR	Knowledge Representation
LDA	Latent Dirichlet Allocation

LEDGAR	Large-Scale Multi-label Corpus for Text Classification of Legal Provisions in Contracts
LIME	Local Interpretable Model-agnostic Explanations
LR	Logistic Regression
LSI	Latent Semantic Indexing
MAE	Mean Absolute Error
MANN	Memory-Augmented Neural Network
MAS	Multiagent System
MCQ	Multiple Choice Question
ML	Machine Learning
NB	Naive Bayes
NDSS	Negotiation Decision Support System
NLP	Natural Language Processing
NN	Neural Network
NSE	Negative Squared Error
OCR	Optical Character Recognition
ODR	Online Dispute Resolution
PC	Principal Component
PCA	Principal Component Analysis
PrInt	Institutional Program for Internationalisation
RBF	Radial Basis Function
ReLU	Rectified Linear Unit
RF	Random Forest
RG	General Repercussion
RMSE	Root Mean Square Error
SAQ	Short Answer Question
SGD	Stochastic Gradient Descent
SHAP	SHapley Additive exPlanations
SLR	Systematic Literature Review
SO	Specific Objective
SS	Silhouette Score
STF	Federal Supreme Court
STJ	Superior Court of Justice
STM	Superior Military Court
SVD	Singular Value Decomposition
SVM	Support Vector Machine
TADRS	Traffic-Accidents Dispute-Resolution System
TF	Term Frequency
TJ	State Court

TM	Text Mining
TRE	Electoral Regional Court
TRF	Federal Regional Court
TRT	Labour Regional Court
TSE	Superior Electoral Court
TST	Superior Labour Court
UFSC	Federal University of Santa Catarina
UMCourt	University of Minho Court
VSM	Vector Space Model
XAI	Explainable Artificial Intelligence
XGBoost	Extreme Gradient Boosting

CONTENTS

1 INTRODUCTION . . . 24
1.1 CONTEXT AND RESEARCH MOTIVATION . . . 24
1.2 RESEARCH PROBLEM, HYPOTHESIS AND OBJECTIVES . . . 25
1.3 METHODOLOGY . . . 26
1.4 ORIGINALITY . . . 28
1.5 WORK STRUCTURE . . . 29
2 LITERATURE REVIEW . . . 30
2.1 BASIC CONCEPTS . . . 30
2.1.1 Alternative Dispute Resolution . . . 30
2.1.2 Artificial Intelligence . . . 34
2.2 RELATED WORKS AND STATE-OF-THE-ART . . . 37
2.2.1 Artificial Intelligence in Alternative (Online) Dispute Resolution . . . 37
2.2.2 Machine Learning in Supreme Courts and other legal applications . . . 41
2.2.3 Considerations on the section . . . 44
3 CASE STUDY . . . 45
3.1 THE SPECIAL CIVIL COURT . . . 45
3.2 NON-PARTICIPANT OBSERVATION OF HEARINGS . . . 46
3.3 DATASET CONTEXT AND CONSTRUCTION . . . 51
3.4 MACHINE LEARNING STEPS AND TECHNIQUES . . . 53
3.4.1 Preprocessing . . . 53
3.4.2 Representation . . . 54
3.4.3 Clustering . . . 55
3.4.3.1 Similarity . . . 56
3.4.3.2 Visualisation . . . 57
3.4.4 Association . . . 57
3.4.5 Classification . . . 58
3.4.6 Regression . . . 62
3.4.7 Training and test . . . 63
3.4.8 Evaluation . . . 63
3.5 EXPERIMENTS, RESULTS AND DISCUSSION . . . 67
3.5.1 Technical support . . . 67
3.5.2 [A] Clustering to guide the judgement factors extraction . . . 68
3.5.3 [B] Association to find relationships between judgement factors and results . . . 78
3.5.4 [C] Classification to predict the categorical judgement result . . . 82
3.5.5 [D] Regression to predict the numerical judgement result . . . 87
3.5.6 Considerations on the section . . . 91

4 PROPOSAL AND VALIDATION **94**
4.1 THE MACHINE LEARNING-BASED MODEL PROPOSED 94
4.2 VALIDATIONS, RESULTS AND DISCUSSION 97
4.2.1 Participant observation of hearings and survey to the parties involved 97
4.2.2 Comparison between predicted and real judgement results of hearing cases **105**
5 CONCLUSION **107**
5.1 FINAL REMARKS AND CONTRIBUTIONS 107
5.2 LIMITATIONS 108
5.3 FUTURE WORK 109
5.4 PUBLICATIONS 109
REFERENCES **111**
APPENDIX A – SYSTEMATIC LITERATURE REVIEW **127**
APPENDIX B – CLUSTERS DESCRIPTIONS OF HIERARCHICAL CLUSTERING **136**
APPENDIX C – CLUSTERS DESCRIPTIONS OF K-MEANS **138**
APPENDIX D – CLUSTERS DESCRIPTIONS OF LINGO **139**
APPENDIX E – CLUSTERS DESCRIPTIONS OF AFFINITY PROPAGATION **141**
APPENDIX F – (JEC/UFSC) EXPLANATORY MATERIAL PRESENTED TO THE PARTIES DURING THE PARTICIPANT OBSERVATION OF CONCILIATION HEARINGS **144**
APPENDIX G – (JEC/UFSC) SURVEY **146**
ANNEX A – (JEC/UFSC) RESEARCHER'S AUTHORISATION RELATED TO THE NON-PARTICIPANT OBSERVATION OF THE CONCILIATION HEARINGS **148**
ANNEX B – (JEC/UFSC) RESEARCHER'S ASSIGNMENT AS CONCILIATOR RELATED TO THE PARTICIPANT OBSERVATION OF THE CONCILIATION HEARINGS **149**

1 INTRODUCTION

1.1 CONTEXT AND RESEARCH MOTIVATION

According to the latest “Justiça em Números” (JN) report, an official publication of the National Justice Council (CNJ), a legal controlling agency, the Brazilian Judiciary ended 2020 with 75.4 million lawsuits in progress, which are waiting for a definitive solution. About 27.9 million lawsuits were filed in 2020, from which approximately 3.8 million were in the Special Civil Courts (CNJ, 2021a).

Considering this high litigation, since its creation instituted by Constitutional Amendment No. 45/2004 (BRASIL, 2004), the CNJ has invested in some policies to improve the procedural management. Two of these are *alternative dispute resolution* (ADR) and the use of *information technology* (E-Justice).

Self-composition was already incorporated as a mandatory initial hearing in Labour Justice, Federal Justice and State Justice, being mainly emphasised in the Special Civil Courts, which have the specific purpose of promoting conciliation in less complex cases (BRASIL, 1995).

However, it was in 2006 that ADR became a policy of the Brazilian Judiciary through the Conciliation Movement, later regulated by CNJ Resolution No. 125/2010. The policy objectives include: (a) promote legal agencies specialised in conciliation and mediation services; (b) train judges, judicial servers, mediators, conciliators, among other professionals on the subject; (c) develop cooperation between agencies and universities to encourage the culture of peace dispute resolution; (d) reduce the Judiciary congestion (CNJ, 2010).

The regulation brought significant advances on ADR, such as the obligation of the courts to institute internal centres with professionals and services for the population focused on conciliation and mediation (BRASIL, 2015a), and the creation of "Resolve" program, whose goal is to promote conciliation in cases related to social security, tax and consumer problems (CNJ, 2018).

Even with all that effort, the caseload in Brazil is still increasing and the culture of peace shows slow evolution since 2006. Only 9.9% of the total ongoing cases were solved through conciliation in the Brazilian Judiciary during 2020. Furthermore, there were reductions or slight increases in the same rate compared to the previous years (12.5% in 2019, 12.7% in 2018, 13.5% in 2017 and 13.6% in 2016) (CNJ, 2021a).

Regarding E-Justice, a legal landmark is Law No. 11.419/2006, whose scope is to materialise the electronic judicial process and to provide the use of electronic means for all procedural fields (civil, criminal, labour, tax, etc.), also for transmission and communication of procedural acts and for digital signature (BRASIL, 2006).

After that, due to their administrative autonomy, computerisation was happening fast in some courts and late in others. Both the public and private sectors began to develop different electronic process systems, leaving it up to each court to choose which one to adopt and when (SANTOS *et al.*, 2020).

Recently, understanding that Artificial Intelligence (AI), especially the Machine Learning (ML) and Natural Language Processing (NLP) domains, can contribute to the celerity and consistency of the decision-making process, the CNJ has been encouraging and regulating its use in the Brazilian Judiciary (CNJ, 2020c, 2020a). It has also instituted committees to advancing innovation management and produce a high impact on the court's productivity (CNJ, 2021b).

Today, the research paradigm in AI and Law has largely shifted. NLP techniques and ML can predict legal outcomes and classify legal texts statistically, directly from quantitative representations of the texts as vectors, that is, as a series of numbers. Unlike knowledge-based approaches, AI can make predictions without identifying elements of legal rules, issues, factors, values, or other kinds of legal knowledge. It can base its predictions on patterns and frequencies of words (ASHLEY, 2022).

However, the development of intelligent solutions in the Brazilian Judiciary faces some limitations. One of them is the lack of management and quality of the data produced, considering the decentralisation of the electronic process systems. To solve this problem without undermining the autonomy of the courts, the CNJ has employed efforts to create a unified base of procedural data from their submission by the courts (CNJ, 2020b).

In this context, the research is motivated to find a solution that combines AI and ADR. A means by which technology helps people to solve their conflicts by themselves, with procedural data transformed into knowledge to support the conciliation. And as a consequence, that contributes to reducing the litigation and slowness of the Brazilian Judiciary.

1.2 RESEARCH PROBLEM, HYPOTHESIS AND OBJECTIVES

Considering the context and the challenges that motivate this research, we present the problem (defined as a question), the hypothesis and the objectives (general and specific) that sets the thesis:

Research problem: Is it possible to apply a ML-based model to predict judgement results in the Special Civil Courts, supporting the parties and improving the conciliation hearings?

Hypothesis: Yes. Since the parties do not have information about the judgement possibilities, the predictions can empower them to make their own decisions and increase the probability of an agreement.

Thus, the *general objective* of this research is to demonstrate that a ML model constructed with the local database can be useful for the parties in the conciliation hearing. To achieve this, we set *specific objectives* (SO) that we considered to be stages of the research:

1. Elaborate literature reviews (systematic and narrative) to understand the AI fields and to survey the contributions on AI usage in ADR environments (state of the art).
2. Collect data into the Special Civil Court of Federal University of Santa Catarina (JEC/UFSC), specifically:
    a) Data from conciliation hearings (the legal subject discussed, the willingness of the parties to agree, conciliator's agreement suggestions and hearing's result, etc.);
    b) Data from judgements about one of the most discussed legal subject in the hearings.
3. Analyse the data collected in 2.a to identify the causes of the low conciliation rates.
4. Apply ML an NLP techniques in the data collected in 2.b (preprocessing, representation, clustering, classification and regression) to construct a model for judgements predictions.
5. Validate the model in the JEC/UFSC conciliation hearings and with the parties involved in them.

1.3 METHODOLOGY

The research methodology is divided into three elements: (1) *world perspective*: how do I see the research problem? (2) *approach method*: how do I ask the question from the research problem in a way that no one else has done before? and (3) *procedures and techniques*: how do I solve the research problem? (verbal information).[2]

As *world perspective*, we adopt the systemic view according to the elements of the theory of Maturana and Varela (2011), namely "organisation", "structure" and "autopoiesis".

"Organisation" of something consists of the relations that must occur for that something to exist. For example, for a chair to be so judged, one must recognise the relationships between its feet, backrest and seat, so that it is possible to sit in it. Whether the chair is made of plastic or wood does not change the fact that it would still be a chair. Then the "structure" corresponds to the components that define the organisation, which can be changed as the environment requires. "Autopoiesis" is the movement that characterises the autonomy of an organisation, which means that it can determine itself (MATURANA; VARELA, 2011).

These concepts are visible in biology, considering the living being as an organisation with the capacity for autopoiesis, which can support structural changes to adapt to

[2] Division instructed by Prof. Aires José Rover in the subject "Research Projects in Law" of the Federal University of Santa Catarina Law School.

the environment disturbs. However, they can also perform in sociology, in social systems such as the Law. Law emerges as an element of society's development process and grows by adapting itself to social needs. These needs point to greater complexity and variability that enriches society with possibilities (LUHMANN, 1983).

Based on this theory, we consider the Brazilian Judiciary as an organisation whose structure can be changed from the inside out due to autopoiesis dynamics. Observing the Brazilian Judiciary from the perspective of the systemic view allows us to identify to what extent human and technological elements lead to changes in its structure for better. Thus, it will be possible to visualise future possibilities for restructuring the judicial system (SARDETO, 2017).

We understand that AI usage can optimise this dynamics with positive structural changes in the organisation, reflecting a higher quality of Brazilian Judiciary service. Moreover, using the data from its caseload to provide knowledge to the parties, which can be useful for them to end the process themselves, is an inside-out solution that strengthens the organisation's autonomy.

As *approach method*, we adopt the inductive. In this way, we achieve more ample conclusions when compared to the initial premises that were the basis for the research. This generalisation occurs by identifying the problem from a social phenomenon (MEZZAROBA; MONTEIRO, 2009).

As *procedures and techniques*, we adopted those described in Table 2 for each specific objective.

Table 2 – Research procedures and techniques

Specific objective	**Procedure**	**Techniques**
1	Literature search	Narrative and systematic review
2	Case study	Non-participant observation and data collection on site
3		Data visualisation
4		Data preprocessing, representation, clustering, association, classification and regression (ML and NLP techniques)
5		Participant observation and survey questionnaire

Finally, we emphasise that this is *qualitative* research. The observations (non-participant and participant) will not consider the quantity of the sample but rather the essential elements of the environment (conciliators, parties, judge) (TRIVIÑOS, 1987).

1.4 ORIGINALITY

The originality of the research is demonstrated by the results of the systematic literature review (SLR). This type of review is a comprehensive summary of primary research on a specific question that attempts to identify, select, synthesise, and appraise all high-quality evidence relevant to answer it. Additionally, SLR identify themes that need evidence and future investigations, minimising bias via transparent and explicit methodology (SAMPAIO; MANCINI, 2007; HARRIS *et al.*, 2014).

We conducted a SLR during in April 2019 (at the beginning of the doctoral course), further updated in July 2022 (at the end of the course). The details of the review are described in Appendix A, and here we will present what we have concluded from it.

Starting from the research question *What are the advances in science on the application of AI techniques in ADR?*, we found systems, experiments and proposed ODR solutions based on several AI fields. We observed that in the majority of the papers synthesised the proposals are based on Knowledge Representation (such as Ontology, Case-Based Reasoning, Rule-Based Systems and Expert Systems) (LODDER; ZELEZNIKOW, 2012; CARNEIRO *et al.*, 2013a, 2014; FERSINI *et al.*, 2014; EL JELALI *et al.*, 2015; CAPUANO *et al.*, 2015; THOMPSON, 2015; CAPUANO; TOTI, 2019). We also found research related to Multi-Agent Systems (CARNEIRO *et al.*, 2011; ABRAHAMS *et al.*, 2012; CARNEIRO *et al.*, 2014), Genetic Algorithms (CARNEIRO *et al.*, 2013b; SIMKOVA; SMUTNY, 2021), Game Theory (GOMES *et al.*, 2014; CARNEIRO *et al.*, 2017) and Blockchain (SAYGILI *et al.*, 2022). Machine Learning and Natural Language Processing was less explored in ADR domain. Its usage was limited to the classification of legal areas to which a conflict belongs for decisions suggesting (FERSINI *et al.*, 2014; EL JELALI *et al.*, 2015), as well as outcome predictions without real application (TSUREL *et al.*, 2020).

To the best of our knowledge, we did not find works that use predictions about outcomes and compensation values of ongoing lawsuits to assist in ADR, by presenting to the parties what is likely to happen with their case. Furthermore, we did not locate institutional efforts on ADR and AI in the Brazilian Judiciary context.

Moreover, in view of the Brazilian Judiciary scenario, where text data is produced on a large scale as litigation grows, exploring ML and NLP in this domain suggests more innovation and usefulness. In addition, considering that legal knowledge is constantly modified by legislation and by court precedents, updating its representation in AI systems and creating rules for each judgement possibility may be unfeasible. Ultimately, since the Brazilian Judiciary incorporated ADR to reduce litigation, we understand that modelling court decisions and predicting results from them as a solution option will provide autonomy and safety to the parties for self-composition. Within this purpose, we can finally reach the culture of pacification. That is the research originality.

1.5 WORK STRUCTURE

We organised this work as follows.

In Chapter 2, to fulfil SO 1, we introduce some concepts to context the two domains that this research crosses: ADR and AI, including an overview of the techniques from each branch (even though in this work we have focused on conciliation and ML). We also describe related work to our research JECs. First, we discuss the papers that resulted from our SLR about intelligent applications to enhance alternative forms of dispute resolution. Second, we expose ML and NLP initiatives in Supreme Courts around the world and other similar contributions to the legal field.

In Chapter 3, we present our case study developed into the JEC/UFSC. First, to comply with SO 2.a and 3, we show through non-participant observation how are the dynamics of conciliation hearings, in particular the obstacles around reaching agreements (diagnosis of the application environment). Second, to carry out SO 2.b and 4, we context and justify our dataset, as well specify the ML and NLP techniques used on it (preprocessing, representation, clustering, association, classification, regression, among others). Then, we describe our experiment setups, show and discuss the results, evaluating the advantages and disadvantages of each approach.

In Chapter 4, to fulfil SO 5, we propose our ML-based model and validate it through participatory observations, presenting and explaining the predictions to the parties in the JEC/UFSC conciliation hearings. Then, we gather responses from the parties and their lawyers about how useful this information is in the negotiation process. We also compare the predicted results with the real results of the hearing cases.

Finally, there are our concluding remarks, contributions and limitations, as well new perspectives of study and projects in Chapter 5.

2 LITERATURE REVIEW

2.1 BASIC CONCEPTS

2.1.1 Alternative Dispute Resolution

Societies usually focus on three bases to dispute resolution: (1) rights-based, (2) power-based and (3) interest-based. The legal system adopts the first under the premise that we have rights protected by laws. Thus, when party A believes that Party B has breached its contract, party A will file a lawsuit to vindicate its rights. One party will be announced by a court as a winner, while the other one as the loser. The second occurs when one party wielding power against another, for example, the employer who cut or refuse to increase the salary of a disfavoured employee. The third emphasises the party's interests in response to the dissatisfaction with solving conflicts through rights or power-based. Interests are those that move the parties, the desires and concerns that underlie the positions they take. Consequently, the solution will accommodate the interests of both, without a winner and a loser. It is what the ADR movement seeks (DONEFF; ORDOVER, 2014).

More than that, the ADR movement has been successful at transforming the legal system. Recent court decisions suggest that the jurists should embrace a vision whose purpose of legal dispute resolution is to achieve social harmony rather than assess factual and legal claims and articulate public norms (HENSLER, 2003).

Lieberman and Henry (1986) define ADR as a set of practices and techniques that aim (1) to permit legal disputes to be resolved outside the courts for the benefit of all disputants; (2) to reduce the cost of conventional litigation and the delays to which it is ordinarily subject; or (3) to prevent legal disputes that would otherwise likely be brought to the court. The ADR roster includes arbitration, mediation, conciliation and negotiation. Changes in procedural rules to provide incentives to the parties to settle and the greater use of partial summary judgement might also be viewed as ADR techniques.

Historically, the movement within the courts gained strength in the United States in the 1980s. What began as an experiment to solve family disputes become the most significant change in civil practice at the time. Now litigants are seeking ADR on their initiative, courts are making ADR available to litigants who request it, and many state and federal courts have implemented mandatory ADR program (ROSENBERG; FOLBERG, 1994).

In Brazil, the movement became important after the 2008 crisis under two trends: the "de-judicialization" and the "proceduralization" of Law. The first is a process of de-formalisation, de-legalisation, and de-constitutionalisation of rights and the creation of alternative means for solving conflict, which usually occurs in parallel with the rupture of state monopolies and the abdication of the power of regulation or interference in the setting of prices, wages, and working conditions by the government. And the second is the State no longer deciding the content of the laws, limiting to establishing procedures in that the different social sectors can discuss and negotiate the normative alternatives that are most appropriate to their respective interests. Instead of making unilateral decisions and imposing them on citizens, companies, associations, and social movements, the legislator thus opts for a negotiated creation of the law (FARIA, 2011).

As introduced in section 1.1, Brazilian Judiciary incorporated the ADR movement as policy in 2010, focusing on *mediation* and *conciliation*. It is these two types of ADR that we will explain in detail below.

Mediation is the informal inclusion of an impartial, neutral third party (without any decision-making power) to assist, facilitate, and encourage those involved in a conflict to achieve a friendly and acceptable solution. It is a procedure facilitated by the intervention of a third party (CALMON, 2007).

Besides dealing with conflict, mediation can also establish or strengthen trusting and respecting relationships between the parties. Or even end relationships in a way that minimises costs and psychological damage. The mediator may alter the social dynamics of the relationship that is the subject of the conflict by influencing the beliefs or behaviours of the individual parties by providing knowledge and information (MOORE, 1998).

Conciliation, in its turn, is a procedure whose purpose is to achieve self-composition also with the help, encouragement and facilitation by an impartial third party, emphasising that self-composition is the prevention or resolution of conflicts practised by the parties involved themselves (CALMON, 2007).

Thus, while conciliation is the means, self-composition is the result. These procedure supposes the agreement between conflicting interests and corresponds to the harmony established between two or more people with dissenting positions. There is an intention to solve the problem peacefully (GOZAÍNI, 1995).

Although the terms are used interchangeably, the concepts allow us to verify a distinction: while mediation focuses on re-establishing communication between the parties, recovering the relationship between them, conciliation aims the self-composition, the agreement. Brazilian Judiciary generally uses mediation for family law and conciliation for contract/civil law.

Conciliation and mediation have common aspects: (1) financial and time economy since conflicts are solved in a less costly way and in less time compared to the Judiciary; (2) oral proceedings, or informality, since it is a moment when the parties freely have the opportunity to discuss the problems that involve them; (3) the autonomy of the decision, which does not need to be approved by the Judiciary, except in cases of bad faith or violation of law; (4) confidentiality, except by the parties desire or by public interest (for example, when a crime is involved) (MORAIS; SPENGLER, 1999).

We detail the stages and actions of mediation and conciliation in Table 3.

Table 3 – Mediation and conciliation stages

Stage	**Mediator and Conciliator actions**
Opening	(1) Prepare the environment; (2) Greet the parties and introduce oneself; (3) Explain the procedure to be used, its methods, characteristics and advantages; (4) Delimit the conciliator/mediator role and its limits; (5) Clarify that communication must be organised and non-violent (without interruption).
Understanding the conflict	(1) Encourage the parties to talk about the conflict; (2) Identify the parties' interests. * In mediation, it is appropriate to: (a) Hold individual sessions with the parties; (b) Active listening with empathy (rapport); (c) Reverse the roles, putting one party in the place of the other (mirror); (d) Normalise (affirm to the parties that the conflict situation is common and should be understood as an opportunity to improve the relationship); (e) Affection (positive response to good behaviour by the party or the lawyer).
Proposals and counterproposals	(1) Encourage cooperative behaviour; (2) Motivate the parties to brainstorm with mutual gain. * In conciliation, it is appropriate to create settlement options and present them to the parties.
Agreement	a) Supervise that the agreement was not concluded under pressure; b) Draft the term together with the parties, prioritising clarity and specificity, to avoid later disagreements; c) Guide the parties, if possible, to solve predictable future problems.

Both procedures can be out-of-court and in-court. Mediation out-of-court is performed by a mediator or private chamber. The parties may choose the mediator and how they wish to conduct it. They can initiate the procedure by a simple invitation from one party to another or a contractual provision. Court mediation, in turn, is hearings conducted by a nominated mediator, with the choice limited to the list of mediators registered in the court (BRASIL, 2015b).

Conciliation out-of-court occurs without a lawsuit filed. Afterwards, the agreement resulting from this conciliation is submitted to the Judiciary for approval. Even so, it is considered non-court, because the procedure was not developed in a judicial environment. Court conciliation, in turn, is hearings at the beginning of a legal process or during its course (CALMON, 2007).

Finally, an emerging and important category in ADR is *online dispute resolution* (ODR), which means that the dispute resolution occurs fully or partially on the Internet. ODR origin is traceable to the early 1990s since the use of the Internet increased, and the environment became not so harmonious. It is might obvious to anyone today when consumer and copyright disputes are commonplace or when identity theft is on the rise. However, this was not clear in the early to mid-1990s, before spam, phishing, music downloading, buying and selling online, multi-player games, etc. Indeed, the hope expressed at that time was that this new online environment for commerce, education and entertainment would find ways to avoid the kinds of conflict. Years ago, many were sceptical of the need and potential of ODR, but today it is well understood. Most importantly, ODR, which focused initially on conflicts related to online activities, is now employed in offline disputes (KATSH, 2011).

In court practise, empirical research on ODR methods indicates that they are most efficient for disputes with a low level of complexity, such as Consumer Law. Legal relationships between conflicting parties with a specific character and type of claim, and based on clear provisions, might be subjected to ODR techniques. For this purpose, it is necessary to create a legal framework involving electronic tools, such as platforms and Internet portals that enable the automatic resolution of disputes without time-consuming hearings (MANIA, 2015).

To understand how technology can assist conflict resolution, Table 4 presents the generations of ODR with current and future perspectives of the subject and the role that technology and humans play in each one.

Table 4 – ODR generations

Generation	**Technology**	**Human**
First	Passive role. Its purpose is to facilitate information and communications between the parties. Examples include instant messaging, video and phone calls, videoconference and mailing lists.	Active role. Conciliators and mediators are carefully chosen according to their skills since they remain the central pieces in the planning and decision-making process.
Second	(1) Intermediary role (fourth part). Its purpose is to assist the conciliators or mediators in the generation, planning, strategy definition and decision-making processes based on knowledge of previous cases. (2) Active role. Its purpose is to represent the parties in the procedure. The technology is configured to act as the party would act or in a more efficient way.	(1) Intermediary role. (2) Passive role.

The 1st generation of ODR already occurs and is commonly practised in the Brazilian Judiciary. The 2nd generation, however, in its two perspectives, has not yet developed. To materialise this generation, we must explorer Artificial Intelligence, which we will explain in the section ahead.

2.1.2 Artificial Intelligence

The main object of research into AI is to enable a computer to perform the remarkable functions that are carried out by human intelligence (TSUJII; SHIRAI, 1985)[3]. On one side, from an engineering perspective, we use AI as an armamentarium of ideas to solve real-world problems. On the other, from a scientific perspective, we want to determine which of these ideas explain various sorts of intelligence (WINSTON, 1992).

An intelligent agent, thus understood as an integrated entity involving a computer system and its users, has (1) autonomy, since the agent operates without the direct intervention of the user or other agents; (2) social ability, since the agent interacts with other agents through some type of communication language; (3) reactivity, since the agent perceives the environment around it and responds opportunely to changes that occur; and (4) proactivity, since the agent not only acts in response to the environment but also takes initiative from a goal (ROVER, 2001).

If we have more than one agent operating in an environment, this refers to Distributed Artificial Intelligence (DAI). It is the study, construction, and application of Multiagent Systems (MAS), that is, systems in which several interacting, intelligent agents pursue some set of goals or perform some set of tasks. A pattern of interaction in MAS is goal and task-oriented coordination, both in cooperative and competitive situations. Then the long-term object of this field is to develop mechanisms and methods that enable agents to interact as well as humans (or even better) and understand the interaction among intelligent entities, whether they are computational, human or both (WEISS, G., 2001).

Back to understanding how an agent becomes intelligent, we can find in the literature three AI paradigms: *symbolic*, *evolutionary* and *connectionist*. The first points to models based on explicit representations that contain symbols organised in specific ways. Aggregate information is explicitly represented with structures constructed from constituent symbols and syntactic combinations of these symbols. Search and Knowledge Representation (KR) are fields that play a central role in symbolic AI (SUN, 1999). The second is based on the mechanisms of natural selection and population genetics, so AI research can use the Darwinian theory by constructing learning systems upon a simulated genetic basis (FENANZO JR, 1986; DASGUPTA; MICHALEWICZ, 2013). The third, which has been gaining attention in the last yeas, focuses on massively parallel models composed of a large number of simple and uniform processing elements interconnected. In these models each processor has a numerical activation value which it communicates to other processors along connections of varying strengths. Unlike the others, the foundation of connectionist models has always been learning. Then, Machine Learning (ML) is the typical field of this paradigm (SMOLENSKY, 1987; SUN, 1999).

[3] Although some AI techniques do not exactly intend to reproduce human intelligence, for example, genetic algorithms.

By the symbolic approach, an agent – as in humans – achieves intelligence thought processes of reasoning that operate on internal representations of knowledge. We can design agents that form representations through inference, using these representations to deduce what to do, that is, deriving new sentences from old. To this end, we develop Logic (set of rules) as a general class of representations to support knowledge-based agents. But we can also put content into the agent, creating representations of abstract concepts, such as events, time, physical objects and beliefs, that occur in many different domains. This is because complex domains require more general and flexible representations, for example, shopping on the Internet or driving a car in traffic. Representing these situations is called Ontology (RUSSELL; NORVIG, 2010).

The agent can also find a sequence of actions that will achieve its goal by search, i.e., a mathematical problem of finding a path from a initial stage to a final stage. However, the difficulty of search and the fact that humans can solve some search problems efficiently suggests that the agent should exploit knowledge about some cases to guide it to a solution. This extra knowledge beyond the search space is Heuristic (POOLE; MACKWORTH, 2010).

By the evolutionary approach, populations of agents are interbred under the influence of a task-based selection pressure. Starting from a population of random individuals, agents capable of performing the task well emerge. Each member of the population is evaluated according to some fitness function. Then, the process repeats to form the next generation. Genetic algorithms are based on this idea (HUSBANDS *et al.*, 1997).

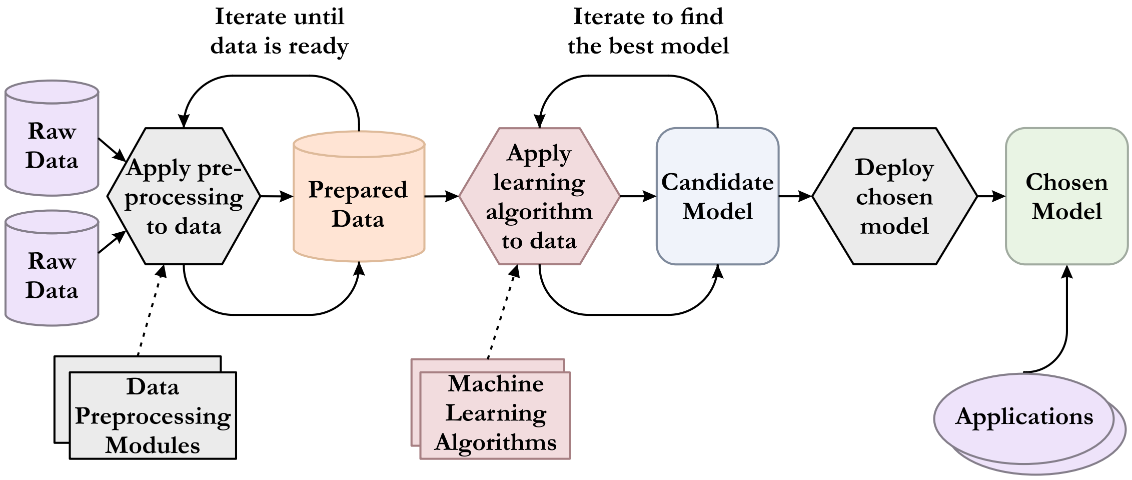
By the connectionist approach, an agent can either gain knowledge through learning, which allows it to operate in initially unknown environments and to become more competent than its initial knowledge alone might allow. The agent is learning if it improves its performance on future tasks after making observations about the world. For example, the programmers cannot anticipate all changes over time; a program designed to predict tomorrow's stock market prices must learn to adapt when conditions change from boom to bust. Therefore, ML is a good option of creating state-of-the-art systems (RUSSELL; NORVIG, 2010).

In practice, the idea is programming computers to optimise a performance criterion using example data or past experience. We have a model defined up to some parameters, and learning is the execution of a computer program to optimise the parameters of the model using the training data or past experience. The model may be predictive to make predictions in the future, descriptive to gain knowledge from data or both (ALPAYDIN, 2009).

Figure 1 briefly demonstrates the steps of the learning process, also called *pipeline*. The result of this step-by-step will be an ML model apt to be applied in a real problem.

Algorithms have been effective for certain types of learning tasks and significant commercial applications have begun to appear. In Data Mining (DM), learning algo-

Figure 1 – Learning process (pipeline)



Source: Chappell (2015).

rithms are being used routinely to discover valuable knowledge from large commercial databases containing loan applications, financial transactions, medical records, etc. It seems inevitable that ML is playing an increasingly central role in computer science and technology (MITCHELL, 1997).

In the context of text data, an example of a learning problem is how to use a finite sample of randomly selected documents, each labelled with a topic, to accurately predict the topic of unseen documents. The larger is the sample, the easier is the task. But the difficulty of the task also depends on the quality of the labels assigned to the documents in the sample, since the labels may not be all correct, and on the number of possible topics (MOHRI *et al.*, 2018).

In this perspective, a recent ML field is Deep Learning (DL), which focuses on creating large neural networks models capable of making accurate data-driven decisions. DL is particularly suited to contexts where the data is complex and where there are large datasets available. Nowadays, most online companies and high-end consumer technologies use DL. For example, Facebook uses DL to analyse text in online conversations. Google and Microsoft also use it for translation. Datasets that contain large numbers of features are called high-dimensional (such as text data). In these cases, it is extremely difficult to hand-engineer features for ML input. Instead, DL takes a different approach by attempting to learn automatically the features that are most useful for a given task from the raw data (KELLEHER, 2019).

Learning can be categorised into *supervised* or *unsupervised*. Supervised learning tries to infer a function or relationship based on labelled training data and uses this function to map new unlabelled data. The objective is to predict the value of the output variables based on a set of input variables. To do this, a model is developed from a training dataset where the values of input and output are previously known, i.e., it requires a sufficient number of libelled records to learn the model from the data. Supervised learning tasks include *classification* and *regression*. Unsupervised learning tries to uncover hidden patterns in unlabelled data. There are no output variables to predict. The objective is to find patterns in data based on the relationship between data points themselves. Unsupervised learning tasks embrace *clustering* and *association* (KOTU; DESHPANDE, 2019)[4].

A newer and open AI field is Explainable Artificial Intelligence (XAI), which is related to the idea of creating a suite of techniques and algorithms that produces more explainable models whilst maintaining high performance levels, i.e., a research area with the aim of making AI results understandable to humans and allow the users to appropriately understand, trust and manage the AI systems (GUIDOTTI *et al.*, 2018; ADADI; BERRADA, 2018). XAI can be also described as the intersection between three distinct areas: Interpretable AI, Human Computer Interaction (HCI) and Social Sciences. The first is responsible for creating interpretable intelligent systems; the second focus on designing the presentation of the predictions and explanations to the users. The last aims to detect whether the users understood the system and its explanations, and trust the results (MILLER, 2019).

2.2 RELATED WORKS AND STATE-OF-THE-ART

2.2.1 Artificial Intelligence in Alternative (Online) Dispute Resolution

The related works exposed below are empirical and applied research resulting from the SLR (Appendix A). In time order and diversified by country, we will explain their purpose in ADR/ODR, the AI techniques used and the legal domain of application.

Firstly, we present the Family-Winner, a Negotiation Decision Support System (NDSS) developed by Bellucci and Zeleznikow (2005) in the context of Australian Family Law, whose purpose is to assist the parties in a divorce and the mediator with advice on trade-offs and compensation strategies. The system is based on Game Theory and Heuristics. They also used the Adjusted Winner algorithm, which resolves a dispute by dividing issues and items among disputants through mathematical manipulation of numeric preferences. The platform asks the parties to indicate the items in dispute and their respective values representing their importance to them, which are used to form trade-off rules. Then it suggests a settlement by sequentially allocating items issues to disputants based on the value of ratings. A rating is a numeral that represents a disputant's want of an item. Ratings often change in response to a previous allocation. The trade-offs are graphically displayed through maps and enable the parties to visualise opportunities relevant to their side of the dispute (LODDER; ZELEZNIKOW, 2012).

[4] We will detail the ML tasks in section 3.4 since it is the AI field explored in this work.

Another NDSS is AssetDivider, the predecessor of Family-Winner, later developed at the request of disputants to arrive at legally fair solutions (also focuses upon the paramount interests of the children and not the interests of the parents). Both systems use the rating (the value indicating the importance of an item to the party) to assign the asset to one of the disputing parties. However, AssetDivider tests whether the asset's dollar value exceeds their allowable amount (given by the percentage split set by the mediator), improving the trade-off strategy (BELLUCCI, 2008; LODDER; ZELEZNIKOW, 2012).

A further extension of that work is the Integrated Online Dispute Resolution Environment (IMODRE), a NDSS that also provides negotiation advice in Australian Family Law disputes. The system upgraded to a multi-agent platform where agents are assigned to perform specific negotiation tasks. One agent uses a Bayesian belief network to recommend a percentage property split. This advice represents a disputants' best alternative to a negotiated agreement (BATNA). Another agent combines this percentage split with Heuristics and Game Theory to facilitate integrative bargaining between parties (ABRAHAMS *et al.*, 2012).

Carneiro *et al.* (2011) proposed an agent-based architecture for dispute resolution, whose objective is to enable a range of services targeted at assisting the disputant parties, independently of the conflict domain. The architecture is abstract and encompasses perceptions that are common to several legal areas. To this end, they used Ontologies for vocabulary, actions, features and theory to specialise a single agent, allowing it to expand to other domains. They developed it from the Portuguese legal system context, covering Family Law, Consumer Law and Labour Law.

Subsequently, that research evolved to University of Minho Court (UMCourt), a multi-agent platform that suggests solutions to new disputes based on the observation of past similar conflicts, improving negotiation and mediation. Among the AI techniques and fields, they apply Case-based Reasoning and Information Retrieval to select past court decisions that may be relevant to solve a given problem, learning with the success or failure of the application. They operated in the context of Portuguese Labour Law, focusing on the relationship between employers and employees, with attention to the scenario of firing an employee (CARNEIRO *et al.*, 2013a).

Considering that legal norms frequently change, rendering past cases potentially useless, Carneiro *et al.* (2013b) proposed a new approach in the UMCourt platform with Genetic algorithms to create possible solutions for a given dispute. This approach can generate a large number of different solutions that cover virtually the whole search space for a given problem. Genetic algorithms provide better results than Case-based ones since it is unrelated of the legal domain and does not depend on the number and quality of cases present in a database. Compared to the previous approach, the bio-inspired model reveals more efficiency, generating more solutions in less time, mainly due to the simplicity of the genetic operators.

At last, Gomes *et al.* (2014) conducted an experiment to identify and catalogue the behaviour of parties during a negotiation. They conceived a negotiation game in which sensors captured the behavioural and contextual information of the disputants. The data analysis, which focused on estimating stress levels, revealed conflict styles and behavioural patterns from the interactions. Then, having structured information detailing each party behaviour towards negotiation helps the mediator develops plans and suggestions for the associated participants. The study emphasises the importance of knowledge about social interactions as a basis for informed decision support in conflicts.

From that, they implemented two solutions working as modules of the UMCourt platform. The first classifies the conflict resolution style of the parties in real-time by analysing the proposals exchanges against boundary values defined in legal documents. The second assesses the level of stress of users based on their interaction patterns with the devices used as interfaces. Both are performed in a non-intrusive way and provide relevant knowledge about the context of interaction to the mediator and to the conflict resolution platform itself. For the classification task, they applied the k Nearest Neighbours (kNN) algorithm through Weka Workbench (CARNEIRO *et al.*, 2014, 2017).

Thompson (2015) presented the Justice Pathway Expert System (JPES), which purpose is to support users dealing with common issues and dispute types. Its functions include problem diagnosis, information, self-help tools and streaming or triage. JPES guides how to manage disputes independently and to engage with various justice system processes, avoiding access to justice problems for users. It also can help the users estimate chances of success in terms of a win/lose in a lawsuit or provide information to formulate a BATNA. Users enter the JPES through questions, which answers corresponding with the system's production rules. JPES is developed in the British Columbia Courts context, embracing various legal domains.

In the Italian mediation context, Fersini *et al.* (2014) proposed a framework based on Ontology and Logic to present strategies to the mediator for improving the chance of an agreement. The framework includes a smart data collection environment to state the essence of the litigation and intelligent retrieval of court decisions to aware the parties about their liability. Regarding the role of the mediator, it addressed an estimation of disputant flexibility to facilitate the optimal mediation strategies. This work was converged in the electronic Justice Relationship Management system (eJRM), an Italian project to enable citizens to personally evaluate the outcome of potential litigation and to be guided to a non-conflict settlement.

Afterwards, they proposed a new approach, the electronic Justice Relationship Management Information Retrieval system (eJRM-IRS), based on ML, Natural Language Processing (NLP) and Information Retrieval (IR). It can capture discriminant terms both in the disputant case description and court decisions, classify the disputant text into a legal field and provide relevant court decisions for consultation purposes. Its architecture consists of four steps: (i) indexing, that store court decisions in a database; (ii) core mining, that trains a classification model to predict the legal field to which a given case description belongs; (iii) query processing, that extracts relevant terms that will be used by core mining to predict the legal field of the text presented; (iv) ranking, that retrieves relevant court decisions (belonging to the foreseen legal field), reporting them to the disputants and mediators. The classifiers include Naive Bayes (NB), Decision Tree, Linear Support Vector Machine (SVM) and with Gaussian Kernel, all combined with Principal Component Analysis (PCA) to reduce the dimensionality of the text. The maximum accuracy obtained was 91.3%, by using the Linear SVM with PCA (EL JELALI *et al.*, 2015).

Another work within the eJRM system is that of Capuano *et al.* (2015), in which they described the creation of an underlying legal Ontology from existing sources and an Ontology integration algorithm used for its production. They also detailed a methodology for generating a training path meant to provide citizens with a better understanding of the legal issues arising from the given case, with corresponding links to relevant laws and jurisprudence retrieved from an external legal repository.

Later, they proposed a Smart Learning system based on Knowledge Discovery and Cognitive Computing. The features implemented within the system include the automatic conceptualisation and classification of textual legal cases, the generation of learning paths by relying upon legal Ontologies, and additional features for managing legal knowledge bases, including editing, versioning, integration and enrichment. The experiments achieved a positive evaluation in terms of effectiveness, efficiency and usability, rendering the system a successful cognitive learning platform for future law professionals and knowledgeable citizens. They concluded that advanced learning content and strategies based on digital storytelling foster the transition toward Personalised Learning Environments, going beyond the limitations of traditional classroom-based paradigms (CAPUANO; TOTI, 2019).

In China, one example of an AI judicial-assistant system based on DL and NLP is the Traffic-Accidents Dispute-Resolution System (TADRS), which has the following features: (i) first, it applies information-extraction algorithms to extract key factual elements from traffic-police records, which include weather data, traffic data, road data, traffic-light and signs data, vehicle data, driver data, accident data, and other relevant data; (ii) second, it associates the factorised factual data with key factors in relevant legal rules, including rules regarding the standard of care, traffic signs, driver's qualification, etc. to construct semantic models; (iii) then it runs deep-learning algorithms to scan tens of thousands of traffic-police decisions and court judgements to find patterns of how decision-makers attribute legal consequences to each factual–legal factor pair; (iv) fourth, it uses automated decision-making algorithms to suggest solutions; (v) finally, these suggestions are either presented to human decision-makers or popped-up to users of online-mediation platforms (ZHENG, 2020).

Also centred on ML and NLP, Tsurel *et al.* (2020) collected a large dataset of disputes from the eBay online marketplace and training classifications models to predict dispute outcomes, whose distribution was 59.6% for seller wins and 40.4% for buyer wins. They reached an accuracy of 86.0%, by using the Extreme Gradient Boosting (XGBoost). Other classifiers include Majority, kNN, Neural Networks (NN), Naive Bayes (NB), Decision Tree and Random Forest (RF). The authors also applied interpretability tools to explain the classifier's decision, such as SHAP and LIME algorithms, which include a component to verify the contribution of different features in the prediction, as well as a component for textual feature interpretation that highlights predictive tokens. In particular, they saw that losing a dispute harms the number of transactions made after the dispute ended and that dispute outcome is reflected in politeness strategies used during correspondence.

Simkova and Smutny (2021) presented a design of the E-NeGotiAtion method for assisted negotiation in business-to-business (B2B) relationships, which uses a genetic algorithm for selecting the most appropriate solution(s). They also evaluated the proposal with students from Slovakia and students from the Czech Republic. The results confirm that the use of evolutionary computation brings advantages compared to case-based and rule-based reasoning, which are by nature more domain-related (e.g., with legislation, rules identified by an expert, or available cases from the past). Their proposed method applies to a wide area of B2B negotiation and offers more flexibility because the specific contextual conditions can be included as the negotiation process is changed without intervention into the genetic algorithm, which focuses on the general problem of finding the optimal solution in the state space of possible solutions.

More recently, Saygili *et al.* (2022) proposed a framework by employing a Decentralised Construction Enabling Transparent Resolution (DCENTR) platform to develop a blockchain network-based ODR system for the construction industry litigation (cases involve an owner and a contractor) and tested it with real ones. The aim was to provide (i) ease and reliability of contract and payment execution and (ii) fast, low-cost and transparent dispute resolution for multiple parties involved. The authors concluded that dispute likelihood can be minimised through reliable contract and payment execution, and if occurred, construction disputes can be resolved with higher transparency and dramatic savings in effort, time, and cost.

2.2.2 Machine Learning in Supreme Courts and other legal applications

Knowing that ML is still under-explored in the ADR domain but, on the other hand, it is widely implemented together with NLP for decision prediction using legal text data, we conducted a complementary literature review (simply narrative) to search works involving ML application in Supreme Courts and for other legal utilities (such as contracts analyses) around the world.

The first large experiment in the judicial environment occurred at the United States Supreme Court, where Katz *et al.* (2014, 2017) constructed a model to predict the court behaviour (cases outcome). For this purpose, they used the votes of several judges who have integrated it over the sample years, in which a decision is confirmed or changed in a higher instance. They also used attributes such as the year of the case, the legal matter discussed and the location of the lower court. The first input included decisions of six decades (1953-2013) and the second of two centuries (1816-2015). The approach with Decision Trees and Random Forest (ensemble method) classifiers achieved 70.2% of accuracy on the case outcome and 71.9% on the judge vote predictions.

At the European Court of Human Rights (a 47-Member body of the Council of Europe), Aletras *et al.* (2016) constructed a model to predict if a Member State act might be a violation or a non-violation of human rights (binary classification), based on the cases and premises of civil and political rights included in the European Convention on Human Rights. The cases comprise specific text parts referring to the fact, applicable law and the arguments presented by the parties involved. The results indicated that the fact section is the most important predictive factor in the text, that is, judicial decision-making is significantly affected by the situation narrated by the parties. The approach with Linear SVM classifier had 79% of accuracy.

At the Supreme Court of France, Şulea *et al.* (2017) explored different text classification techniques to predict the (2) law area of a case, (2) the court ruling based on the respective case description, and (3) when a case description and a ruling were issued. They used cases and decisions from the 1880s to 2010 as input. The approach with Linear SVM classifier reported 96% of F1-Score in predicting a case ruling, 90% of F1-Score in predicting the law area of a case, and 75.9% of F1-Score in estimating the decade to which the case belongs.

At the Supreme Court of the Philippines, to reduce court congestion and problems with pending cases, Virtucio *et al.* (2018) conducted experiments to predict the case outcome in the criminal context. They used public processes filed from 1987 to 2017 as input. The work includes extensive data preparation, with a categorisation of legal area, case type, laws and crimes. As result, they obtained 59% of accuracy with Random Forest classifiers.

To speed up the appeals examination, Nilton Correia da Silva *et al.* (2018) conducted a project to automatically linking cases to a “general repercussion” (RG) issue in the Brazilian Supreme Court. The RG is a local procedural instrument that acts as a “filter”, allowing the court to select the appeals that it will judge according to its previous criteria of legal, political, social or economic relevance. Then the model should classify an entering appeal in one of the RG issues. The approach based on DL had an accuracy of 90.35% in a preliminary evaluation with Convolution Neural Networks (CNN). This was the first ML project in the Brazilian Judiciary, called VICTOR.

Lastly, at the Indian Supreme Court, Sharma *et al.* (2022) enabled a prediction model and its operational prototype, eLegPredict, which successfully predicts court decisions. The attribute of importance is 'appeal', which is mainly classified into three categories: (i) allow, (ii) dismiss, and (iii) dispose. They achieved 76% of accuracy by applying the X Gradient Boost classifier. The eLegPredict is equipped with a mechanism to aid end-users, where as soon as a document with a new case description is dropped into a designated directory, the system quickly reads through its content and generates prediction.

ML is also explored in other legal services, just as contractual analysis. The classification of clauses or sentences is helpful for informing a layperson of their rights and obligations. Glaser *et al.* (2018) performed experiments to verify the portability of ML models in different document types. They trained different ML classifiers on a dataset about Tenancy Law (sentences extracted from the German Civil Code - BGB) and applied the best models on a rental agreements dataset. The model's performance varies on the contract set. Some models perform significantly worse, while some reveal portability, such as the Extra Trees classifier (ETC), with 82.7% of F1-Score in train BGB and 82.5% of F1-Score in test rental agreements. Furthermore, they trained and evaluated the same models in a dataset consisting solely of contracts to observe a reference performance. The results demonstrated that ML models are portable depending on the document type used for training.

In Italy, Lippi *et al.* (2019) proposed CLAUDETTE, a web server based on ML and DL models, which can detect potentially unfair clauses in Terms of Service, often not read by consumers due to the large volume of texts. The dataset includes Terms of Service from Microsoft, Amazon, Airbnb, Spotify, Facebook, Dropbox, etc. The types of unfair clauses cover Consumer Law issues, for instance, arbitration, unilateral change, content removal, jurisdiction, choice of law, limitation of liability and unilateral termination. They applied different ML and DL classifiers (SVM, LSTM, and CNN), achieving an average of 80% precision. The results were satisfactory, thus becoming an empowering tool for consumers. Recently, they performed a new approach by training a Memory-Augmented Neural Network (MANN) to identify unfair clauses using as facts the legal *rationales* (justifications provided by legal experts motivating their conclusion to consider a given clause as unfair) behind the unfairness labels, then a possible explanation of an unfairness prediction could be constructed based on the list of memories (RUGGERI *et al.*, 2022).

Tuggener *et al.* (2020) constructed LEDGAR, a freely available multilabel corpus of legal provisions in contracts. The dataset includes over 12,000 labels annotated in almost 100,000 provisions in over 60,000 contracts, then it is interesting for research in the field of NLP, text classification, as well as for legal studies. In their classification experiments, they used Logistic Regression, Neural Networks and BERT approaches, obtaining satisfactory metrics depending on the subcorpora.

Finally, one similar and recent work is that of Hendrycks *et al.* (2021). They created the Contract Understanding Atticus Dataset (CUAD), a new dataset for legal contract review. The purpose is to highlight salient portions of a contract that are important for a human to review. Contracts often contain main clauses that warrant review or analysis by lawyers. CUAD consists of over 500 contracts, each carefully labelled by legal experts to identify 41 different types of relevant clauses, for a total of more than 13,000 annotations. With CUAD, models can learn to extract and identify significant clauses from contracts automatically. They applied BERT approaches and achieving reasonable metrics depending on the labels.

2.2.3 Considerations on the section

From the reading and evaluation of the related papers, we conclude that:

1. Proposing AI-based systems to encourage agreements between disputing parties is not something new, including the idea of presenting them a previous case to guide their decisions.

2. Predicting court decisions to reduce court congestion and improve the consistency of judgements, in the analysis of long legal texts, is not new either.

However, to the best of our knowledge, predicting judicial decisions aiming to construct a ODR solution, i.e., to achieve more agreements by providing real possibilities for parties in conflict, allowing them to feel safe and confident to agree, has not yet been attempted by the Law and AI scientific community (at least not in Brazil).

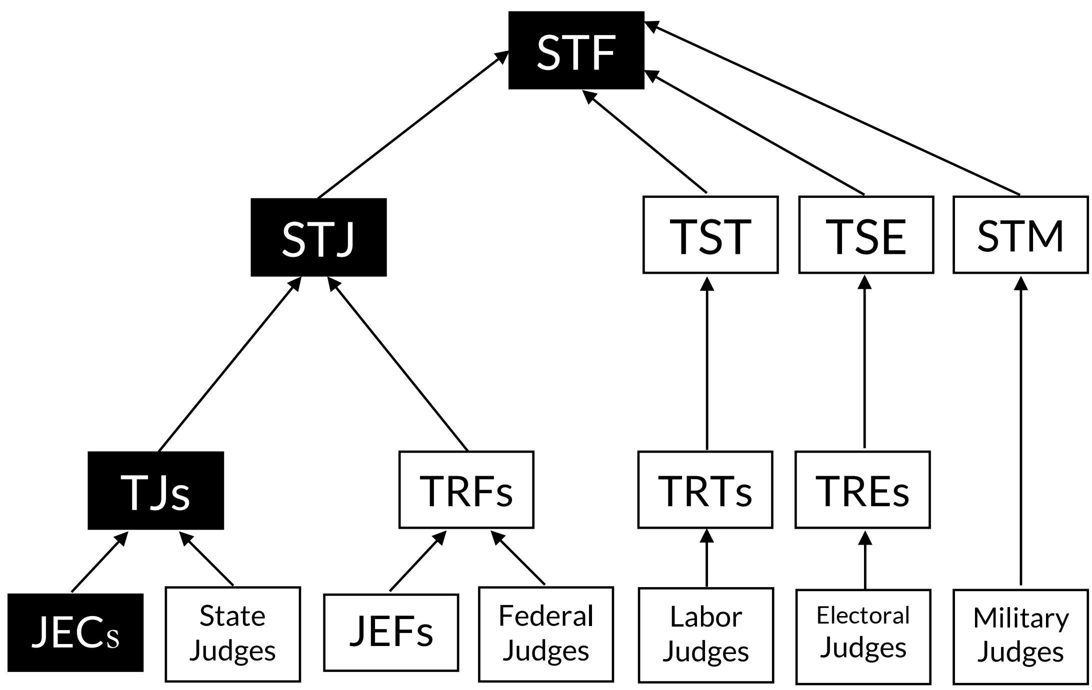
To present real possibilities for the parties in conflict is necessary to work with ML tasks, which requires a large dataset. It is fully possible in the Brazilian Judiciary, considering its high litigiousness rates. We will explore this experience in the case study that forms this thesis.

3 CASE STUDY

3.1 THE SPECIAL CIVIL COURT

According to Brazilian Federal Constitution, the Brazilian Judiciary is composed of: (1) Federal Supreme Court (STF); (2) Superior Court of Justice (STJ); (3) Federal Regional Courts (TRFs) and Federal judges; (4) Superior Labour Court (TST), Labour Regional Courts (TRTs) and Labor judges; (5) Superior Electoral Court (TSE), Electoral Regional Courts (TREs) and Electoral judges; (6) Superior Military Court (STM) and Military judges; (7) State Courts (TJs) and State judges. Also, Federal and State Courts can, within their jurisdiction, create the Special Courts (JECs and JEFs), which are responsible for judging local less complex cases (BRASIL, 1988). Figure 2 illustrate this organisation.

Figure 2 – Brazilian Judiciary organisation chart



Under this structure, the Special Civil Courts (JECs) are agencies of the Brazilian State Courts (7), created in 1984 with the purpose is to facilitate the citizen's access to justice. To help citizens solve their problems, they provide the remission of lawsuit costs, procedural simplification and incentive to conciliation between the disputant parties. JECs tend to get closer to the legal demands of the ordinary citizen, who is involved in daily and minor conflicts (WATANABE, 1985).

Parties do not need a lawyer to enter the JECs. When one party enter the Special Court, a summons is issued for the other with the date of the conciliation hearing. Both parties are obliged to attend it. Conciliation can be conducted by: (1) the chief judge (who is responsible for the court and is judges the lawsuits); or (2) the voluntary judge (a person who has a law degree but is not invested in the position); or (3) the conciliator (a person who has a law degree and is qualified to conciliate). Once the session is opened, whoever conducts it shall explain to the parties the advantages of conciliation, as well the risks and consequences in going ahead with the lawsuit (BRASIL, 1995).

However, as introduced in this work, conciliation rates in JECs are still low. Thus, before proposing solutions to modernise the judicial environment, we must investigate closely how conciliation hearings work and diagnose what problems they currently face. It means that we conduct a case study, a research procedure whose object is a unit that will be analysed in depth (TRIVIÑOS, 1987). The unit we will study is the Special Civil Court located at the Federal University of Santa Catarina (JEC/UFSC).

3.2 NON-PARTICIPANT OBSERVATION OF HEARINGS

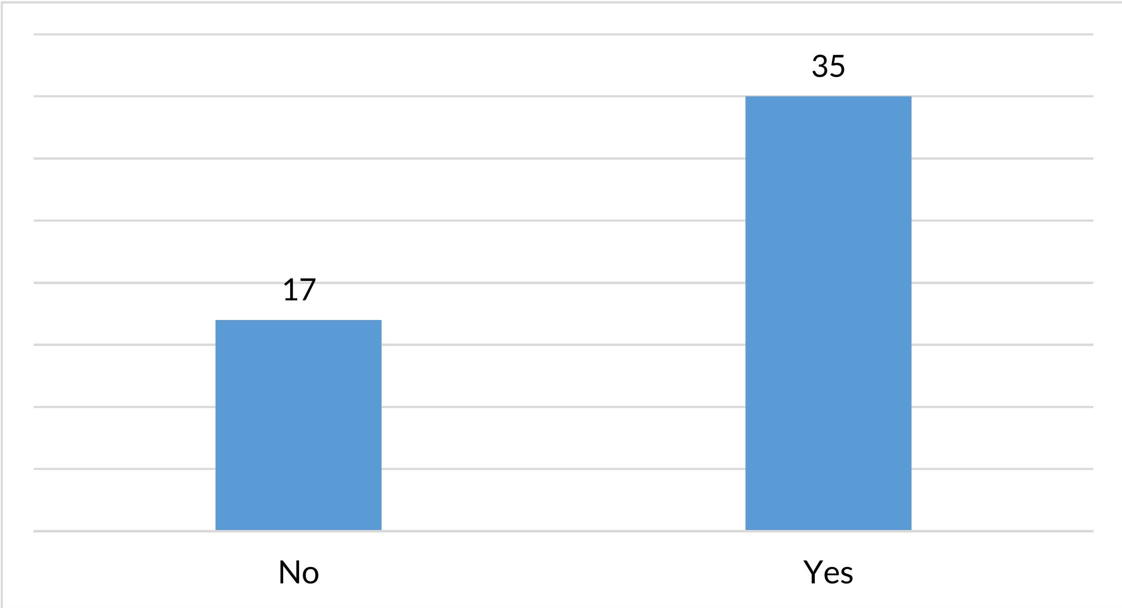
To better identify the causes of the low conciliation rates, we first employed the technique of free observation (non-participatory). We observed 52 conciliation hearings held in JEC/UFSC in the period between 09/04/2019 to 18/07/2019, based on the following questions. The official documents relating to our non-participant observation of the conciliation hearings are listed in Annex A.

1. Did all parties attend the conciliation hearing?
2. Did they have a lawyer?
3. What is the legal area and sub-area?
4. Did the conciliator suggest any form of agreement?
5. Did the conciliator refer to previous similar cases to guide the parties?
6. Were the parties interesting to settle?
7. What was the outcome of the conciliation hearing?

All hearings were conducted by conciliators. We emphasise that we do not intend to conduct quantitative analysis since the sample is not representative (there are conciliation hearings every day). The objective of the non-participant observations is to experience the dynamics of the conciliation hearing, knowing the place and the people involved. In other words, this analysis is qualitative. We expose and organise the responses to the questions in the charts below, followed by our analysis.

Figure 3 shows that although the attendance of the parties at the hearing is mandatory, there is a significant absence. The absence may be due to disinterest or failure to summon. In the first situation (disinterest), if the absence is of the plaintiff, the case is

Figure 3 – (Non-participant observation) Q1. Did all parties attend the conciliation hearing?



dismissed without prejudice. If the defendant is absent, the case is judged by default. In the second case (failure to summon), the hearing is postponed. In any case, if one of the parties is absent, the hearing is prejudiced, which means that it does not continue.

Figure 4 – (Non-participant observation) Q2. Did the parties have a lawyer?

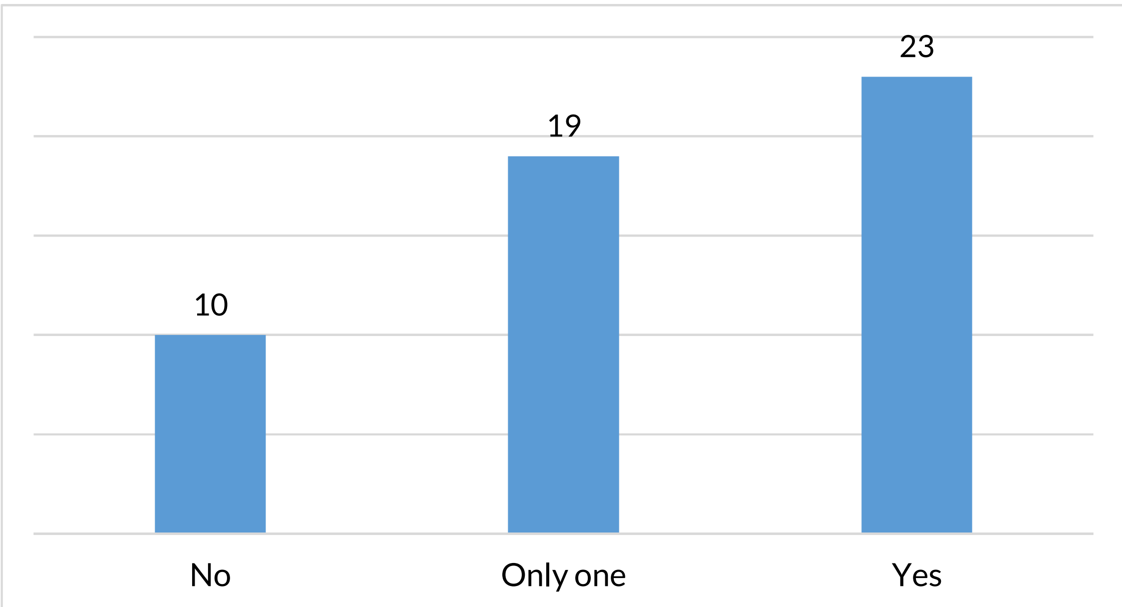


Figure 4 indicates that it is common to have a lawyer representing at least one of the parties in the hearings, although this is not mandatory in the JECs to facilitate access. The lawyer is an important element in the negotiations due to his/her legal knowledge and experience. However, we have also observed that due to his/her interest (payment of his legal fees), he/she may direct the parties not to accept smaller amounts of compensation.

Figure 5 – (Non-participant observation) Q3a. What is the legal area?

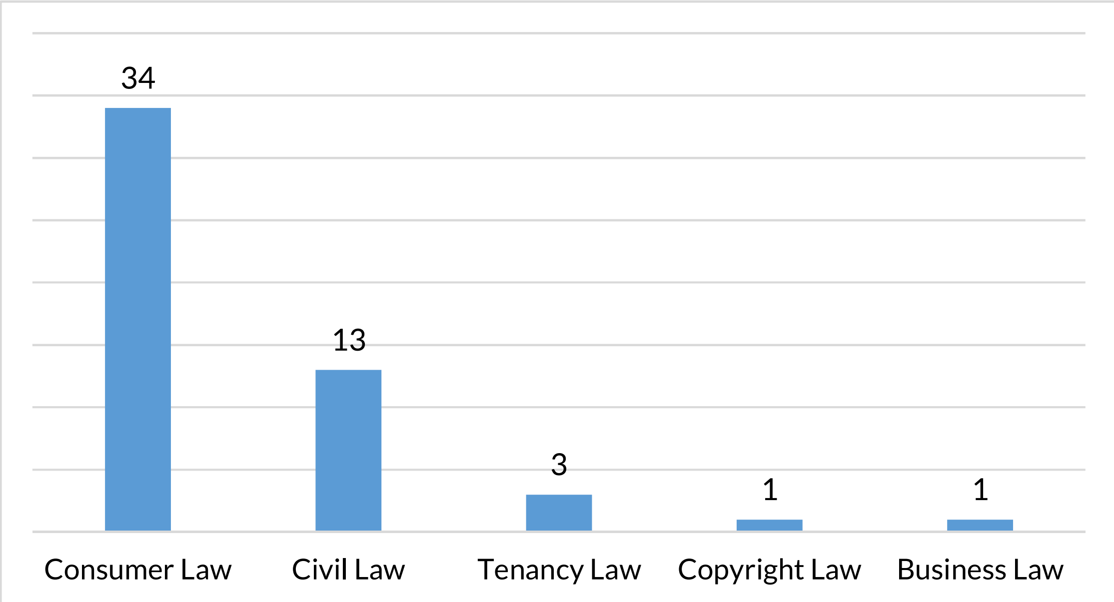
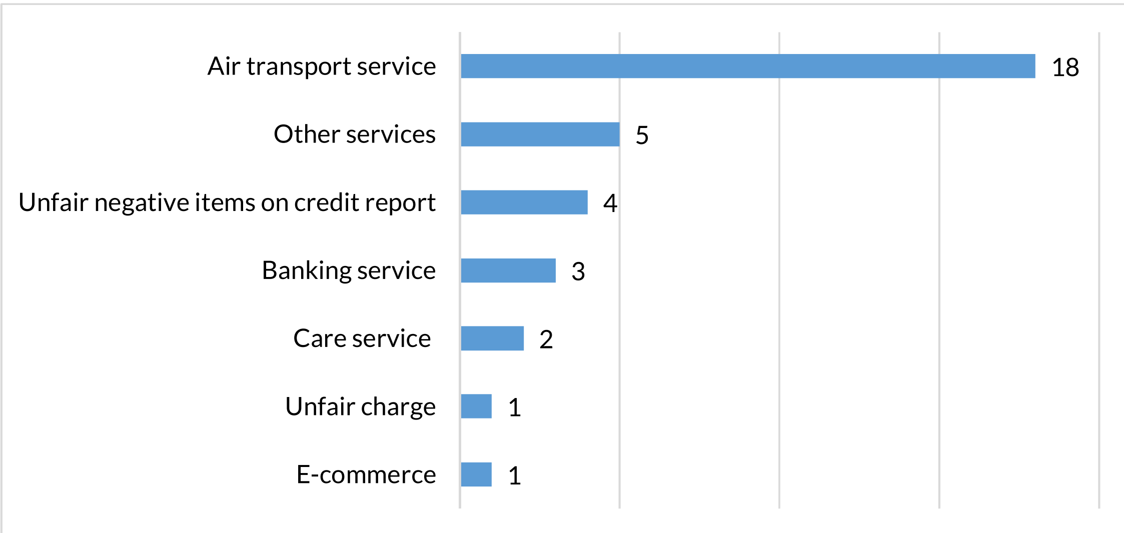


Figure 6 – (Non-participant observation) Q3b. What is the Consumer Law area?



Figures 5 and 6 demonstrate that the legal area with the highest incidence of conflicts discussed in the hearings is Consumer Law. Within this area, the predominant subject is failures in the air transport service. Considering the high incidence of lawsuits filed on this subject, some of the hearings were held by joint effort.

Figure 7 – (Non-participant observation) Q4. Did the conciliator suggest any form of agreement?

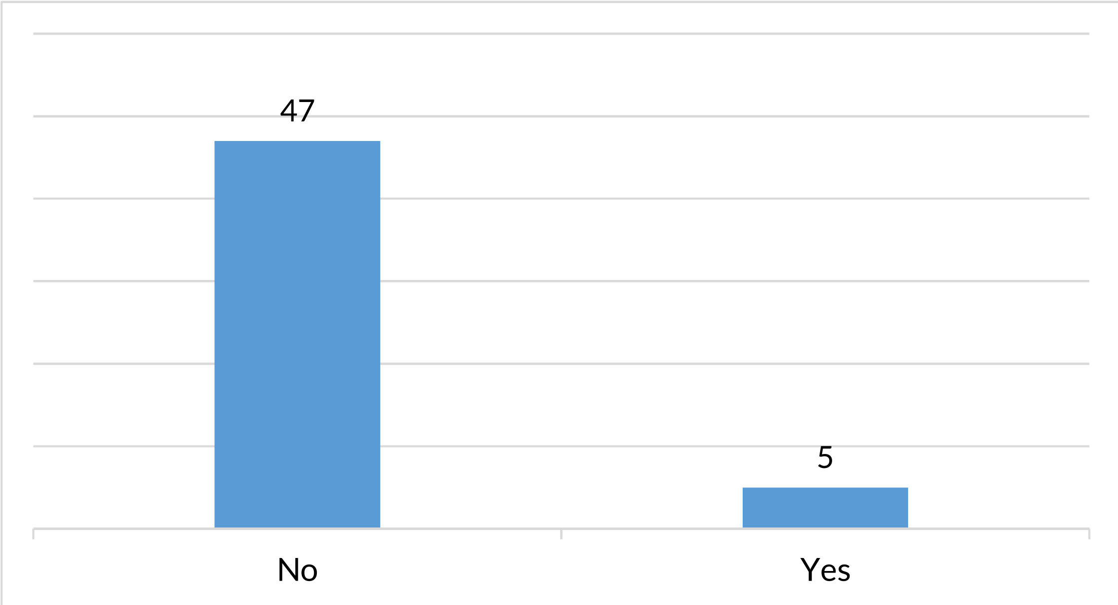
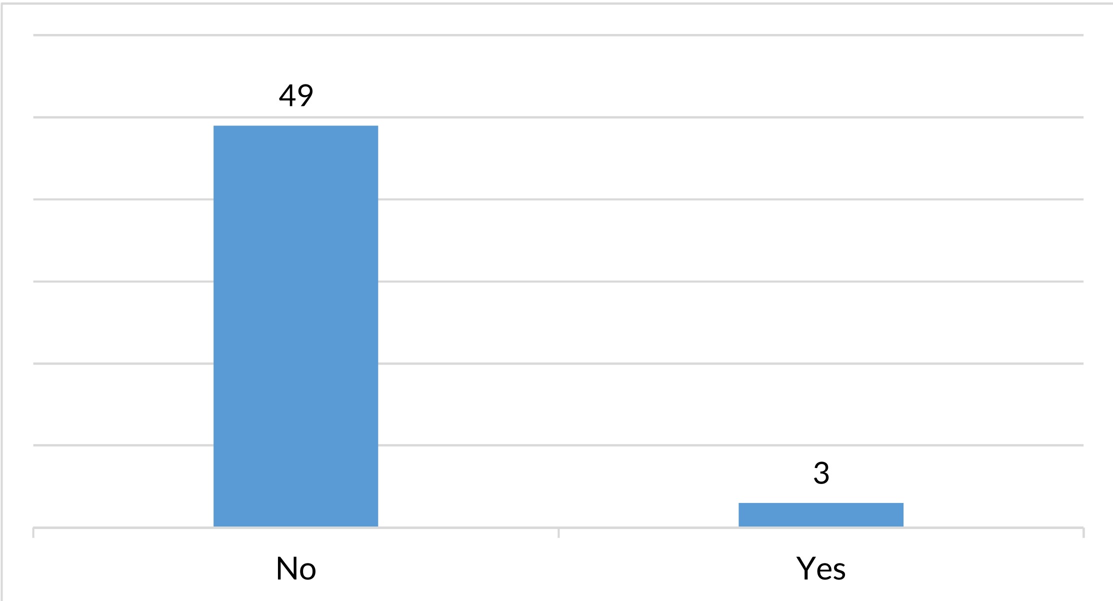
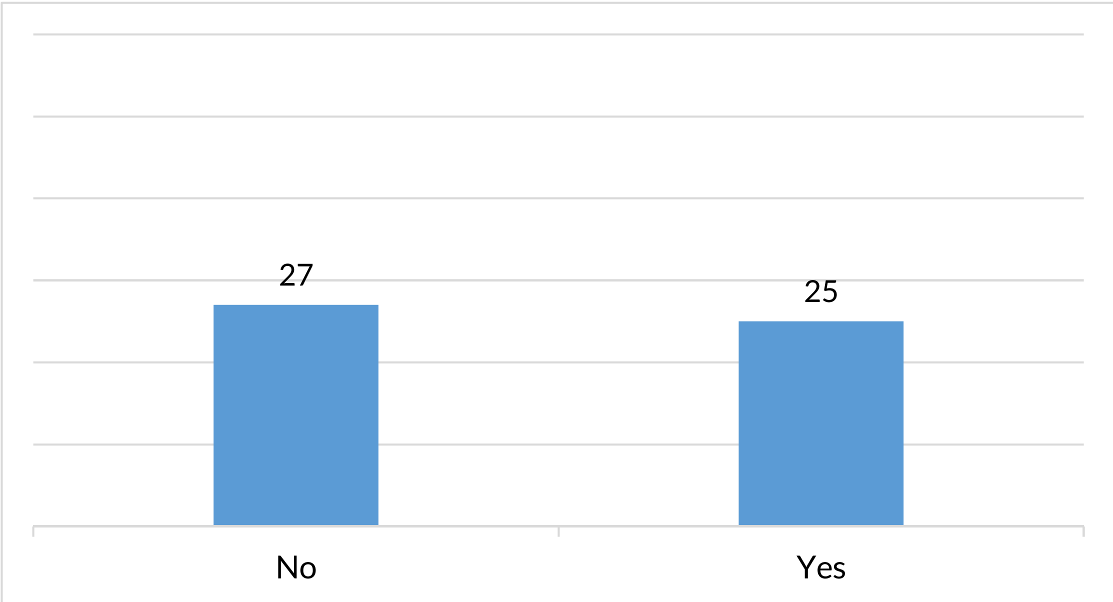


Figure 8 – (Non-participant observation) Q5. Did the conciliator refer to previous similar cases to guide the parties?



Figures 7 and 8 show very few situations the conciliator suggests a form of agreement or refers to a similar previous case to assist the parties. This diagnosis allows us to affirm that one of the reasons for the low incidence of agreement is that the parties do not know about the judgements, encouraging our thesis proposal. If the conciliator does not bring some judgement possibility at the initial hearing, the party will prefer to wait for the real judgement, even more so in the face of the no-cost process. Furthermore, without information it is more difficult to start a negotiation.

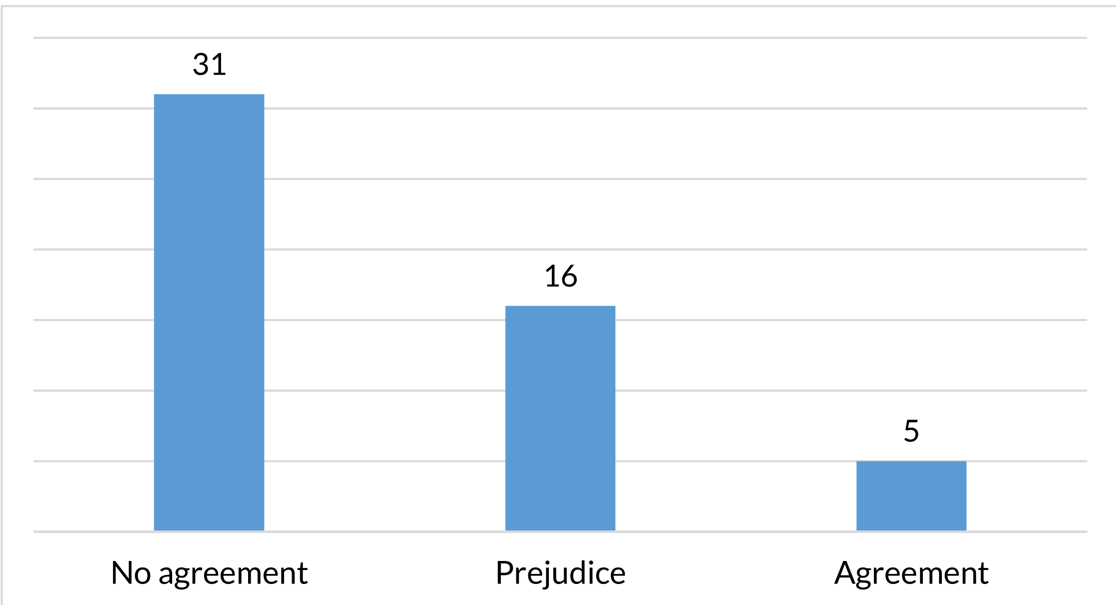
Figure 9 – (Non-participant observation) Q6. Were the parties interesting to settle?



Although the conciliator is trained to deal with the parties, we observed that due to the high number of cases submitted to the hearings, he or she has no time to study each lawsuit in depth. This situation creates insecurity for him or she in suggesting settlements or mentioning judgement possibilities to the parties.

Figure 9 summarises the intention of the parties during the hearing according to

Figure 10 – (Non-participant observation) Q7. What was the outcome of the conciliation hearing?



what we observed. Even though the parties must attend the hearing, there are situations such as being absent or not presenting proposals to start a negotiation, which initially eliminates the possibility of an agreement. It shows that the litigation culture remains in Brazilian Justice since some of the parties (in this sample, a half) prefer to wait for the judgement. Finally, Figure 10 confirms that as happens in general in Brazilian Justice, also in JEC/UFSC there are few agreements reached during the conciliation hearing.

Among the variables found that interfere in the non-realisation of the agreement, the one we will focus on in this thesis proposal is the lack of information (judgement possibilities) provided by the conciliator during conciliation hearings.

3.3 DATASET CONTEXT AND CONSTRUCTION

As verified during the observations, the subject of air transport service (Consumer Law) is one of the most recurrent in the lawsuits filed at the JEC/UFSC.

Consumer Law is a legal area that came to prominence in Brazil in the early 1990s, which recognises that the individual interests of daily life and political struggles of several types converge in more active participation of citizens for the democratisation of society in times of globalisation (LOPES; LANIADO, 2010). In this direction, the Brazilian Code of Consumer Protection provides basic consumer's rights, such as effective compensation for material and immaterial damages (BRASIL, 1990).

There is no difficulty in understanding the material damage, which is a decrease in the injured party's assets or a loss of earnings (DINIZ, 2020). In cases of air transport service failure, the material damage may refer, for example, to the purchase of a new ticket, hotel and food expenses due to flight delays and cancellations, among others. But conceptualising immaterial damage is complex, even from a legal perspective.

In Brazil, where we commonly use the expression "moral damage", the concept refers either to the intimate mental sphere injured by the pain a person suffers or to the injury to his social reputation. It is related to the growing legal protection of personal rights, whose injury creates a duty to compensate (REALE, 1992). In other words, immaterial damage is an injury to personality rights, such as honour, dignity, intimacy, image, name (GONÇALVES, 2020). Regarding the failures in air transport service, for example, flight delay, flight cancellation, overbooking, baggage loss, etc., the courts have been decided that these events can generate immaterial damage and consumer compensation (BENJAMIM, 2015).

Compensation for immaterial damage is usually monetary. It is not possible to evaluate the painful sensation experienced by the injured person. As a mean of mitigating the consequences, money can play a satisfactory role (DINIZ, 2020). Immaterial damage compensation is one of the most controversial matters in the judicial practice of several law systems due to a lack of criteria for its assessment. There are some circumstances considered by the judge when fixing the value, such as the person's age, health status, person's gender, place and time of injury. Anyway, these variables are weighted by the judge in a free assessment, according to his/her interpretation of each case (SADIKU, 2020).

Since it offers an unbureaucratic and no-cost way out to solve this kind of problem, it is natural that the injured consumer goes to the Special Civil Court to claim for material and immaterial damage. The last one becomes an interesting machine learning problem and application, because the amount of compensation set by the judge is unknown, while its prediction may help in reaching agreements.

To construct the dataset, we collected all the documents manually and personally into the JEC/UFSC, through the e-SAJ[5] and e-Proc[6] (Brazilian electronic judicial process systems). We had the support of the local chief judge in this step to avoid repeated judgements or judgements about a subject not related to failures in air transport service. Although the Brazilian Judiciary indexes its processes according to the subject, the indexation may be incorrect due to human error. It whats occurs, for example, when a lawsuit is registered by different operators (lawyers, parties or Judiciary employees). Efforts to unify data management in the Brazilian Judiciary are very recent[7]. Thus, it was possible to get more outliers than expected if we did not make a personalised collection.

The structure of the collected document (a 1st-degree judgement or court decision), consists of:

1. *Report*: the summary of what happened according to the allegations and evidence of the parties.

2. *Reasoning*: the reasons that formed the judge's conviction about how the facts occurred.

3. *Result*: the verdict, which can be:

    a) *Well founded*: The consumer wins the lawsuit.

    b) *Not founded*: The consumer loses the lawsuit.

    c) *Partly founded*: The consumer partially wins the lawsuit (e.g., when he/she pleads for greater compensation than the assigned value by the judge).

    d) *Dismissed without prejudice*: The consumer makes a procedural error (e.g., when he/she indicates as a defendant the wrong airline company). So the consumer can file a new lawsuit.

We note that this is a legal structure, based on art. 489 of the Brazilian Code of Civil Procedure (BRASIL, 2015a), not a structure in data terms, since this is not tagged/annotated

[5] "Sistema de Automação da Justiça": https://esaj.tjsc.jus.br.
[6] "Sistema de Transmissão Eletrônica de Atos Processuais": https://eproc1g.tjsc.jus.br.
[7] In August 2020, the Brazilian Judiciary instituted "Datajud", a unified national database, and the courts are required to send data of all processes following pre-established standards (CNJ, 2020b).

in the text. After the data collection, we reached a final dataset consisting of 1163 judgements specifically on failures in air transport service issued between February 2011 to September 2020, and imported in XML and TXT format with UTF-8 encoding.

3.4 MACHINE LEARNING STEPS AND TECHNIQUES

In this section we explain all the ML steps and techniques used during the dataset experiments, including other related areas, such as Text Mining (TM) and Natural Language Processing (NLP). We will limit it to conceptualising and explaining the functioning of the techniques without delving into mathematical terms.

3.4.1 Preprocessing

We refer to data preprocessing or data preparation as the set of techniques that initialise the data properly to serve as input for a certain DM and ML algorithms. In recent years, this area has become of great importance because these algorithms require meaningful and manageable data to operate correctly and to provide useful knowledge, predictions or descriptions (GARCÍA *et al.*, 2015).

The legal judgements are submitted to the following preprocessing techniques:

1. *Normalisation*: converting all characters inside the document to lowercase (JURAFSKY; MARTIN, 2019).

2. *Tokenisation*: grouping characters in a string into meaningful pieces. This technique identifies important components or tokens using a set of delimiters, such as space and punctuation (LEE, M. L. *et al.*, 1999). We used regular expressions, an algebraic notation for characterising a set of strings, particularly useful for searching in texts when we have a pattern to search for and a corpus of texts to search through. A regular expression search function will search through the corpus, returning all texts that match the pattern (JURAFSKY; MARTIN, 2019).

3. *Stemming*: reducing a word variant to its stem by removing any attached suffixes and prefixes (affixes). The stem does not need to be an existing word in the dictionary, but all its variants should map to this form after the stemming has been completed (JIVANI *et al.*, 2011). We used the Porter stemming, which works on a bunch of rules where the basic idea is to remove and/or replace the suffix of words. For example, one rule would be: to stem all terms ending in s by removing the s as in algorithms to algorithm. While the method is extremely efficient, it can make mistakes that could prove costly (KOTU; DESHPANDE, 2019). It happened in our dataset, in which the word “morais” (a Portuguese word meaning moral in plural) was reduced to “morai” (Portuguese verb meaning living).

4. *Filtering*: removing such terms that do not convey specific meaning (stopwords), such as articles, conjunctions, pronouns and prepositions (KOTU; DESHPANDE, 2019).

5. *N-grams*: grouping of words that generally appear together. It is usually the final preprocessing step (KOTU; DESHPANDE, 2019). In our dataset, we defined the *n* in 2 (bigrams).

3.4.2 Representation

After the preprocessing step, it is also necessary to transform the corpus to a numerical representation, which will serve as input to some ML techniques. A simple way to model documents is to transform them into sparse numeric vectors and deal with them with linear algebraic operations. This representation is called *Bag of Words (BOW))* or *Vector Space Model (VSM)*. In the BOW model, a word is represented as a separate variable having a numeric weight of varying importance, which may be its frequency in the text. Extracted N-grams are also considered units in the BOW (HU; LIU, H., 2012).

Thus, the most common parameter of this count in the BOW model and that we adopted in this work is the Term Frequency (TF), which is simply the ratio of the number of times a keyword appears in a given document, (n_k, where k is the keyword), to the total number of terms in the document (n), as shown in Equation 1:

$$TF = \frac{n_k}{n} \tag{1}$$

However, in our dataset context, a common and generic word such as “process” will have a fairly high TF score and a specific word such as “baggage” will have a much lower TF score. For this problem we have another count parameter, the Inverse Document Frequency (IDF). Considering the mentioned example, IDF is defined as the Equation 2, where N is the number of documents under consideration and N_k is the number of documents that contain the keyword (k).

$$IDF = \log_2 \frac{N}{N_k} \tag{2}$$

Again, a word such as “process” would arguably appear in every document and, thus, the ratio ($\frac{N}{N_k}$) would be close to 1, and the IDF score would be close to zero for. However, a word like “baggage” would possibly appear in a relatively fewer number of documents and so the ratio ($\frac{N}{N_k}$) would be much greater than 1. Thus, the IDF score would be high for this less common keyword.

Finally, by weighting this two counts we can calculate the TF-IDF, which is expressed as the simple product as shown in Equation 3.

$$TF - IDF = \frac{n_k}{n} . \log_2 \frac{N}{N_k} \tag{3}$$

Other and more recent text representation techniques have been proposed to improve this example, such as *Word Embeddings*, which is capable of capturing syntactic and semantic linguistic patterns in the text (BOJANOWSKI *et al.*, 2017). However, in this work, we focus on the BOW model to represent our corpus, due to its simplicity and easy interaction with the other techniques used.

3.4.3 Clustering

As introduced in the 2.1.2, one of the ML unsupervised learning tasks is clustering, which helps to identify patterns from the dataset. Clustering algorithms can be divided into two categories: hard and soft. In the hard clustering, each document will be assigned to only one cluster, while in the soft the documents can be assigned to one or more clusters. Soft clustering may be suitable in applications where the overlap between clusters is reasonable, and outliers and uncertain cluster memberships may happen (PETERS *et al.*, 2013).

We applied four types of clustering algorithms: Hierarchical and Lingo (soft clustering), K-means and Affinity (hard clustering), which are explained as follows:

1. *Hierarchical Clustering*: This soft clustering algorithm builds tree structures merging documents and clusters according to their similarities (AGGARWAL, 2018). There are two kinds of hierarchical clustering: agglomerative and divisive. Agglomerative hierarchical clustering is a bottom-up approach that starts with many small clusters and iteratively merges selected clusters until reaches a single root cluster. On the other hand, divisive hierarchical clustering is a top-down approach that starts with a single root cluster and iteratively partitions existing clusters into sub clusters (CICHOSZ, 2015).

2. *Lingo*: This more recent soft clustering algorithm initially identifies the label of each group and then assigns documents to them. Specifically, it extracts frequent phrases from the input documents, hoping they are the most informative source of human-readable topic descriptions. Next, by performing the reduction of the original term-document matrix using a factorisation method such as Singular Value Decomposition (SVD), it tries to discover any existing latent structure of diverse topics in the search result. Finally, Lingo matches clusters descriptions with the extracted topics and assigns relevant documents to them. To select the best label for each cluster, it calculates a score between an abstract concept from the factorisation method and the phrase vectors. Then, to calculate the score for the cluster, it simply multiply the label score to the number of documents assigned to the cluster. Thus, Lingo gives priority to the readability of the label and the size of the cluster (OSIŃSKI *et al.*, 2004).

3. *K-means*: This simple algorithm of hard clustering defines a prototype in terms of a centroid, which is usually the mean of a group of points and is typically applied to data objects in a continuous n-dimensional space. We first choose K initial centroids, where K is a user-specified parameter, namely, the number of clusters desired. Each point is then assigned to the closest centroid, and each collection of points assigned to a centroid is a cluster. The centroid of each cluster is then updated based on the points assigned to the cluster. We repeat the assignment and update steps until no point changes clusters, or equivalently until the centroids remain the same (TAN *et al.*, 2014).

4. *Affinity Propagation*: This more recent hard clustering algorithm has a distinct approach to form the clusters when compared to K-means. Affinity Propagation considers all the data points as nodes in a network with the potential to form a cluster. By the exchange of real value messages along the edges of the network, it searches for a good set of exemplars, or clusters centres. The value of the messages reflects the affinity of a data point to choose another as its exemplar and it is updated along the iterations in the algorithm. We do not set any number of clusters to the algorithm, Affinity Propagation automatically finds the ideal number of clusters (FREY; DUECK, 2007).

3.4.3.1 Similarity

To group documents in the clustering application, it is necessary to define the similarity or dissimilarity between them numerically. In this work we relied in these two measures:

1. *Cosine Similarity*: This is a similarity measure that has gained high popularity, particularly in text clustering applications, which calculates the angular distance by comparing two numerical vectors with one or more dimensions (CICHOSZ, 2015). For this, we must represent the text numerically. Considering two vectors a and b in a BOW, we calculate the cosine according to Equation 4.

$$Cosine(a, b) = \frac{a \cdot b}{|a||b|} \tag{4}$$

2. *Negative Squared Error (NSE)*: This metric compares two vectors and measures the negative of the square of Euclidean distance between them. NSE is suitable for clustering optimisation problems where the goal is to minimise the square error, such as the Affinity Propagation clustering (FREY; DUECK, 2007). Considering two vectors a and b in a BOW, we calculate the NSE according to Equation 5:

$$NSE(a, b) = -|x_a - y_b|^2 \tag{5}$$

3.4.3.2 Visualisation

Throughout the clusters analysis process, we used a set of visualisation tools to better understand our data and the cluster generated by the clustering algorithms. The idea is to use those techniques to identify the common reason that relates the documents with the others in the same cluster.

1. *Word Cloud*: This technique creates an image of words in different sizes and colours with the goal of showing in its center the words with the biggest weights. In the borders, the less important words stand (KWARTLER, 2017). Then, Word Cloud helps to verify the predominant words or N-grams in a set of documents.

2. *Latent Dirichlet Allocation (LDA) and Latent Semantic Indexing (LSI)*: These are topic modelling techniques that help to identify the main topics of each cluster. LDA states that each document in a corpus is a mixture of a set of topics and each topic is characterised by a distribution of words. Thus, for each document, LDA shows the corresponding list of topics and relevant words. LSI also reveals topics in the documents, but from the relationship between words (CAMPBELL *et al.*, 2015; PAPADIMITRIOU *et al.*, 2000).

3. *Principal Component Analysis (PCA)*: This statistical technique allows to visualise the distribution of clusters in a two-dimensional representation. PCA reduces the dimensionality of a dataset consisting of a large number of interrelated variables (for example, a text dataset) while retaining as much of its variation. This is achieved by transforming to a new set of variables, the principal components (PCs), which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables (JOLLIFFE; CADIMA, 2016).

3.4.4 Association

Another ML unsupervised learning tasks is association, which is used to find hidden associations in large sets of data items. An association rule is an implication expression of the form $x \rightarrow y$, where x and y are disjoint itemsets (TAN *et al.*, 2014).

One of the popular applications of this ML task is the “market basket analysis”, which finds co-occurrences of an item along with another within the same transaction. Thus, the rule indicates that, based on the history of all transactions, when x is found in a transaction, there is a strong propensity for y to occur within the same transaction. The x is the antecedent or premise of the rule and the y is the consequent or conclusion of the rule. The antecedent and consequent of the association rule can contain more than one item (KOTU; DESHPANDE, 2019).

The most common association algorithm is the *Frequent Pattern Growth (FP-Growth)*, which generates a tree from frequent patterns by scanning the whole dataset following the support threshold. Then the rules are formed by constructing a conditional tree, which saves the costly dataset scans in the subsequent mining processes (HAN *et al.*, 2000).

3.4.5 Classification

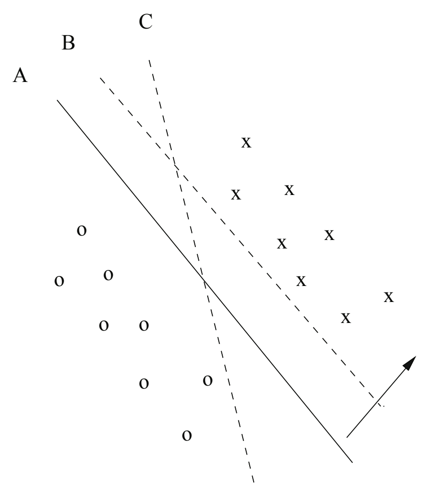
As introduced in the 2.1.2, one of the supervised learning tasks is classification. While in clustering we do not know the classes or categories, in classification we divide the corpus into classes that are previously defined. For example, we can extract examples from a news portal on political matters that might attach one of three labels: “senate”, “congress”, and “legislation”. Then, for a given set of examples in which labels are not available, the goal is to place them in one of these categories (AGGARWAL, 2018).

We employed these approaches for the classification task: (1) linear-based models: Support Vector Machine, Logistic Regression and Naive Bayes; (2) tree-based model: Random Forest; (3) instance-based model: k Nearest Neighbours; and (4) neural-based model: Neural Networks.

1. *Support Vector Machine (SVM)*: The main goal of this classifier is to determine separators in the search space which can best separate the different classes. For example, in Figure 11, there are two classes denoted by ’x’ and ’o’ and three separation hyperplanes denoted by A, B and C. Hyperplane A provides the best separation between the different classes because the normal distance of any of the data points from it is the largest (AGGARWAL; ZHAI, 2012). The data that is not linearly separable in the original input space may be easily separable in the higher dimensional space. When we cannot separate the classes by a line as in Figure 11 , we can add a Kernel function to the SVM, which embeds the data into a higher-dimensional space. From this, we can build new forms of hyperplanes (RUSSELL; NORVIG, 2010). One popular Kernel function is the Radial Basis Function (RBF), which is particularly useful in supervised settings like classification in which one can measure the algorithm performance on the labelled data to tune parameters (AGGARWAL, 2018).

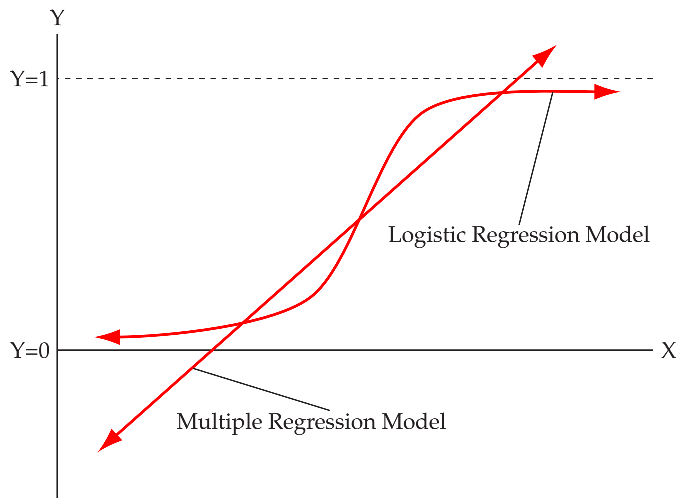
2. *Logistic Regression (LR)*: This classifier is a statistically based technique with the purpose of classifying objects into distinct groups based on the characteristics of the object. Like most classification models, LR is designed to predict a probability value of an event occurring (i.e., the probability of an observation being in the group coded 1 versus the group coded 0). However, there are fundamental differences between LR and other models predicting metric outcomes, such as multiple regression. Because the binary dependent variable has only the values of 0 and 1, the predicted value (probability) must be bounded to fall within the same range. To define this relationship bounded by 0 and 1, LR uses the logistic curve to represent the relationship between the independent and dependent variables

Figure 11 – (SVM) Which is the best separation hyperplane?



Source: Aggarwal and Zhai (2012).

Figure 12 – Logistic Regression versus Multiple Regression



Source: Hair *et al.* (2019).

(shown in Figure 12). Using this form of the relationship allows for a direct estimation of a nonlinear relationship (HAIR *et al.*, 2019). Although this higher flexibility may be desirable in general, it carries with it a higher risk for model overfitting ("memorizing the training cases"), which can potentially reduce a model's accuracy on previously unseen cases. In predictive modelling, fitting the training cases is just part of the task: correctly classifying new cases is the most important goal (DREISEITL; OHNO-MACHADO, 2002).

3. *Naive Bayes (NB)*: This classifier uses a probabilistic generative model, which assumes that the corpus is generated from a mixture of different classes. The gener-

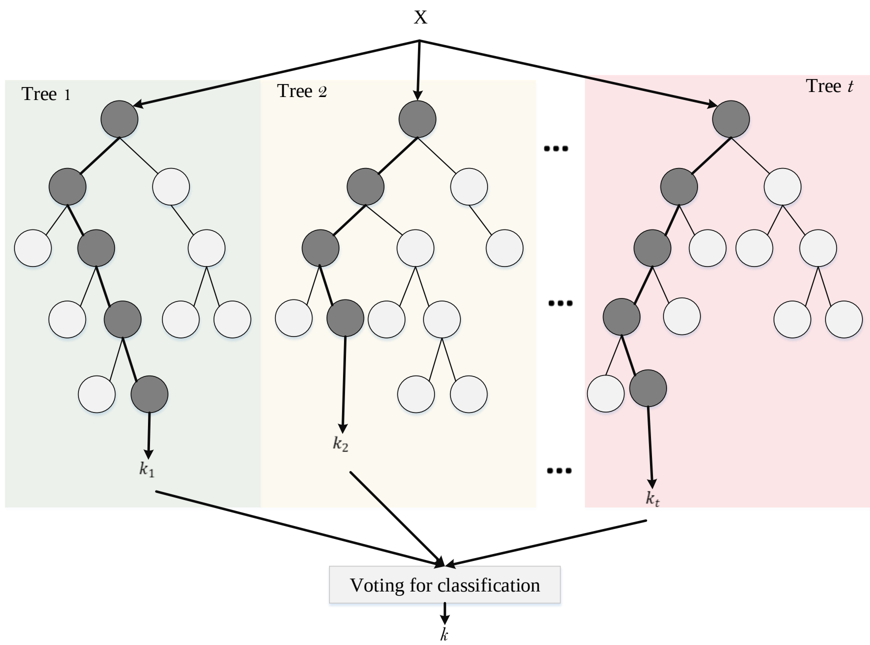
ative process, which is applied once for each observed document, is a) select the r-th class (mixture component) C_r with prior probability $\alpha_r = P(C_r)$; b) generate the next document from the probability distribution for C_r. The observed (training and test) data are assumed to be outcomes of this generative process, and the parameters of this generating process are estimated so that the log-likelihood of this data set being created by the generative process is maximised. Generally, only the training data is used to estimate the parameters, because the training data contains additional information about the identity of the mixture component that generated each document. Subsequently, these parameters are used to estimate the probability of the generation of each unlabelled test document from each mixture component (class). This results in a probabilistic classification of unlabelled documents (AGGARWAL, 2018).

4. *Random Forest (RF)*: This classifier is an ensemble of decision trees. The idea of a decision tree is to partition the data space based on a series of split conditions on the attributes. In the training phase, the data space is partitioned into attribute regions that are heavily biased towards a particular class label. Then, during the testing phase, it identifies the relevant partition of the data space for the test instance and returns a label. Each node in the decision tree corresponds to a region of the data space defined by the split conditions at its ancestor nodes, and the root node corresponds to the entire data space (AGGARWAL, 2018). RF combines decision trees classifiers as shown in Figure 13. Each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. After training all trees as forests, we have predictions assigned based on voting the most popular class (KOWSARI *et al.*, 2019).

5. *k Nearest Neighbours (kNN)*: The basic idea of this classifier is to identify the k nearest neighbours of a test point and compute the number of points that belong to each class. The class with the largest number of points is reported as the relevant one. We must use metrics to compute the nearest neighbours, such as Cosine similarity or Euclidean distance. The kNN classification can be used for both binary classes and multi-way classes (AGGARWAL, 2018). Given a test document x, the similarity of x and each neighbour document is the score of the category of the neighbour document. Several of the k nearest neighbour documents may belong to the same category. Thus, the sum of the score of that category would be the similarity score of class k concerning the test document x. By sorting the scores of the candidate categories, the algorithm assigns the candidate category with the highest score to the test document x (JIANG *et al.*, 2012).

6. *Neural Networks (NN)*: These are computational models inspired by the nervous system of living beings. They have the ability to acquire and maintain knowledge

Figure 13 – Random Forest

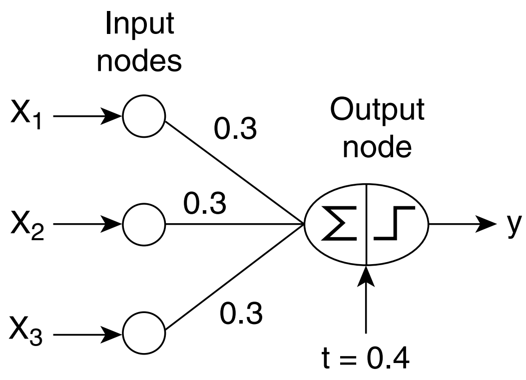


Source: Kowsari *et al.* (2019).

(information based) and can be defined as a set of processing units, represented by artificial neurons, interlinked by a lot of interconnections (artificial synapses), implemented by vectors and matrices of synaptic weights (SILVA, I. N. d. *et al.*, 2018). Figure 14 illustrates a simple NN architecture known as a perceptron. The perceptron consists of two types of nodes: input nodes, which are used to represent the input attributes, and an output node, which is used to represent the model output. The nodes in a neural network architecture are commonly known as neurons or units. Each input node is connected via a weighted link to the output node. The weighted link is used to emulate the strength of synaptic connection between neurons. As in biological neural systems, training a perceptron model amounts to adapting the weights of the links until they fit the input-output relationships of the underlying data. The perceptron is a single-layer NN because it has only one layer of nodes. However, we can add several intermediary layers between the input and output nodes, performing more complex mathematical operations (TAN *et al.*, 2014).

Recently, DL models have been proven to be effective in text classification, especially the *Convolution Neural Networks (CNN)*. This classifier use convolutional masks to sequentially convolve over the data. For texts, a simple mechanism is to recursively convolve the nearby lower-level vectors in the sequence to compose higher-level vectors. Similar to images, such convolution can naturally represent different levels of semantics shown by the text data (PENG *et al.*, 2018).

Figure 14 – Perceptron



Source: Ivan Nunes da Silva *et al.* (2018).

3.4.6 Regression

Another ML supervised learning tasks is regression. As in classification, we use regression to predict a variable. The main difference between them is that the output variable in regression is numerical (or continuous) while that for classification is categorical (or discrete) (AGGARWAL, 2018).

Classification models can be adapted for the regression task. Thus, we employed these approaches: a) linear-based model: SVM; b) tree-based/ensemble models: RF, Bagging and Boosting algorithms (Gradient Boosting, Adaboost and Extreme Gradient Boosting); and c) neural-based model: NN. Since the algorithms SVM, RF and NN have been explained in the previous section, we will only discuss the others below.

1. *Bagging*: Bagging and Boosting are two popular ensemble techniques that utilise different re-sampling methods to create diverse training data for obtaining different base models (such as Decision Trees, NN). To produce diverse base models despite similar training datasets, we often forced them to be weak (moderately accurate). Thus, for a given training data set with sample size n, Bagging generates k new training set, each with sample size n, by sampling from the original training data set uniformly and with replacement. Through sampling with replacement, some observations appear more than once in the Bagging sample, while other observations will be 'left out' of the sample. Then, k base models are trained using the newly generated k training set and combined through averaging (regression problem) or majority voting (classification problem). Bagging improves prediction accuracy through diversity of each base models (ZHANG; HAGHANI, 2015).

2. *Boosting*: Different from Bagging, the Boosting method generates base models sequentially. Prediction accuracy is improved through developing multiple models in sequence by putting emphasis on these training cases that are difficult to estimate. In the Boosting process, examples that are difficult to estimate using the

previous base models appear more often in the training data than the ones that are correctly estimated. Each additional base model is aimed to correct the mistakes made by its previous base models. While the Boosting method strategically resamples the training data to provide the most useful information for each consecutive model, in the Bagging method each sample is uniformly selected to produce a training dataset (ZHANG; HAGHANI, 2015). In this work, we applied three Boosting algorithms, *Gradient Boosting (GBoost)*, *Adaboost* and *Extreme Gradient Boosting (XGBoost)*, whose difference is the mathematical function.

Finally, due to the inner differences among the regression techniques, they can achieve better or worse performances in different situations. Thus, it may be useful to apply some of those models together, so they complement one another. The final prediction of this combination is the average output among the models. This approach is called *Ensemble Voting* (MENDES-MOREIRA *et al.*, 2012).

3.4.7 Training and test

There are different sampling methods for creating training and test datasets. We can divide the dataset using Hold-out or Cross-validation.

1. *Hold-out*: In this method, we split a fraction of the dataset for training while the remaining fraction we use for testing. The model learns on the training set, and then we use the test set to evaluate how well that model performs on unseen data. A common split is using 70% of data for training and the remaining 30% of the data for testing (AGGARWAL, 2018). However, this method can introduce some bias in the pipeline, because the distribution of the examples may not be similar in those two subsets, specially in small datasets (HAWKINS, 2004).

2. *Cross-validation*: In this method, we divide the dataset into different fractions of approximate sizes. Then at each training stage, one set is used for testing and the others for training, in such a way that each set created will be used both for training and testing (WONG, 2015). By evaluating the models several times using different and random train and test sets, the prediction quality measurements could be more precise (KUHN; JOHNSON, *et al.*, 2013).

3.4.8 Evaluation

Each ML task (clustering, association, classification and regression) has metrics to evaluate the quality of its output, which we explain below.

For the clustering task (ROUSSEEUW, 1987; MANNING *et al.*, 2010; VANDEGINSTE *et al.*, 1998):

1. *Clustering Tendency (CT)*: A general metric based on Hopkin's Statistic, which states that if there is a clustering tendency in the data, the distance between a point and its nearest neighbours will be much smaller than the distance between randomly selected artificial points. The value of CT varies from zero to one, where a higher value means higher tendency and vice versa. However, we need values significantly bigger than 0.5 to affirm the existence of tendency. CT is calculated as shown in Equation 6, where d_i is the distance to the nearest data points for each selected data point (we have to randomly select a small number of the real data points, for example 5%); and u_j is the distance to the nearest real data point for each artificial point.

$$CT = \frac{\sum u_j}{\sum u_j + \sum d_i} \tag{6}$$

2. *Silhouette Score (SS)*: An internal metric used to evaluate the degree of cohesiveness of the clusters. SS indicates whether the documents lie well within their clusters or they are placed between clusters. The SS is higher when the documents are close to the centroid of a cluster and far from all the others and it is lower when these distances are similar. This metric is very useful to select the best number of clusters. The SS is calculated using the mean intra-cluster distance a and the mean nearest-cluster distance b for each sample, as shown in Equation 7.

$$SS = \frac{(b - a)}{max(a, b)} \tag{7}$$

The value of SS varies from -1 to 1. If the score is 1, the cluster is dense and well-separated than other clusters. A value near 0 represents overlapping clusters with samples very close to the decision boundary of the neighbouring clusters.

3. *Entropy*: An external metric use to scale the amount of disorder in a system. Having the correct cluster for each document, entropy is higher when the documents inside a cluster have a higher variety of distinct ground truth labels. So, when the entropy value is lower, the clustering quality is better. First, we calculated the entropy of each cluster separately, according to Equation 8, where ω is a given cluster, c is the description we assigned to the this cluster (see Appendixes B, C, D and E), C is the set of all our descriptions, $|\omega_c|$ is the count of documents classified as c in cluster ω and n_ω is the count of documents in cluster ω.

$$E(\omega) = -\sum_{c \in C} \frac{|\omega_c|}{n_\omega} \log_2(\frac{|\omega_c|}{n_\omega}) \tag{8}$$

Then, we calculated the total entropy in the results of each technique applied (Hierarchical Clustering, Lingo, K-means and Affinity Propagation), according to

Equation 9, where Ω={$\omega_1, \omega_2, ..., \omega_k$} is the set of clusters, $E(\omega)$ is a single clusters entropy, N_ω is the number of documents in cluster ω and N is the total number of documents.

$$E(\Omega) = \sum_{\omega \in \Omega} E(\omega)\frac{|N_\omega|}{N} \tag{9}$$

4. *Purity*: Another external metric relates to the proportion of documents assigned to the most frequent ground truth in the cluster. In other words, the degree of uniformity of the documents in a cluster. Thus, a higher purity implies a better clustering. We calculate the Purity according to Equation 10, where N is the number of documents, k is the number of clusters, c_i is a cluster in C, and t_j is the description which has the maximum count for cluster c_i.

$$Purity = \frac{1}{N}\sum_{i=1}^{k} max_j|c_i \cap t_j| \tag{10}$$

For the association task (TAN *et al.*, 2014):

1. *Support*: The ratio of transactions containing the set of items of the association rule. Given the form $x \rightarrow y$, is the frequency that a combination between x and y occurs considering all cases in the dataset (N), according to Equation 11.

$$Support = \frac{Frequency(x, y)}{N} \tag{11}$$

2. *Confidence*: The ratio of correct results of the association rule, given the set of items of the support. Considering the form $x \rightarrow y$, is the frequency that x cases have y in relation to the frequency in which x occurs, according to Equation 12.

$$Confidence = \frac{Frequency(x, y)}{Frequency(x)} \tag{12}$$

For the classification task (AGGARWAL, 2018):

1. *Precision*: The percentage of positive instances correctly predicted (true positives) which belong to all positive instances predicted (true positives and false positives), as shown in Equation 13.

$$Precision = \frac{TP}{TP + FP} \tag{13}$$

2. *Recall*: The percentage of positive instances correctly predicted (true positives) which belong to all instances predicted recommended as positives (true positives and false negatives), as shown in Equation 14.

$$Recall = \frac{TP}{TP + FN} \tag{14}$$

3. *F1-score*: The harmonic mean between the precision and the recall, as shown in Equation 15.

$$F1 = \frac{2 * TP}{2 * TP + FP + FN} \tag{15}$$

4. *Accuracy*: The percentage of test instances in which the predicted value matches the ground-truth value, that is, the ratio of all hits to the total, as shown in Equation 16.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{16}$$

For the regression task (AGGARWAL, 2018; CHAI; DRAXLER, 2014; DEVORE, 2011):

1. *Root Mean Square Error (RMSE)*: The average of the errors of the square differences between the predicted (y_i) and the actual ($\hat{y}_i$) values, as shown in Equation 17. This metric is more sensitive to outliers and tend to penalise more bigger errors.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{17}$$

2. *Mean Absolute Error (MAE)*: The average of the errors when predicting the dependent variable, as shown in Equation 18. MAE is also simple to interpret and it is less sensible to outliers than RMSE.

$$MAE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)}{n} \tag{18}$$

3. R^2: The proportion of observed variation in the predicted values that can be explained by the regression model, as shown in Equation 19. So, the higher R^2, the better the model can explain the variation in the predictions (y).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{19}$$

3.5 EXPERIMENTS, RESULTS AND DISCUSSION

In this section, we present the experiments performed with the techniques previously conceptualised. We note that we conducted the experiments from 2019 to 2021, while the dataset was increased over time. Our experimental setup follows these four goals and ML tasks:

(1) First, since the data is textual and unstructured, we intend to discover some patterns that help us in extracting attributes (judgement factors) and labels (judgement results).

(2) Second, once we extracted and structured judgement factors and results, we want to find some relationships between them.

For the 1st and 2nd goals, we resorted to unsupervised learning, respectively clustering and association tasks.

(3) Third, we aim to predict the verdict, i.e., whether the consumer wins or not the lawsuit (categorical judgement result)

(4) Finally, once the consumer wins, we want to predict how much he or she will receive as immaterial damage compensation (numerical judgement result).

For the 3rd and 4th goals, we relied on supervised learning, respectively classification and regression tasks.

3.5.1 Technical support

To perform the experiments, we reused existing implementations and standard methods, including Orange 3 (DEMŠAR *et al.*, 2013), Carrot² (OSIŃSKI; WEISS, D., 2019), Python language and a set of open-source libraries. Both software have a user-friendly interface that allows professionals from other areas of knowledge to handle it independently, as well as tutorials that facilitate understanding the techniques.

To introduce the Orange 3 and Carrot² use and to implement the experiments in Python language it was indispensable the collaboration of another researcher with programming skills and ML knowledge. For more details about mathematical explanations, we refer to his dissertation (DAL PONT, 2021), which is related to the same interdisciplinary project. Here we will show the experiments with best results involving direct participation of the PhD student in Law. Other applications using complex techniques (such as CNN and Word Embedding) can be found in his research.

Additionally, this work required the use of a high-performance computer available from the E-government, Digital Inclusion and Knowledge Society (EGOV) research group at UFSC.

3.5.2 [A] Clustering to guide the judgement factors extraction

In this experiment[8] we worked with a dataset of 665 documents. At the end of preprocessing step, the dataset resulted in a corpus of 910 types of tokens, with a total of 401,932 tokens. We first calculated the Clustering Tendency, which resulted in 0.88, indicating a high potential for the application of clustering techniques.

By applying the clustering algorithms (Hierarchical and Lingo - soft clustering; K-means and Affinity - hard clustering), we evaluated the results based on the following criteria: (1) entropy and purity; (2) algorithm's ability in providing descriptions/topics; (3) legal evaluation; and (4) experimental complexity. The criteria 3 corresponds to our evaluation of each cluster in order to identify a law event in common between the documents belonging to it. For this, we considered the output quality (algorithm mistakes and successes) and the our difficulty in identifying a common topic. Table 5 the parameters we used in each technique. In those implemented through Orange 3 and Carrot[2], we have adopted the parameters recommended by the tools.

Table 5 – Clustering setup parameters

Tool	Technique	Parameters
Orange	Hierarchical Clustering	Representation: TF; Distance Metric: Cosine Similarity; Linkage: Ward's method; Height Ratio: 14%.
Orange	K-means	Representation: TF; Distance Metric: Cosine Similarity; K values: from 4 to 30; Iteration Limit: 100.
Carrot	Lingo	Representation: TF-IDF; Distance Metric: Cosine Similarity; Cluster Count Base: 30%; Factorisation Method: Nonnegative Matrix (ED Factory).
Python Programming Language and Scikit-Learn Library	Affinity Propagation	Representation: TF; Distance Metric: NSE; K values: unlimited; Iteration Limit: 5.000.

The list of all clusters generated by each algorithm, with the appropriate descriptions (those given by us and those given by Lingo), are in the Appendix B, C, D and E. Here we present and discuss some of the clusters generated in each approach. These descriptions will later be used as attributes/independent variables/judgement factors and as labels/dependent variables/judgement results, which will be explained in the next subsection.

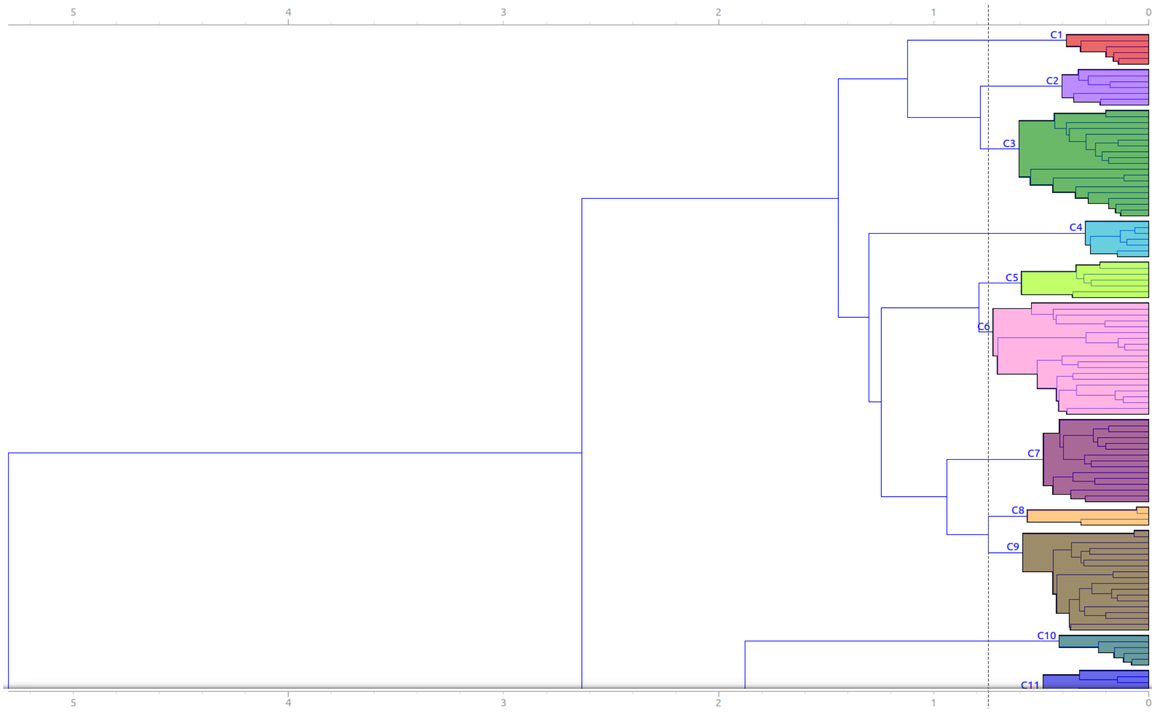
Results from Hierarchical Clustering

Hierarchical clustering resulted in 76 clusters and is represented by a partial dendrogram in Figure 15. The dotted line, called height ratio, signals the point of cluster

8 The results discussed below are adapted from the paper Sabo *et al.* (2021).

selection. Each horizontal path of the dendrogram that crosses the dotted line raises a cluster. The clusters described are in the Appendix B.

Figure 15 – (Hierarchical clustering) Partial dendrogram in Orange 3



1. *Entropy and purity*: Hierarchical clustering achieved an entropy of 0.6021 and purity of 0.7875.

2. *Algorithm's ability in providing descriptions*: Not available.

3. *Legal expert's evaluation*:

   a) *Output quality*: Hierarchical clustering generated a satisfactory output, grouping judgements with common law events. To exemplify this, we emphasise two clusters, C10 and C30.

   - **C10. Promotional ticket offer not fulfilled by a specific airline / Well-founded or partly founded (6 documents)**: This cluster contains all judgements about a ticket sale by a specific airline on the day known as "Cybermonday". In this case, the airline decided, unilaterally and without any justification, not to issue all the tickets already purchased. In Brazilian Consumer Law, once something is offered, the airline is obliged to fulfil it.
   - **C30. Notice about the existence of a previous lawsuit that is identical or similar to the current one (2 documents)**: C30 contains two specific judgements about cases dismissed without prejudice. One is about *lis pendens* (or notice of pending action), which means in Brazilian Procedure Law that a previous identical lawsuit

was filed. The other is about *connection* (or related lawsuits), which means in Brazilian Procedure Law that a previous similar lawsuit was filed. However, the word 'connection' is also related to the word 'flight'. It has been widely used in other legal judgements (e.g., when the consumer misses the connecting flight due to the delay or cancellation of the previous flight). Hierarchical clustering was able to group that legal judgement in the correct cluster.

b) *Difficulty in identifying law events*: Analysing many clusters with fewer documents was less complex. In addition, word cloud and topic modelling (LSI and LDA) supported the identification of law events. To exemplify this, we present the word cloud and the topic keywords generated in two clusters, C1 and C47.

- **C1. Permanent baggage loss / Well-founded or partly founded (6 documents)**: Figure 16 shows the word cloud output, and Table 6 indicates the LSI and LDA topics. With these highlighted keywords, we can easily recognise the C1 law event and assign it as the cluster description.

Figure 16 – (Hierarchical clustering) C1 Word cloud

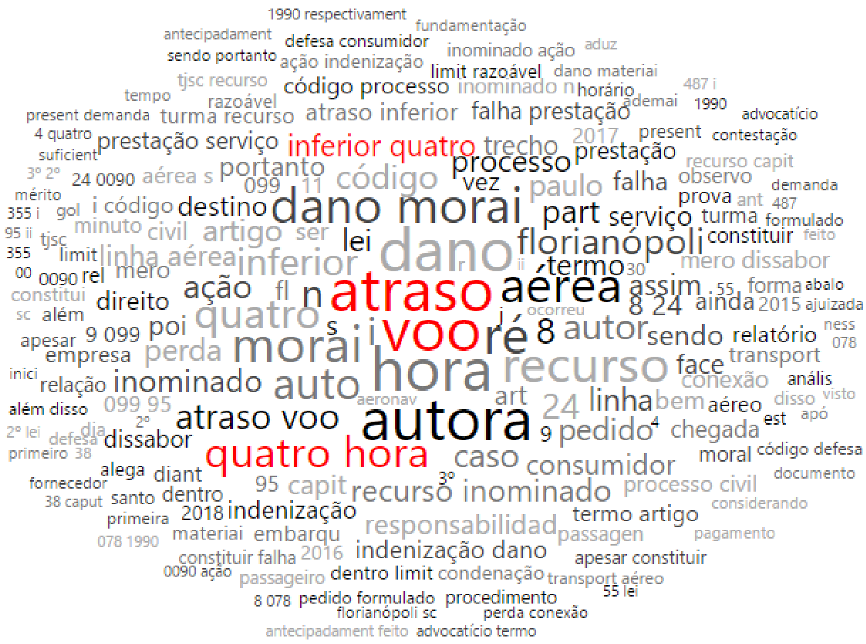
* Translation of the words in red: baggage loss; baggage; loss.

Table 6 – (Hierarchical clustering) C1 Topic keywords

LSI	damage; civil; plaintiff; **baggage**; moral; moral compensation; defendant; material; compensation; air
LDA	damage; plaintiff; civil; **baggage**; appeal; material; moral; compensation; moral damage; **loss**

- **C47. Under-four-hour flight delay / Not founded (8 documents)**: Figure 17 shows the word cloud output, and Table 7 indicates the LSI and LDA topics. With these highlighted keywords, we can easily recognise the C47 law event and assign it as the cluster description.

Figure 17 – (Hierarchical clustering) C47 Word cloud



* Translation of the words in red: under four; delay; flight; four hours.

Table 7 – (Hierarchical clustering) C47 Topic keywords

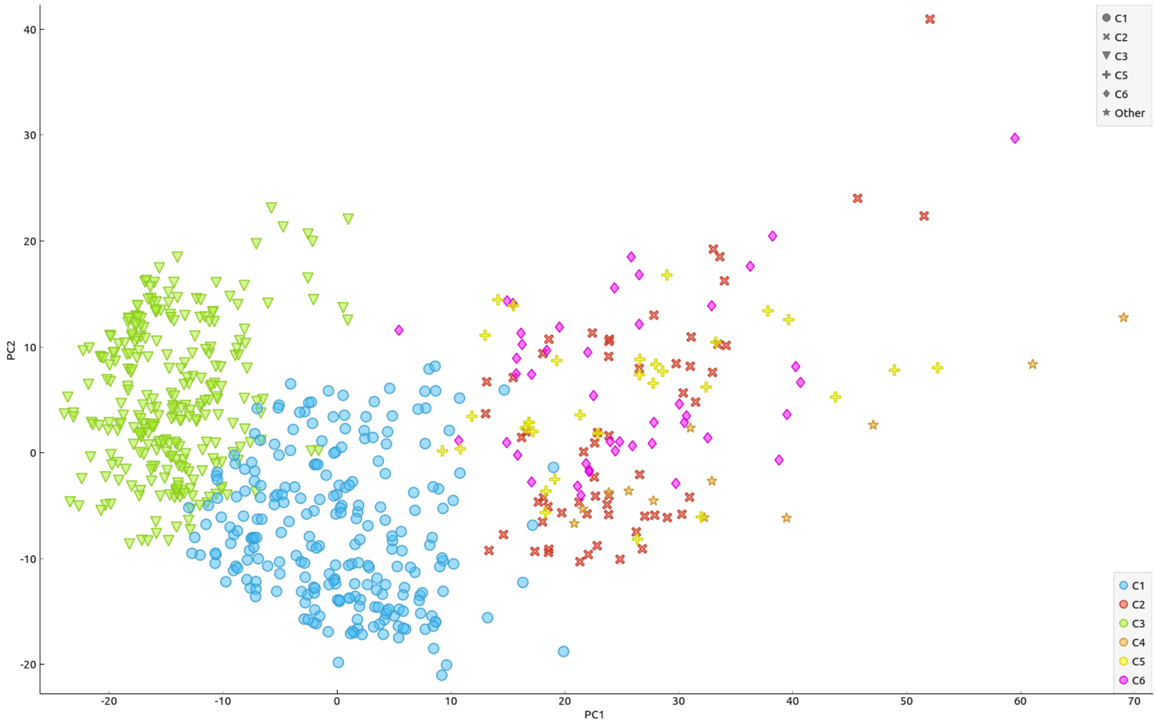
LSI	**flight**; **delay**; hour; damage; plaintiff; appeal; air; moral; defendant; moral damage
LDA	**delay**; **flight**; hour; appeal; damage; plaintiff; **four hour**; four; moral

4. *Experimental complexity*: Using Orange 3 we run the experiments of the Hierarchical clustering without any difficulties, since the tools make it easier to test different parameters and visualise the results after the processing step. Also, the fact the hierarchical clustering runs only once makes it significantly faster than K-means, but similar to Lingo.

Results from K-means

Clustering with K-means resulted in 6 clusters and is represented in Figure 18. Using PCA, we show the clusters in a two-dimensional representation. The number of clusters is the product of the best SS calculation (0.081 for a k value of 6 clusters). The clusters described are in the Appendix C.

Figure 18 – (K-means) Two-dimensional representation of documents and their clusters in Orange 3



1. *Entropy and purity*: K-means clustering achieved an entropy of 1.7242 and purity of 0.5854.

2. *Algorithm's ability in providing descriptions*: Not available.

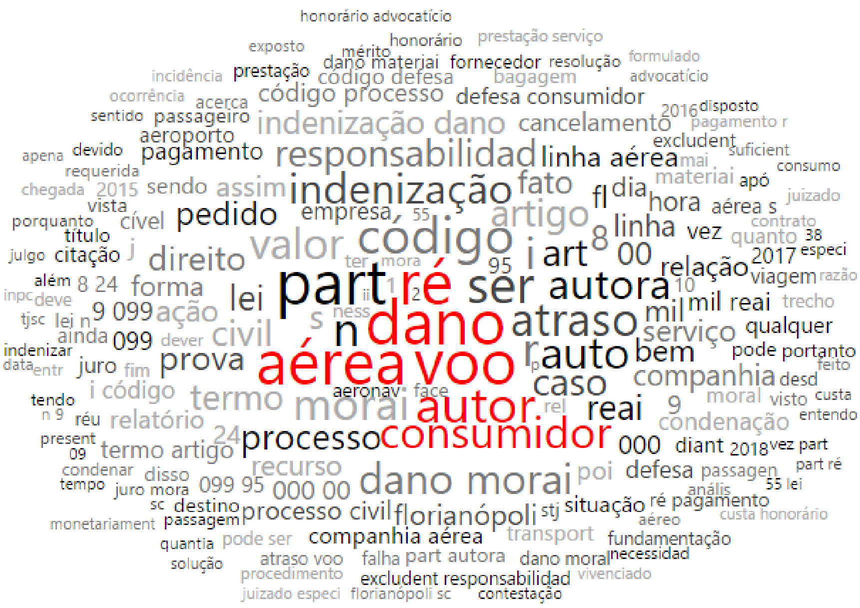
3. *Legal expert's evaluation*:

   a) *Output quality*: K-means generated an unsatisfactory output, because the algorithm groups many judgements with different law events. An exception to this is cluster C4, in which the algorithm groups judgements prepared by an assistant judge. Considering the small number of clusters, we expected K-means to generate a cluster with all cases not founded. However, this did not happen.

   b) *Difficulty in identifying law events*: In contrast to hierarchical clustering, analysing few clusters with many documents was more complex. This analysis was the most difficult of all. In four clusters, we was unable to identify only one law event in common. Also, word cloud and topic modelling (LSI and LDA) did not assist the analysis, because the keywords were too generic. To exemplify this, we present the word cloud and the topic keywords from the cluster C1, in which there are several law events.

      - **C1. Different law events (243 documents)**: Figure 19 shows the word cloud output, and Table 8 indicates the LSI and LDA topics. These keywords appear in all 665 judgements, making it impossible to identify any shared law event from them.

Figure 19 – (K-means) C1 Word cloud



* Translation of the words in red: defendant; damage; air; plaintiff; consumer.

Table 8 – (K-means) C1 Topic keywords

LSI	defendant; damage; flight; code; moral; party; file
LDA	defendant; damage; flight; plaintiff; code; file; air; moral

4. *Experimental complexity*: As described, using Orange to run the K-means is a simple process. However, K-means tends to take considerably longer time than hierarchical clustering and Lingo, because the software executes the algorithm several times with multiple k values.

Results from Lingo

Clustering with Lingo resulted in 63 clusters and is represented in Figure 20. The larger the region representing a cluster, the larger the number of documents in it. We suppressed the personal names and replaced with "***". The clusters described are in the Appendix D.

1. *Entropy and purity*: Not available.

2. *Algorithm's ability in providing descriptions*: Available.

Figure 20 – Lingo clustering in Carrot²

3. *Legal expert's evaluation*:

   a) *Output quality*: Lingo generated a satisfactory output. To exemplify this, we cite three situations. The first is that Lingo created five clusters whose description is the defendant airline name (C1, C6, C8, C11, C19). One of them has the highest score (C1). This is a different result compared to the others approaches. The second is that Lingo, as the other approaches, also grouped judgements referring to problems such as flight delays, cancellations and changes (C4, C15, C16, C26, C33, C42, C56, C61) and baggage loss or stolen items (C2, C23, C30, C34, C41, C51). The third situation is that Lingo created clusters whose descriptions are irrelevant, such as the page number of a procedural file (C38, C50). This last situation was less frequent.

   b) *Difficulty in identifying law events*: Considering the algorithm's ability in providing descriptions, our analysis consisted only in verifying the consistency of those generated by Lingo (whether they make legal sense). We observe that some descriptions may be confusing to those who are not familiar with the subject of the database due to removal of stop words in the preprocessing step. This reinforces the need for a legal expert analysis.

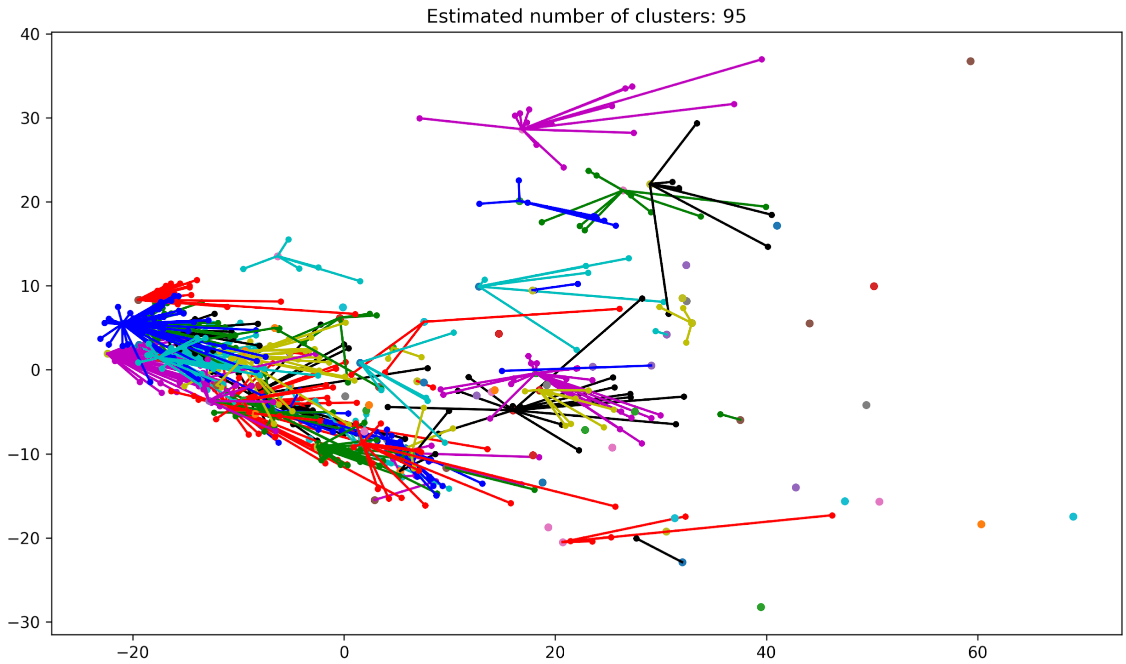
4. *Experimental complexity*: Using Carrot² to run the Lingo algorithm is also simple, although we needed to adjust the preprocessing step in this case. It takes similar

time as from hierarchical clustering to run the algorithm.

Results from Affinity Propagation

Clustering with Affinity Propagation resulted in 95 clusters and is represented in Figure 21. Using PCA, we show the clusters and their corresponding documents represented as stars, which are the documents of each cluster connected to the central document in that cluster. As well as in K-means, the number of clusters is the product of the best SS calculation (0.081 for a k value of 95 clusters). The clusters described are in the Appendix E.

Figure 21 – Affinity Propagation clustering



1. *Entropy and purity*: Affinity Propagation achieved an entropy of 1.3909 and purity of 0.6062.

2. *Algorithm's ability in providing descriptions*: Not available.

3. *Legal expert's evaluation*:

    a) *Output quality*: Affinity Propagation generated an unsatisfactory output, because the algorithm created many clusters with only one document (38 clusters of 95). Some of these single-document clusters could be grouped together. Examples are clusters C4, C5, and C6; and also clusters C21, C23, and C26.

    b) *Difficulty in identifying law events*: As opposed to K-means and similar to hierarchical clustering, analysing many clusters with fewer documents was less complex. Likewise, word cloud and topic modelling (LSI and LDA) supported the identification of law events. To exemplify this, we present the word cloud and the topic keywords generated in two clusters, C29 and C52.

- **C29. Baggage damaged / International flight / Cases subject to the Montreal and Warsaw Convention / Partly founded (2 documents)**: Figure 22 shows the word cloud output, and Table 9 indicates the LSI and LDA topics. With these highlighted keywords, we can easily recognise the C29 law event and assign it as the cluster description.

Figure 22 – (Affinity Propagation) C29 Word cloud

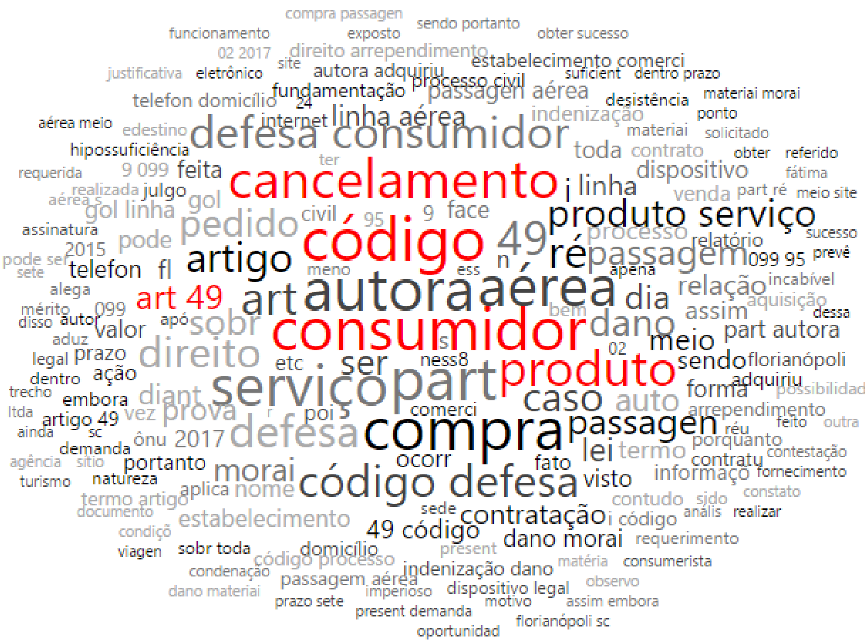
* Translation of the words in red: international; break; damage; baggage; Warsaw.

Table 9 – (Affinity Propagation) C29 Topic keywords

LSI	**damaged**; **baggage**; plaintiff; appeal; case; defendant; material; compensation; **break**; value
LDA	damaged; baggage; plaintiff; appeal; case; defendant; material; compensation; **break**; value

- **C52. Right to regret denied / Not founded (3 documents)**: Figure 23 shows the word cloud output, and Table 10 indicates the LSI and LDA topics. The keywords "49" and "art" correspond to the article number in the Brazilian Code of Consumer Protection that provides the right to regret. As explained in Section 2.1, this is the possibility for the consumer to waive the purchase within seven days and be repaid. With these highlighted keywords, we can easily recognise the C52 law event and assign it as the cluster description.

Figure 23 – (Affinity Propagation) C52 Word cloud



* Translation of the words in red: cancellation; code; art 49; consumer; product.

Table 10 – (Affinity Propagation) C52 Topic keywords

LSI	**consumer**; plaintiff; **purchase**; party; **code**; service; **cancellation**; air; **product**; **49**
LDA	**consumer**; air; plaintiff; service; **code**; **product**; party; **purchase**; **49**; **art**

4. *Experimental complexity*: Affinity Propagation required the implementation of the whole pipeline from preprocessing to visualisation of the clusters, which makes it the most laborious algorithm. It takes a little longer when compared to Lingo and hierarchical clustering.

We summarise the evaluation of the four approaches into Table 11. Weighing the four criteria, we can state that the most advantageous was hierarchical clustering, since the best entropy and purity, the least difficulty for us to analyse the clusters, and the least experimental complexity.

Table 11 – Clustering evaluation overview

	Entropy	**Purity**	**Algorithm's ability in providing descriptions**	**Legal expert's evaluation**		**Experimental complexity**
				Output quality	Difficulty in identifying law events	
Hierarchical clustering	0.6021	0.7875	Not available	Satisfactory	Less difficult	Less complex
K-Means	1.7242	0.5854	Not available	Unsatisfactory	More difficult	More complex
Lingo	-	-	Available	Satisfactory	Less difficult	Less complex
Affinity Propagation	1.3909	0.6062	Not available	Unsatisfactory	Less difficult	More complex

3.5.3 [B] Association to find relationships between judgement factors and results

Considering the clustering results, which enabled us to identify relevant law events on the content of the judgements, we extracted some information from the text and structured it in a XLS format document. We divided this information into two groups:

- **Attributes | Independent Variables | Judgement Factors**: In ML, attributes are the predictors that affect a given outcome. Some authors use feature as a synonym for attribute (e.g., in feature-subset selection) (KOHAVI; PROVOST, 1998). In Statistics, they correspond to the independent variables, whose value does not depends on other variables, assuming to have a direct effect on the dependent variable (MORGAN *et al.*, 2004). Bringing to the Law context, this refers to the factors of the judicial decision, which the judge takes into consideration in his/her decision-making, and which do not depend on his/her verdict.
- **Labels | Dependent Variables | Judgement Results**: In ML, a label value is what will be predicted, the outcome (KOHAVI; PROVOST, 1998). In Statistics, it corresponds to the dependent variable, whose value depends on changes in the independent variables, assumed to be the cause of the independent variables (MORGAN *et al.*, 2004). In Law, this refers to the judgement result, the judge's verdict after weighing the factors.

As attributes / independent variables / judgement factors, we extracted the following:

1. *Date of judgement*: The judge's perspectives may change over time. Consequently, the amount of compensation may vary by date. In the dataset, this is a numerical continuous variable, represented by day, month, and year.
2. *Judge*: Since each judge is free to set the amount of compensation according to his/her conviction on the case, this is a important attribute. In the dataset, this is a categorical variable, represented by the name of the thirty one judges who prepared the collected judgements.
3. *Type of judge*: Since in the JEC/UFSC there are three types of judges (chief, assistant, and voluntary) this is a important attribute. The chief judge is responsible for the court and is the one who, as a rule, judges the lawsuits. The assistant or substitute judge is the one who judges when the chief judge needs to be absent. And the voluntary judge is the one who has a law degree but is not invested in the position. He or she voluntarily prepares judgements that are submitted to the approval of the chief judge. An assistant judge can freely fix a different value of compensation than a chief judge. The voluntary judge can do this too, but the chief judge can

modify the value. In the dataset, this is a categorical variable, represented by those three types.

4. *Permanent baggage loss*: It is an event that can generate compensation for immaterial damage. In the dataset, this is a categorical variable, represented by “yes” (when there was a loss) and “no” (when there was no loss).

5. *Tampered baggage*: Depending on the level of damage or in case of missing consumer’s belongings (theft), it is an event that can generate compensation for immaterial damage. In the dataset, this is a categorical variable, represented by “yes” (when there was tampering) and “no” (when there was no tampering).

6. *Temporary baggage loss*: It is an event that can generate compensation for immaterial damage. In the dataset, this a categorical variable, represented by “yes” (when there was a loss) and “no” (when there was no loss).

   - *Loss interval*: It is a sub-attribute. The longer the delay in returning the baggage to the consumer, the greater can be the value of the compensation for immaterial damage. In the dataset, this is a discrete numerical variable, represented by days.

7. *Flight cancellation*: It is an event that can generate compensation for immaterial damage. We consider as flight cancellation those cases with no rebooking or when the destination is changed. In the dataset, this is a categorical variable, represented by “yes” (when there was cancellation) and “no” (when there was no cancellation).

8. *Flight delay*: It is an event that can generate compensation for immaterial damage. We consider as flight delay those cases with rebooking. In the dataset, this is a categorical variable, represented by “yes” (when there was a delay) and “no” (when there was no delay).

   - *Delay interval*: It is a sub-attribute. The longer the delay in rebooking (that is, the longer the interval between the initially contracted flight and the actual flight operated), the greater can be the value of the compensation for immaterial damage. In the dataset, this is a numerical continuous variable, represented by hours and minutes.

9. *Adverse weather conditions*: It is an event that excludes the possibility of compensation for immaterial damage because it is an unpredictable situation. Even the airline effort is not capable of overcoming them, so there is no way to impute liability to it. In the dataset, this is a categorical variable, represented by “yes” (when there was proven bad weather) and “no” (when there was no proven bad weather).

10. *Consumer fault*: It is an event that excludes the possibility of compensation for immaterial damage because it removes the airline's liability. An example of this situation is when the consumer does not arrive at the airport in plenty of time to check his/her flight and bags. In the dataset, this is a categorical variable, represented by "yes" (when there was the consumer fault) and "no" (when there was no consumer fault).

11. *Overbooking*: Selling more tickets for a flight than are available is considered an abusive practice. Thus, it is an event that can generate compensation for immaterial damage. In the dataset, this is a categorical variable, represented by "yes" (when there was overbooking) and "no" (when there was no overbooking).

12. *No show*: Cancellation of the return ticket unilaterally when the consumer does not show up on the outward flight is considered an abusive practice. Thus, it is an event that can generate compensation for immaterial damage. In the dataset, this is a categorical variable, represented by "yes" (when there was cancellation by no show) and "no" (when there was no cancellation by no show).

13. *Right to regret and repayment claim*: Hindering the consumer's repayment when he/she decides to cancel the acquired ticket is an event that can generate compensation for immaterial damage. This situation is known by a sequence of bad experiences (called *via crucis* by judges) that the consumer must face getting the repayment. In the dataset, this is a categorical variable, represented by "yes" (when repayment was hindered) and "no" (when the repayment was not hindered or when there was no claim).

14. *Downgrade*: The airline changes a business class passenger to economy class. Besides a breach of contract, it is also a breach of the consumer's expectation, and, therefore, it is an event that can generate compensation for immaterial damage. In the dataset, this is a categorical variable, represented by "yes" (when there was a downgrade) and "no" (when there was no downgrade).

As labels / dependent variables / judgement results, we extracted the following:

1. *Verdict*: It is the final decision of the judge in relation to the lawsuit (1st-degree decision). As indicated in Section 3.3, it can be *well founded*, *not founded*, *partly founded* and *dismissed without prejudice*. In the dataset, this is a categorical variable, represented by those four possibilities.

2. *Immaterial damage*: It is the amount fixed by the judge as compensation for immaterial damage. In the dataset, this is a numerical continuous variable, represented by the monetary value in Brazilian Reais.

By applying the association algorithm (FP-Growth), we evaluated the results based on the support and confidence. Nevertheless, we did not find any relationship between the categorical and some numerical variables of interest (*delay interval* and *immaterial damage*). Thus, we categorised the *delay interval* and *immaterial damage* into five ranges, assisted by the quantile calculation[9]. These last two labels were represented in Table 12:

Table 12 – Association rules ranges

Range	Delay interval	Range	Immaterial damage
1	0:00:00 - 4:00:00	1	R$ 0,00
2	4:01:00 - 8:00:00	2	R$ 1,00 - R$ 2.000,00
3	8:01:00 - 12:00:00	3	R$ 2.001,00 - R$ 5.000,00
4	12:01:00 - 24:00:00	4	R$ 5.001,00 - R$ 8.000,00
5	24:01:00 - 72:00:00	5	R$ 8.001,00 - R$ 25.000,00

Table 13 presents the parameters we used to select the rules, which we manipulated to find specific ones with more than one attribute / independent variable / judgement factor as antecedent, and the two labels / dependent variables / judgement results (*verdict* and *immaterial damage*) as a consequence, preferably together.

Table 13 – Association setup parameters

Tool	Technique	Parameters
Orange	FP-Growth	Min. Support: 1%
		Min. Confidence: 70%
		Min. Antecedent: 3
		Max. Consequent: 2

Table 14 shows the implication rules selected in this experiment after applying the parameters, ordered by the highest range of immaterial damage.

Overall, we can confirm a relationship between the variable "delay interval" and "immaterial damage", that is, the more an airline delays a consumer's flight, the greater their monetary compensation, and vice versa. Specifically, there is also a relationship between the permanent loss of the consumer's baggage and the highest range of compensation, whereas a temporary loss of 3 days will result in reasonable indemnities. We also note that the practice of unilaterally cancelling the return flight due to the consumer's no-show on the outward flight is associated with a middle range. The null compensation is related to the delay interval of fewer than 4 hours (this was also observed in the results in the clustering, see Figure), and to some culpable practice of the consumers, i.e., when they are late for the flight by themselves (in this case, the lawsuit will also be unfounded).

These rules will serve as an explanation for the parties. Once we can predict a certain result for their process, we will be able to indicate some factors connected to it, based on the relationships verified through the association task.

[9] Given a data sample, quantile of order p or pth quantile, denoted by $q(p)$, is a location measure where p is any proportion, $0 < p < 1$, such that $100p\%$ of observations are smaller than $q(p)$. Some quantiles have particular names, for example, the $1st$ quantile or $25th$ percentile, where $q(0.25) = q1$, which means that each quantile will contain approximately 25% of the data (BUSSAB; MORETTIN, 2010).

Table 14 – Association rules selected

Rule	Antecedent	→	Consequent	Supp.	Conf.
1	Judge=V.P. Type of judge=Chief Flight delay=Yes Delay interval=5	→	Verdict=Partly founded Immaterial damage=5	3%	81,4 %
2	Verdict=Partly founded Judge=V.P. Type of judge=Chief Permanent baggage loss=Yes	→	Immaterial damage=5	1,2%	87,5 %
3	Judge=V.P. Type of judge=Chief Flight cancellation=Yes No show=Yes	→	Verdict=Partly founded Immaterial damage=3	1,4%	72,7%
4	Judge=V.P. Type of judge=Chief Temporary baggage loss=Yes Loss interval (days)=3	→	Immaterial damage=3	1,2%	73,7%
5*	Immaterial damage=2 Judge=V.P. Type of judge=Chief Flight delay=Yes	→	Verdict=Partly founded Delay interval=2	1,8%	75%
6	Judge=V.P. Type of judge=Chief Consumer fault=Yes	→	Verdict=Not founded Immaterial damage=1	1%	92,3%
7	Verdict=Not founded Flight delay=Yes Delay interval=1	→	Immaterial damage=1	1,2%	100%

3.5.4 [C] Classification to predict the categorical judgement result

We classified the judgements into four labels that correspond to the four possible categorical judgement results (*well founded*, *not founded*, *partly founded* and *dismissed without prejudice*). We also classified the judgements into two labels, *consumer wins* (which is a merger of the classes *well founded* and *partly founded*) and *consumer loses* (which is a merger of the classes *not founded* and *dismissed without prejudice*). That way, we removed the final part of the document that corresponds to the judgement result, since the class is indicated there textually.

In this experiment[10] we worked with a dataset of 849 documents. At the end of preprocessing step, the dataset resulted in a corpus of 13,993 types of tokens, with a total of 585,268 tokens. We observe that some judgements contain more than one result (more than one labels). We duplicated the judgement in these cases. Thus, in the 4-class classification, the dataset resulted in 860 documents for training and is divided into: (1) 222 documents for the *well founded* label; (2) 89 documents for the *not founded* label; (3) 537 documents for the *partially founded* label; (4) 12 documents for the *dismissed without prejudice* label. In the 2-class classification, the dataset resulted in 857 documents for training and is divided into: (1) 98 documents for the *consumer loses* label and (2) 759

[10] The results discussed below are adapted from the paper Sabo *et al.* (2019).

documents for the *consumer wins* label.

By applying the classification algorithms (kNN, LR, NB, NN, RF and SVM) and using only TF as the count in the text representation step, we evaluated the results based on the accuracy, precision, recall and F1-score. As a sampling method for creating the training and test sets, we used the cross-validation. Table 15 presents the parameters we used in each technique. Since this experiment was run only on Orange 3, we adopted the parameters recommended by the tool.

Table 15 – Classification setup parameters

Tool	Technique	Parameters
Orange	kNN	Number of Neighbors: 4; Distance Metric: Euclidean; Weight: Uniform
	LR	Regularisation type: Ridge (L2); C (strength): 1
	NB	–
	NN	Hidden Layers: 2 Neurons in each layer: 100, 50; Activation Function (Hidden): tanh; Solver: Stochastic Gradient Descent (SGD) Iteration Limit: 1000 Early Stopping: Deactivated.
	RF	Number of Trees: 10; Minimum subset size: 5. Max Depth: Unlimited. Max Leaf Nodes: Unlimited.
	SVM	C (cost): 1.0; ε (Regression loss): 0.1; Kernel: RBF; Iteration Limit: 100
	Cross-validation	Number of Folds: 10; Sampling type: Stratified.

We emphasise that to perform these experiments we manually labelled the verdict (categorical judgement result) to serve as an adjacent input to the dataset.

Results from 4-class classification and 2-class classification

Table 16 and Figure 24 present the results obtained by each classification model in the “4-class classification”. These numbers represent an average over the classes.

Table 16 – Results from 4-class classification

Model	Accuracy	F1-Score	Precision	Recall
LR	**0.777**	**0.770**	0.768	**0.777**
NN	0.775	0.768	0.766	0.775
RF	0.756	0.736	0.745	0.756
kNN	0.700	0.662	0.690	0.700
SVM-RBF	0.556	0.542	0.616	0.556
NB	0.144	0.212	**0.821**	0.144

Figure 24 – Results from 4-class classification

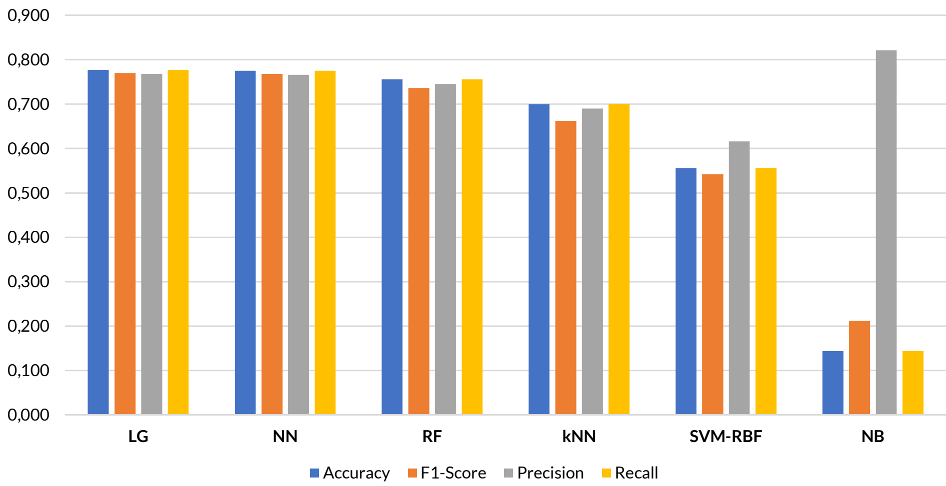
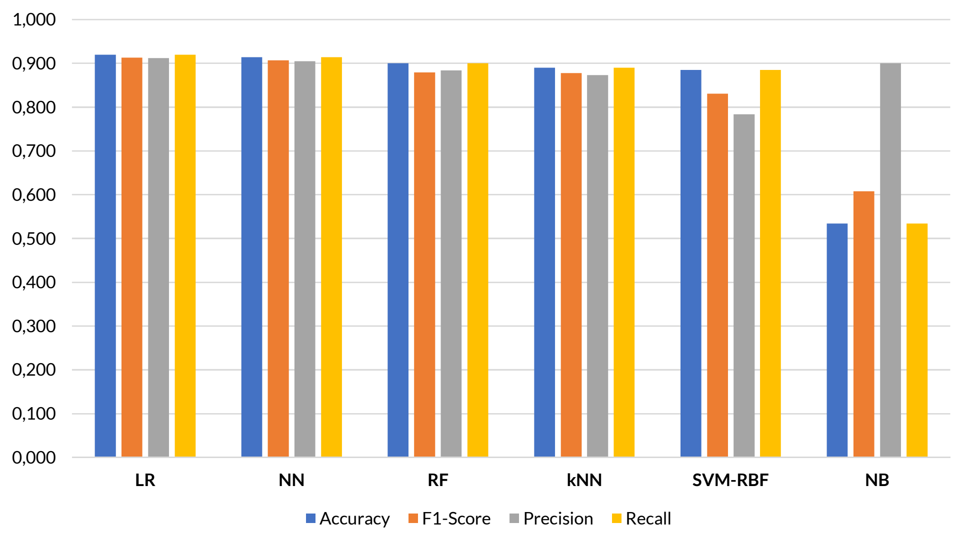


Table 17 and Figure 25, in turn, show the results obtained by each classification model in the "2-class classification". These numbers represent an average over the classes.

Table 17 – Results from 2-class classification

Model	Accuracy	F1-Score	Precision	Recall
LR	**0.920**	**0.913**	**0.912**	**0.920**
NN	0.914	0.907	0.905	0.914
RF	0.900	0.879	0.884	0.900
kNN	0.890	0.878	0.873	0.890
SVM-RBF	0.885	0.831	0.784	0.885
NB	0.534	0.608	0.900	0.534

Figure 25 – Results from 2-class classification



From Table 16 and Figure 24 to Table 17 and Figure 25, we can note significant improvements on the four metrics for the classification models. In terms of techniques, we can state in both situations that LR, followed by NN and RF, achieved the best results, while the NB presented poor results. In terms of strategy, we can affirm that the most appropriate is the 2-class classification.

This differences between the two strategies occurs because it is easier for models to learn with fewer classes. Moreover, we are working with small subsets (in the 4-class classification, *not founded* (89 examples) and *dismissed without prejudice* (12 examples); in the 2-class classification, *consumer loses* (98 examples)), then the chances of false positives and false negatives (the number of wrong predictions) are higher.

To better understand this situation, it is necessary to verify the number of predictions the model got right and the number it got wrong. The confusion matrices represented in Tables 18 and 19 give the number/proportion of instances between the predicted and real class by the best and the worst model (LR and NB) in the 4-class classification, while Tables 20 and 21 give the same in the 2-class classification.

Table 18 – Confusion matrix for Logistic Regression (4-class)

		Predicted				
		partly founded	*well founded*	*dismissed without prejudice*	*not founded*	Σ
Real	*partly founded*	**472**	53	0	12	537
	well founded	72	**147**	3	0	222
	dismissed without prejudice	5	5	**0**	2	12
	not founded	29	8	2	**50**	89
	Σ	578	213	5	64	860

Table 19 – Confusion matrix for Naive Bayes (4-class)

		Predicted				
		partly founded	*well founded*	*dismissed without prejudice*	*not founded*	Σ
Real	*partly founded*	**38**	14	464	21	537
	well founded	3	**60**	157	2	222
	dismissed without prejudice	0	1	**9**	2	12
	not founded	0	2	70	**17**	89
	Σ	41	77	700	42	860

In the 4-class, we can see that the best model (LR - Table 18) got a proportional number of correct predictions for each label in relation to the total number of examples in each subset/class (472 predicted examples of 537 real examples, 147 of 222, 0 of 12 and 50 of 89). However, as said, it did not get any correct predictions for the smallest subset (*dismissed without prejudice*). On the other hand, the worst model (NB - Table 19) got few correct predictions in general, but at least got 9 correct predictions of 12 for the smallest subset (*dismissed without prejudice*).

Table 20 – Confusion matrix for Logistic Regression (2-class)

		Predicted		
		consumer loses	*consumer wins*	Σ
Real	*consumer loses*	**48**	50	98
	consumer wins	18	**741**	759
	Σ	66	791	857

Table 21 – Confusion matrix for Naive Bayes (2-class)

		Predicted		
		consumer loses	*consumer wins*	Σ
Real	*consumer loses*	**95**	3	98
	consumer wins	396	**363**	759
	Σ	491	366	857

When we change to 2-class classification, we see that although the best model (LR - Table 20) had a satisfactory accuracy (92%), it got only half correct predictions of the smallest subset. The worst model (NB - Table 21), on the opposite, because it got a significant number of wrong predictions in the largest subset, also got 95 predicted examples of 98 real examples in the smallest subset.

We need to pay attention to this kind of situation when we work with smaller subsets, in particular when we adopt the binary classification. Even if the model got all the examples of the label *consumer loses* wrong, the overall accuracy would not be low. This is what happened, for example, with SVM-RBF, which obtained an accuracy of 88.5% but did not memorise any examples of the smaller subset, as shown in the confusion matrix in Table 22.

Table 22 – Confusion matrix for SVM-RBF (2-class)

		Predicted		
		consumer loses	*consumer wins*	Σ
Real	*consumer loses*	**0**	98	98
	consumer wins	0	**759**	759
	Σ		857	857

Considering the LR performance and our experience in the legal environment, this model can be sufficient for a real application, whereas predicting whether the consumer will win compensation or not (2-class) is quite useful. Furthermore, from the Consumer Law principles (especially the consumer's vulnerability), we understand that is not unfair a situation of more false positives in relation to false negatives, since our model tends to predict that the consumer will win the lawsuit. Finally, these classification results are on the same level as the accuracies presented in the related work.

3.5.5 [D] Regression to predict the numerical judgement result

Considering classification results, once we can predict if the consumer will win the lawsuit (categorical judgement result), the next step is to predict how much compensation will be arbitrated by the judge in his or her favour for the immaterial damage suffered (numerical judgement result). For this, we used only the judgements in which there is the compensation, i.e., those that are well founded and partial founded. As in the classification, we also removed the final part of the text from the document, since the amount of compensation is also specified there.

In this experiment we worked with a dataset of 928 documents. At the end of preprocessing step, the dataset resulted in a corpus of 15,258 types of tokens, with a total of 715,269 tokens.

By applying the regression algorithms (SVM, RF, NN, Bagging, GBoost, Adaboost and XGBoost) and using only TF as the count in the text representation step, we evaluated the results based on the RMSE, MAE and R^2, what we call "baseline". Table 23 presents the parameters we used in each technique for the "baseline", noting that in this experiment we adopted a complex setup. We adopted the parameters recommended by the Scikit-Learn Library.

Table 23 – Regression setup parameters (baseline)

Tool	Technique	Parameters (complex)
Python Programming Language and Scikit-Learn Library	Adaboost	No. Trees: 100.
	Bagging	No. Trees: 100.
	NN	Hidden Layers: 5; Neurons in each layer: 512; Activation Function (Hidden): ReLU; Solver: ADAM; Iteration Limit: 100; Early stopping: Deactivated.
	RF	No. Trees: 100. Minimum subset size to split: 2. Max Depth: Unlimited. Max Leaf Nodes: Unlimited.
	SVM	C (cost): 1.0; ε (Regression loss): 0.1; Kernel: RBF; Iteration Limit: 100.
	GBoost	No. Trees: 100. Max Depth: Unlimited. Max Leaf Nodes: Unlimited.
	XGBoost	No. Trees: 100. Max Depth: Unlimited.
	Ensemble Voting	Bagging; Neural Network; GBoost; XGBoost.

Furthermore, we evaluated the application of some adjustments (some ML and NLP techniques to improve the results, and also the addition of the attributes/factors extracted), what we denominated "full pipeline". We explain these adjustments ahead:

1. *N-grams* (preprocessing step): Varying from 1 to 4.

2. *Feature selection* (representation step): Using the Mutual Information method, it maps the relationship between each feature (unit in the BOW) and the dependent variable (the amount of immaterial damage compensation) (COVER; THOMAS, 2006).

3. *Overfitting avoidance* (regression step): Overfitting occurs when the model is too specialised in the train data and it achieve a poor prediction quality when evaluated in the test set (KARYSTINOS; PADOS, 2000). A possible adjustment to reduce overfitting is to reduce the complexity of the models, that is, check whether simpler models perform as well as the complex ones (LIU, R.; GILLIES, 2016). Thus, Table 24 presents the parameters we used in each technique for the "full pipeline", noting that they were simplified. In this case we manipulated the previous parameters.

4. *Cross-validation* (training step): As explained before, cross-validation uses multiple combinations of the train and test sets and the resulting metrics will be averaged. In this work, we set the number of folds to 5.

5. *Removal of outliers* (training step): Outliers are very distinctive examples in the dataset, and by removing them, we make it easier for the models to learn. To detect outliers, we used the Isolation Forest (LIU, F. T. *et al.*, 2008) with contamination set to 10% per cent. The former intends to remove outliers from the whole dataset, while the latter, from the train set. However, by removing outliers from all dataset, we imply that our future cases for prediction will not contain outliers.

6. *Addition of attributes/judgement factors*: Considering the attributes/dependent variables/judgement factors extracted in subsection 3.5.3 (from 1 to 14), we proposed them as a special adjustment. Here we do not intend to predict them, but only evaluate the gains when we add them.

Table 24 – Regression setup parameters (full pipeline)

Tool	Technique	Parameters (full pipeline/simple)
Python Programming Language and Scikit-Learn Library	Adaboost	No. Trees: 50.
	Bagging	No. Trees: 50.
	NN	Hidden Layers: 5; Neurons in each layer: 256; Activation Function (Hidden): ReLU; Solver: ADAM; Iteration Limit: 50; Early stopping: Activated.
	RF	No. Trees: 50. Minimum subset size: 2 Max Depth: 10. Max Leaf Nodes: 100.
	SVM	C (cost): 1.0; ε (Regression loss): 0.1; Kernel: RBF; Iteration Limit: 50.
	GBoost	No. Trees: 50 Max Depth: 10. Max Leaf Nodes: 100.
	XGBoost	No. Trees: 50. Max Depth: 10.
	Ensemble Voting	Bagging; Neural Network; GBoost; XGBoost.

We emphasise that to perform these experiments, besides the manually extraction of attributes/factors, we also manually labelled the amount of the immaterial damage compensation (numerical judgement result).

Results from baseline and full pipelines

Table 25 and Figure 26 present the results obtained by each regression model in the “baseline”, which setup does not include the adjustments.

Table 25 – Results from regression baseline

Model	RMSE	MAE	R^2
Ensemble Voting	**2,784**	**2,006**	**0.36**
GBoost	2,902	2,122	0.30
Bagging	2,908	1,997	0.30
RF	2,914	2,004	0.30
XGBoost	2,916	2,059	0.30
NN	3,031	2,273	0.24
Adaboost	3,276	2,596	0.11
SVM-RBF	3,553	2,750	-0.04

Figure 26 – Results from regression baseline

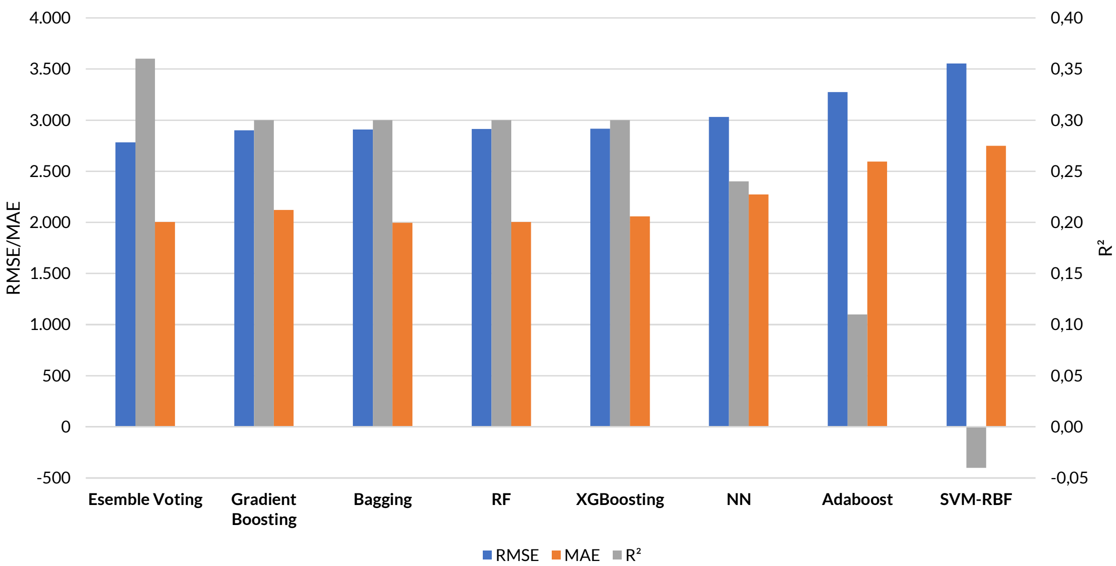
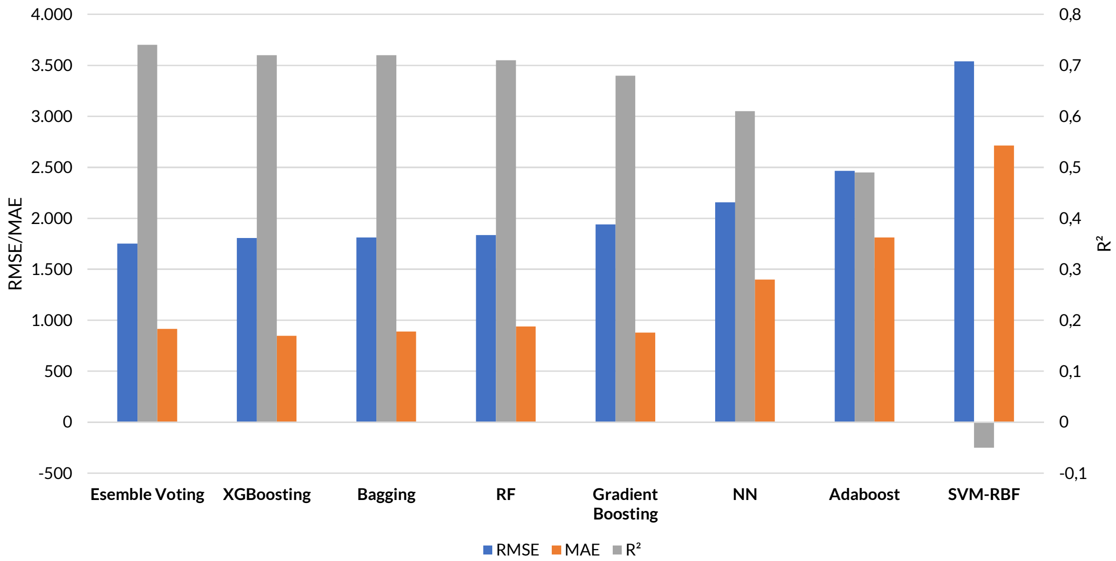


Table 26 and Figure 27, in turn, show the results obtained by each regression model in the "full pipeline", which setup includes all the additions (except the outliers removal in training data).

Table 26 – Results from regression full pipeline

Model	RMSE	MAE	R^2
Ensemble Voting	**1,750**	915	**0.74**
XGBoost	1,807	**847**	0.72
Bagging	1,812	890	0.72
RF	1,837	939	0.71
GBoost	1,940	879	0.68
NN	2,157	1,399	0.61
Adaboost	2,464	1,813	0.49
SVM-RBF	3,540	2,714	-0.05

Figure 27 – Results from regression full pipeline



From Table 25 and Figure 26 to Table 26 and Figure 27, we can note significant improvements on the three metrics for most of the regression models, except for SVM with RBF Kernel. In that case, we can affirm that SVM is underfitted, since the poor results stood regardless the pipelines we would apply. On the other hand, in terms of techniques, we can realise that Ensemble Voting achieved the best results among the others according the RMSE and R^2. XGBoost achieved the best prediction quality in terms of MAE. As described before, RMSE tends to penalise bigger errors, while MAE does not, so we can state that Ensemble Voting has fewer large errors than XGBoost. Still, it predicts incorrectly more examples than XGBoost.

As expected, we can affirm that the "full pipeline" leads to better results than "baseline", which also means that the extraction of attributes helped the performance of the techniques. Moreover, from our experience in the legal environment, a MAE of less than 1.000 can be considered irrelevant. When the parties are negotiating the amount of immaterial damage compensation, it is possible and reasonable for one of them to make a concession of approximately 1.000,00 Brazilian Reais.

3.5.6 Considerations on the section

Regarding our clustering experiments, we verified that Hierarchical Clustering performed best, both in terms of metrics and legal analysis, although the expert has to examine all clusters and assign a description to them. On the other hand, the Lingo algorithm, which has the ability to generate descriptions, allowed us to identify situations that, in theory, should not be a variable/factor in the judgement. Can a certain airline as defendant influence the judge? This was a question that Lingo made us think about. Moreover, the activity of an algorithm in generating descriptions is something we consider relevant for the legal text analysis, because it helps the expert to perceive unseen weighty law events, minimising biases in the information extraction.

Likewise, considering the complexity of legal text data, which may have more than one descriptions (as can be seen in the Appendices), soft clustering methods are more advantageous and useful in identifying legal variables. Since this approach allows manipulation of the number of clusters, the expert can run different attempts until a cluster size in which it is possible to detect a common law event. In this perspective, we can state that these techniques are still limited in explaining a law event. An clustering algorithm output must be critically analysed by a legal professional before implementing or applying anything.

About the association experiments, once we organise the descriptions into factors and judgement results and extract rules that relate them, we can model the judge's reasoning (or at least some part of it). Knowing what the judge takes into consideration when deciding and presenting comprehensibly it to the parties is a way to explain possible predictions, which can be highly useful in an ODR system. In other words, the symbolic approach remains significant in this domain.

Concerning our classification experiments, we found that classical learning models can obtain satisfactory results for legal texts (such LR and non-complex NN). Nevertheless, we emphasise the importance of the NNs when applied to legal texts. As shown in the related work, successful results are reached by using even more complex networks (such as recurrent and convolutional). With respect to the number of classes, when we are dealing with non-large datasets and disproportionate splits, adopting a binary classification strategy (i.e., addressing the problem in a simplified way) can improve the performance of the models, enabling applications in the legal environment.

In relation to the regression experiments, although Ensemble Voting showed the best overall performance, we note that XGBooting (an optimised version of the tree-based models that has also been successful) achieved the lowest MAE (847.00), which is the simplest measure of error. Ultimately, we believe that our results can encourage the parties involved (consumer and airline) to an agreement. For example, the consumer who will earn R$ 5.000,00 only at the end of the lawsuit, will agree more easily to being compensated in R$ 4.000,00 in the beginning, so the case is closed immediately. By obtaining more agreements, we also generate a positive impact on the Justice response time.

Lastly, Table 27 summarises all ML tasks, stages and techniques employed in the case study.

Table 27 – ML tasks vs. ML stages and techniques (overview)

ML stages and techniques

ML tasks		Preprocessing	Representation	Learning and Training	Evaluation
	Clustering	Normalisation Tokenisation Stemming Filtering 2-grams	Bag of Words (TF and TF-IDF)	Hierarchical clustering K-means Lingo Affinity propagation	Clustering tendency Silhouette score Entropy Purity
	Association	-	-	FP-Growth	Support Confidence
	Classification	Normalisation Tokenisation Stemming Filtering 2-grams	Bag of Words (TF)	SVM-RBF kNN Logistic Regression Neural Networks Random Forest Naive Bayes Cross-validation	Accuracy F1-score Precision Recall
	Regression	Normalisation Tokenisation Stemming Filtering 4-grams	Bag of Words (TF) Mutual information	SVM-RBF Neural Networks Random Forest Bagging Adaboost GBoost XGBoost Ensemble voting Cross-validation	RMSE MAE R^2

4 PROPOSAL AND VALIDATION

4.1 THE MACHINE LEARNING-BASED MODEL PROPOSED

Given the satisfactory results obtained in the experiments, our thesis proposal is a ML-based model to predict the judgement result and, with this, to assist the conciliation hearing in the Special Civil Courts. In other words, we envision that the knowledge generated from the data by applying the ML techniques can be provided to the parties to help them reach an agreement. Our proposed model includes everything from data collection to application in real cases. For this reason, we prefer to call this an ML-based model, rather than an ML model, because we understand that our model is more extensive, whereas an ML model is about fitting a set of ML techniques to a given dataset. Figure 28 presents the proposed ML-based model. Therefore, as a result of the case study, we consolidated the model into four stages:

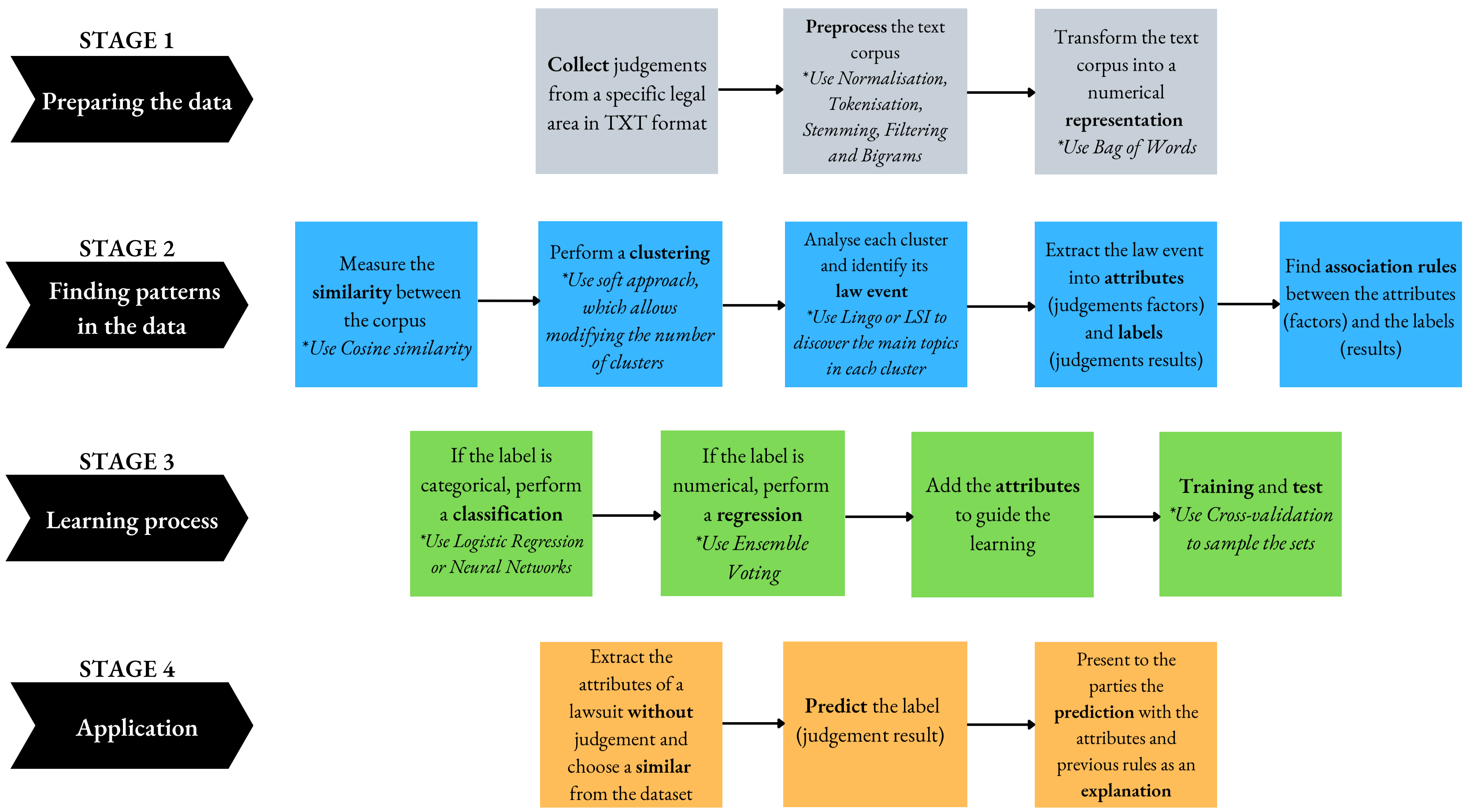
- **Stage 1 - Preparing the data**: This first and important stage consists of collecting court decisions by importing them in TXT format, which is acceptable by most tools for implementing ML techniques. This data will be cleaned by applying basic NLP techniques, such as *normalisation*, *tokenisation* and *stemming*, and also by filtering the *stopwords* and uniting *bigrams*. Afterwards, we have to transform this data into a numerical representation through the *BOW*, which is a technique that consists simply on counting the frequency of terms. Thus, the data is ready to serve as input for the application of supervised and unsupervised ML techniques.

- **Stage 2 - Finding patterns in the data**: This is a stage that we consider essential for further structuring the data. The use of unsupervised techniques will serve to verify patterns in the documents when we do not know about the text subject and, whether we know, it will avoid bias in the information extraction. For the clustering application, we first determine the similarity or dissimilarity between the documents using the *cosine* and subsequently group them. As a clustering technique, the *soft approaches* (Hierarchical Clustering and Lingo algorithms) is advantageous in the case of complex documents (such as legal documents), because a document can belong to more than one group, or it can belong to a large group and a subgroup at the same time. In our context, for example, a document can belong to the group “substitute judges” and at the same time from the group “flight delay”. Also, a document may belong to the general category “baggage loss” and the subcategory “temporary baggage loss”. To identify this common law event in the clusters and extract descriptions, the clustering performed by Lingo (which has the ability to already provide clusters descriptions) or topic modelling techniques such as LSI facilitate the analysis. Then, these extracted descriptions should be organised into judgement factors and results (attributes and labels) in XLS or CSV format, and it is also possible to extract association rules (antecedent → consequent) if there is a strong relationship between them.

- **Stage 3 - Learning process**: Once we have the text corpus and also structured data about it, we proceed to the stage of supervised learning. If the extracted possible results are categorical, then we should apply classification techniques (*NN* and *LR* proved to be more effective); if we are facing numerical results, then regression techniques are appropriate (*Ensemble Voting* based on other techniques - *NN*, *Bagging*, *G Boosting* and *XG Boosting*). In addition to the text corpus itself, it is important to include as input the extracted attributes/factors, because these are extra information about the documents, which will enhance the learning process. To separate the train and the test sets, we understand that the *cross-validation* technique is the most suitable, because it prevents the classification and regression algorithms from memorising only the training documents (overfitting), improving the results at test time. In both regression and classification, we do not need parameters that make the models complex and require high computational performance.

- **Stage 4 - Application**: Lastly, to enable a real application of the model in a legal environment, i.e., a judgement result prediction (the verdict and the amount of compensation for immaterial damage) in cases in which there is still no judgement, we should follow this strategy: (a) select lawsuits that were awaiting judgement (*new cases*), extract the same attributes/factors from each of them and search in the dataset for similar situations (*past cases*). That way, the attributes/factors of the new case and the document related to the past case will serve as input to the model; (b) predict the label (categorical and numerical) of each new case. In the case of the numerical value, it is appropriate to predict it in different rounds, calculating a minimum, a maximum and an average, and compare it with the value of the past case; (c) present the predictions to the parties by using the attributes/factors and the association rules extracted, since they can serve as an explanation to the parties and their lawyers.

Figure 28 – ML-based model proposed



The model can be replicated and tested in other legal domains. We believe that it is essential having both a legal expert and a professional with knowledge of Python to perform the regression task. However, for the other tasks (preprocessing, clustering, association, and classification), the software used (Orange 3 and Carrot[2]) have helpful components, didactic interface, and tutorials that allow the legal expert to analyse the data and understand the techniques without needing advanced technical skills. Moreover, the techniques that compose the model are not complex and do not require high computational cost.

4.2 VALIDATIONS, RESULTS AND DISCUSSION

In this final section we explain how we made a real application of the model, allowing us to validate the thesis proposal. It is necessary to emphasise that this step called ‘validation’ should not be confused with ‘evaluation’ (Subsection 3.4.8), which correspond to the metrics defined in the literature to evaluate the quality of the results (output) after the application of ML techniques on the corpus. By ‘validation’ we mean verify whether the model is consistent and able to comply with the general objective of the thesis (demonstrate that a ML-based model can be useful for the parties in the conciliation hearing). We also can name this step as an ‘empirical verification’ of the model, since this will occur in a legal environment. We validate it in two steps:

(1) First, as opposed to the beginning of the case study (non-participant observation), we participated in conciliation hearings at the JEC/UFSC, applying the model and presenting its predicted results to the parties involved. We also invited them (parties and their lawyers) to answer an anonymous questionnaire evaluating the usefulness of the information provided.

(2) Second, after a few months, we consulted the lawsuits submitted to the conciliation hearings and verified the real results, comparing them to the ones predicted by the model.

4.2.1 Participant observation of hearings and survey to the parties involved

To enable our participant observation, the JEC/UFSC staff selected 13 lawsuits that were awaiting judgement and the chief judge assigned a conciliation hearing for each of them, giving us the role of conciliator. The official documents relating to this assignment and our participant observation of the conciliation hearings are listed in Annex B. We also sent an invitation letter to the parties, and in the case of the airlines (defendants), we asked them to send agreement offers to the hearing to start a conversation. This is the first part our ML-based model validation, which aims to examine the parties reaction when putting them in contact with judgement predictions about their cases.

We observed them in the period between 13/09/2021 to 16/09/2021. Due to the Covid-19 pandemic, all conciliation hearings were held virtually. To introduce to the parties the predicted values, we prepared simple language material about the developed project (named *Intelligent Conciliation* - Appendix F), in which we explained:

1. *Purpose*: To provide data-based information to the parties in order to assist them in the negotiation.
2. *Advantages of the agreement*: Informality, celerity, simplicity, among others.
3. *Judgement influencing factors*: Flight delay and cancellation, waiting time at the airport (which corresponds to the extracted judgement factors/rules). With this, we are able to provide some explanation to the parties regarding the predicted outcome.
4. *Overview of the amounts fixed as compensation for immaterial damage*: Which amounts are most incident, to demonstrate that large compensation values are exceptions in the JEC/UFSC.
5. *High compensation cases*: Which are the circumstances and factors related to those past cases, so the parties can compare them to their case.
6. *Individual predicted value*: Which are the predicted values for the new case discussed and its specific judgement factors.

After the conciliation hearing we send a voluntary and semi-structured survey to everyone involved in the hearings (plaintiffs, defendants, and lawyers) in order to evaluate the impact of the information provided. The survey includes multiple choice (MCQ) and short answer (SAQ) questions (closed-ended), and a final open-ended question. We used Likert scale to better measure the answers to some of the questions.

1. What is your position in the case submitted for the conciliation hearing? (SAQ)
   a) Plaintiff
   b) Defendant
   c) Plaintiff's lawyer
   d) Defendant's lawyer
2. Did you reach an agreement during the conciliation hearing? (SAQ)
   a) Yes
   b) No
3. If the previous answer was "no", what reason(s) would you attribute to the failure of the conciliation? (MCQ)

a) Absence of one or both parties
b) No plaintiff's interest
c) No defendant's offer
d) The defendant's offer did not meet the plaintiff's expectation

4. If the previous answer was "yes" or "partial", what reason(s) would you attribute to the success of the conciliation? (MCQ)

   a) Parties' interest
   b) Conciliator's data information
   c) Defendant's offer
   d) The parties waived part of the value

5. Was the information presented by the conciliator about local judgements helpful to you? (SAQ)

   a) Extremely helpful
   b) Very helpful
   c) Somewhat helpful
   d) Slightly helpful
   e) Not at all helpful

6. Were the values predicted of compensation for immaterial damage helpful to you? (SAQ)

   a) Extremely helpful
   b) Very helpful
   c) Somewhat helpful
   d) Slightly helpful
   e) Not at all helpful

7. Could you please measure your level of satisfaction with the activity provided? (SAQ)

   a) Very satisfied
   b) Somewhat satisfied
   c) Neither satisfied nor dissatisfied
   d) Somewhat dissatisfied
   e) Very dissatisfied

8. Would you like to leave a comment, critic or praise? (open-ended)

The completed form is in Appendix G (in Brazilian Portuguese language). We do not collect personal data in the responses.

We expose and organise the responses from the survey in the charts below, followed by our analysis, emphasising that the responses were individual, anonymous and voluntary.

Figure 29 – (Participant observation) Q1. What is your position in the case submitted for the conciliation hearing?

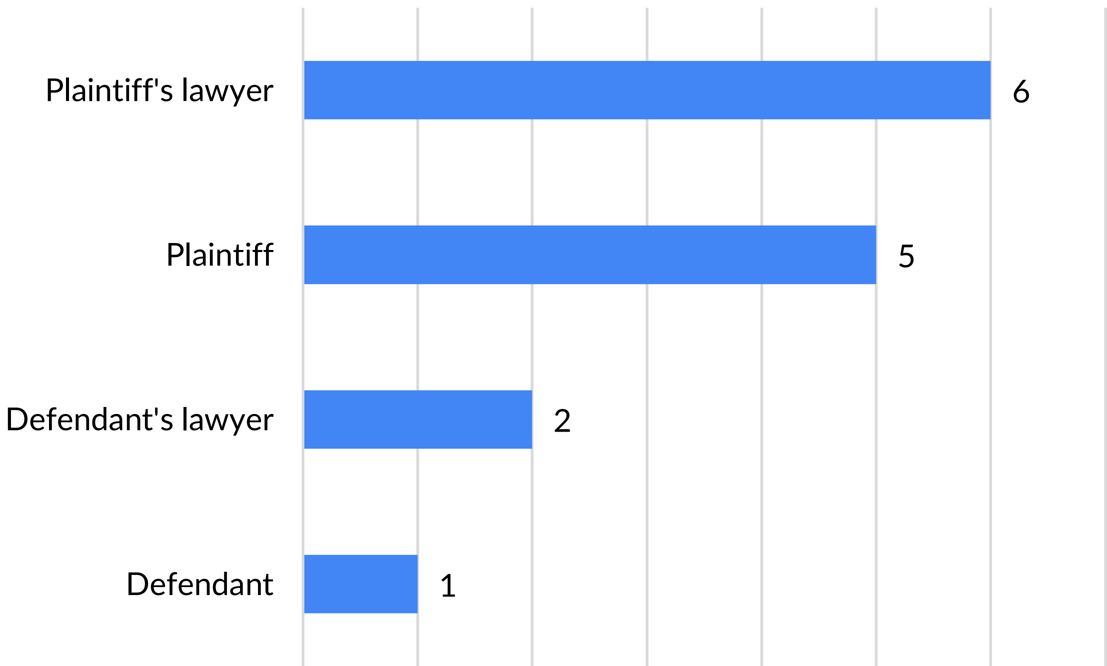


Figure 29 shows that we got more responses from plaintiffs and their lawyers, which may imply that those involved are more interested in obtaining the available information about the immaterial damage compensation. Of the 14 conciliation hearings held, we reached an agreement in 1 of them, and for this reason the greater amount of “no” in the answers, as shown in the Figure 30.

Figure 30 – (Participant observation) Q2. Did you reach an agreement during the conciliation hearing?

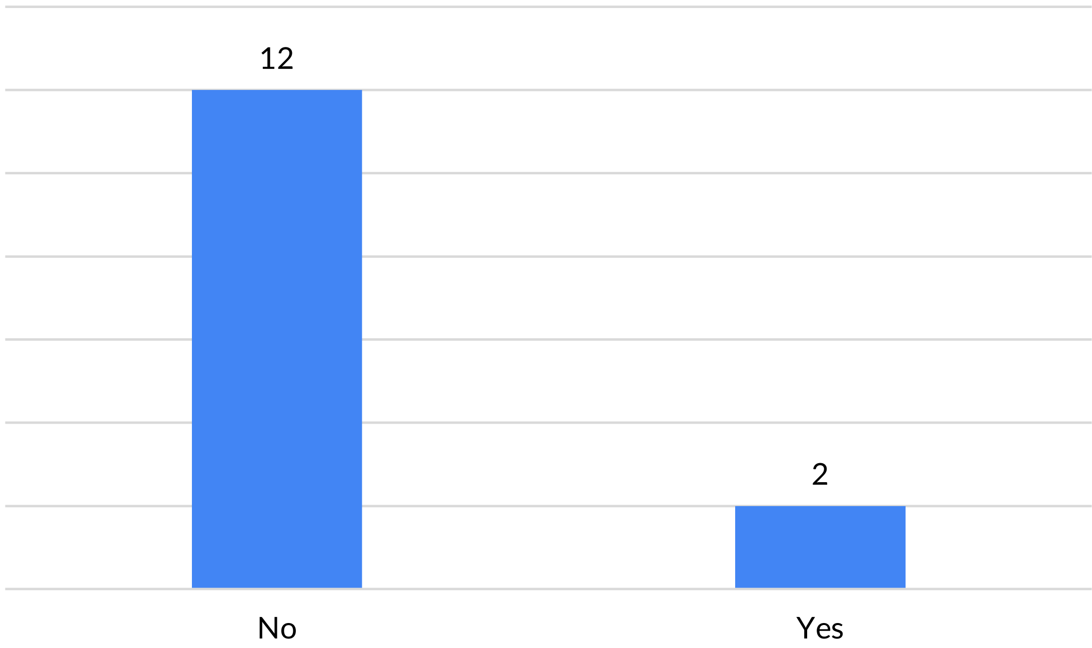
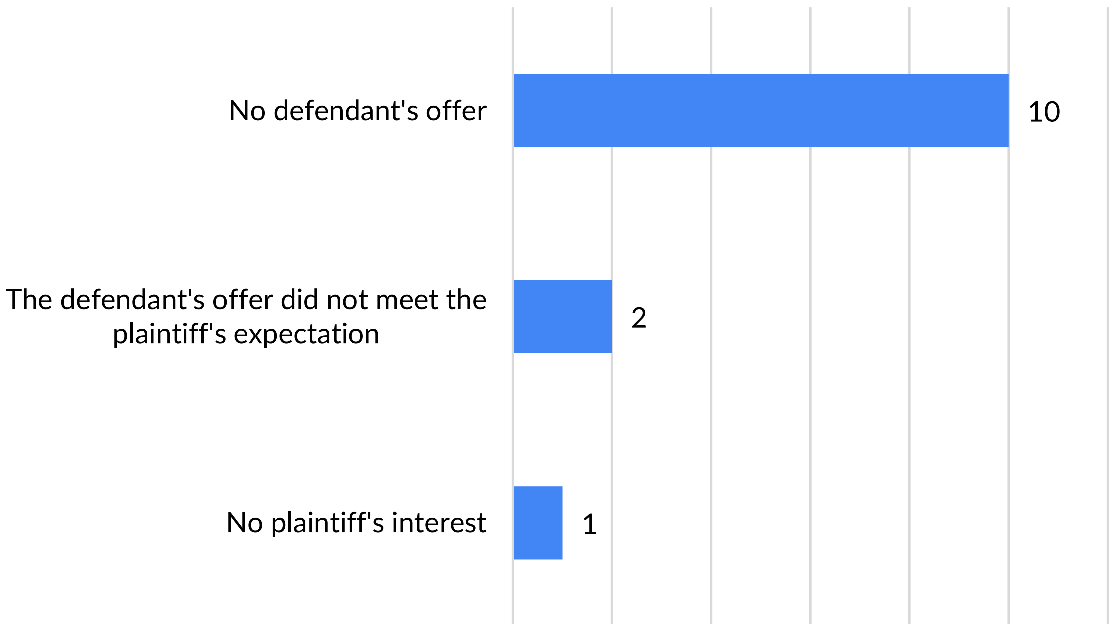
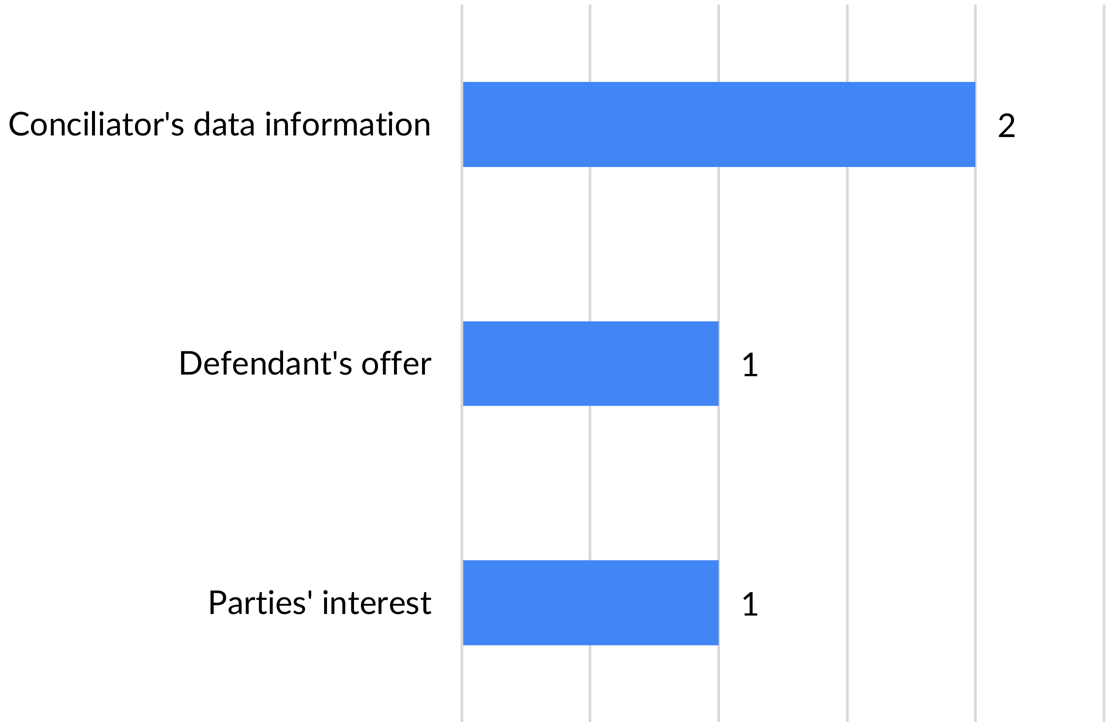


Figure 31 – (Participant observation) Q3. If the previous answer was "no", what reason(s) would you attribute to the failure of the conciliation?



We perceived it is essential the sending of offers by the airlines to start at least a negotiation between the parties, even without an agreement at the final. As the Figure 31 indicates, the main reason why we have not reached agreements is that the defendants attend the hearings without proposals. This is an situation that limit the performance of our research. In a future round of hearings, we will consider other external means to encourage airlines to submit offers, such as sending prior information to the airlines' legal departments.

Figure 32 – (Participant observation) Q4. If the previous answer was "yes" or "partial", what reason(s) would you attribute to the success of the conciliation?



On the other hand, as the Figure 32 indicates, at the hearing with agreement the parties credited the success to the information we provided, to their interest in negotiating and also to the fact that the airline made an initial offer. According to our impression, in this specific case, we observed that the predicted value made the defendant improve its initial offer and, as a consequence, the second offer was closer to the plaintiff's expectation, resulting in the settlement.

Figure 33 – (Participant observation) Q5. Was the information presented by the conciliator about local judgements useful to you?

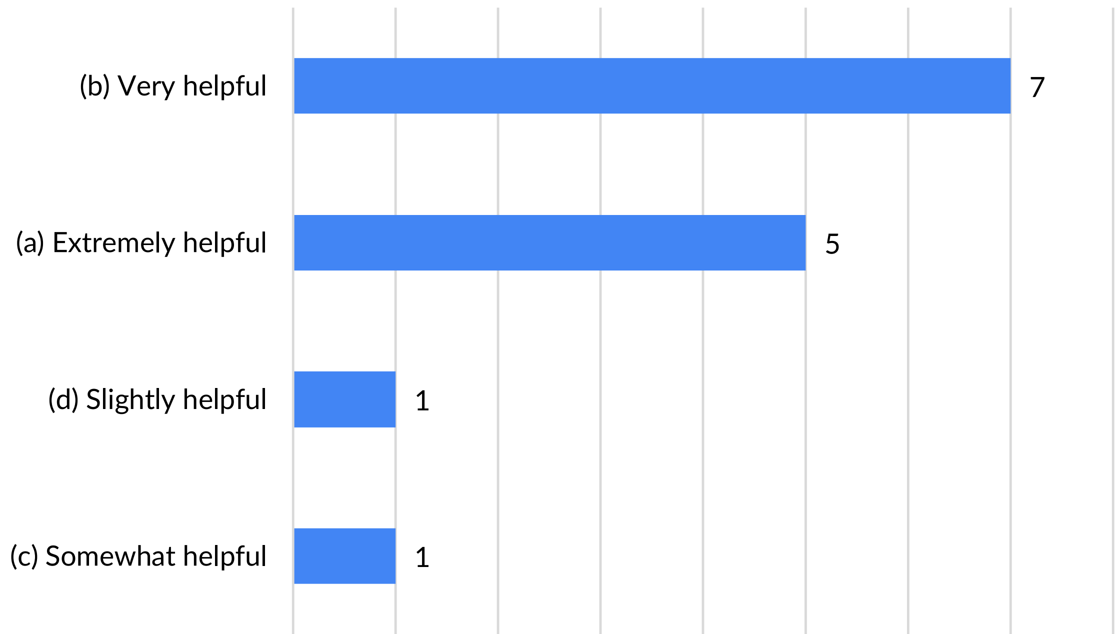
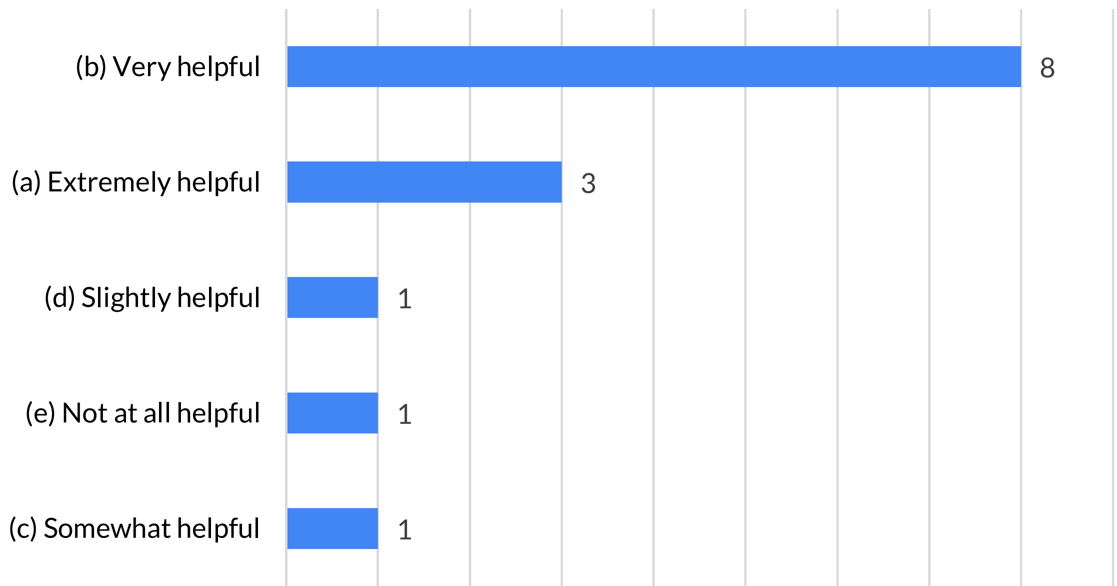


Figure 34 – (Participant observation) Q6. Were the values predicted of compensation for immaterial damage helpful to you?



In most of the hearings, we observed that the information presented, including the predicted amounts of compensation for immaterial damages, was well received and appreciated by the parties involved and their lawyers. However, we observed that some authors and lawyers were disappointed with the predicted value, because they imagined it would be higher. This is why providing factors and rules related to the predicted outcome is important not only to inform the parties about them, but especially for their acceptance.

Figure 35 – (Participant observation) Q7. Could you please measure your overall level of satisfaction with the activity provided?

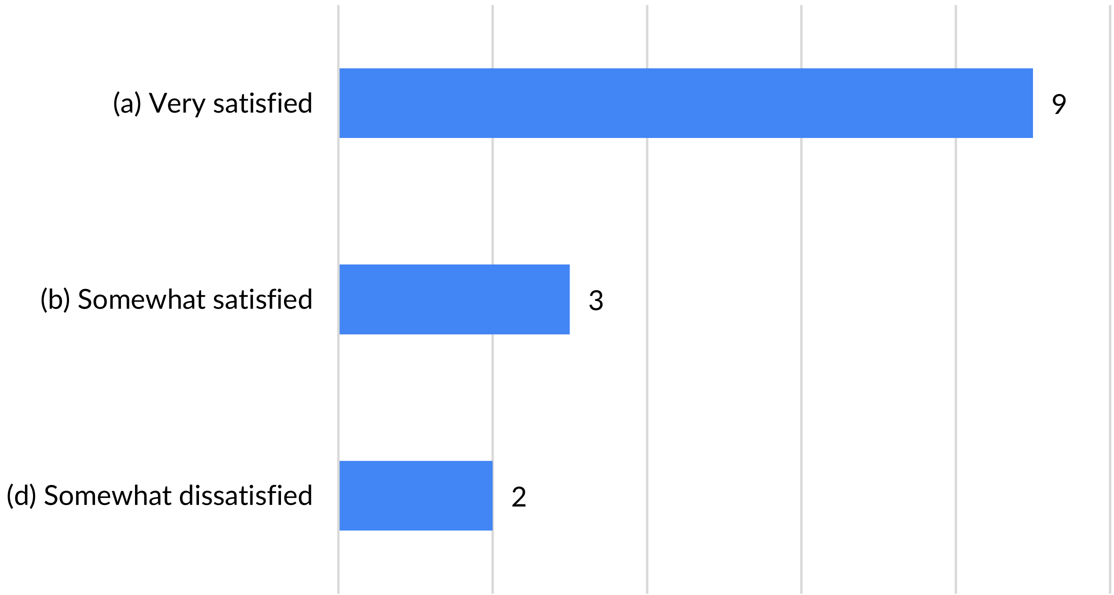


Table 28 – (Participant observation) Q8. Would you like to leave a comment, critic or praise?

	Number of responses	Comments
Praise	5	Congratulations on the initiative; This kind of information provides subsidies for the parties to decide.
Critic	3	The predicted amounts were low in relation to the circumstances of the case; The process could be held up for a long time awaiting the hearing.
Suggestion	1	Previously check with the airline if they will offer an initial value, otherwise do not hold the hearing.

As seen in Figure 35 and Table 28, we can state that our thesis proposal obtained satisfactory results according to our observation and the responses of the parties involved. We consider the criticism and suggestions valid and constructive, although we do not agree with the hypothesis of not holding the hearing, since conciliation is the main objective of the JECs. This means that when the party proposes a lawsuit in this way, it should be aware that it will have to try to negotiate.

In addition to the usefulness, we can state that the prediction of a lawsuit outcome is also fair when provided to the parties involved, who can make their own decisions (choosing to settle or wait for the judgement). Therefore, the proposed model was applied respecting the parties' autonomy and the judicial procedure.

4.2.2 Comparison between predicted and real judgement results of hearing cases

The participation in the hearings was conducted without the presence of the chief judge, so she did not have access to our predictions. Approximately nine months after our participant observation, we consulted these lawsuits to verify their judgement results and compare them with the predictions we have provided to the parties in the conciliation hearings. This is the second part of our ML-based model validation, which aims to check how close or not our predictions were to the actual judgement results.

Table 29 shows the predictions we reported to the parties and the actual results in the first-degree judgement. We remark that in the prediction of immaterial damage compensation (numerical judgement result), we ran the regression pipeline around 20 times and provided a range to the parties, which corresponds to the minimum and maximum obtained values.

Table 29 – Predicted judgement results x Real judgement results

Lawsuit	Predicted verdict	Predicted immaterial damage (range)	Real verdict	Real immaterial damage
1	Partly founded	R$ 4.000,00 - R$ 5.900,00	Not founded	R$ 0,00
2	Partly founded	R$ 2.100,00 - R$ 4.000,00	Partly founded	R$ 1.000,00
3	Partly founded	R$ 6.200,00 - R$ 7.800,00	*Pending judgement*	
4	Partly founded	R$ 2.400,00 - R$ 4.300,00	Partly founded	R$ 3.000,00
5	Partly founded	R$ 9.000,00 - R$ 10.100,00	*Agreement in conciliation hearing*	
6	Partly founded	R$ 2.200,00 - R$ 4.300,00	Partly founded	R$ 1.000,00
7	Partly founded	R$ 7.500,00 - R$ 8.900,00	Partly founded	R$ 6.000,00
8	Partly founded	R$ 4.900,00 - R$ 6.500,00	Partly founded	R$ 8.000,00
9	Partly founded	R$ 4.400,00 - R$ 5.500,00	Partly founded	R$ 5.000,00
10	Partly founded	R$ 4.400,00 - R$ 6.000,00	Partly founded	R$ 0,00
11	Partly founded	R$ 5.300,00 - R$ 8.700,00	Not founded	R$ 0,00
12	Partly founded	R$ 5.100,00 - R$ 5.600,00	Partly founded	R$ 8.000,00
13	Partly founded	R$ 4.200,00 - R$ 6.400,00	Partly founded	R$ 5.000,00

Some observations about these cases:

- **Lawsuits no. 4, 9 and 13**: In these cases the real value is exactly within the predicted range, i.e., the model predicted the correct value.
- **Lawsuits no. 2, 6, 7 and 8**: In these cases the real value is somewhat above or below the predicted range, i.e., the model had an average error of approximately R$ 1.325,00 in the predicted value. This error is close to the one obtained during the experiments.
- **Lawsuits no. 1 and 12**: In these cases the real value is distant from the predicted range, i.e., the model predicted the wrong value.

- **Lawsuits no. 10 and 11**: These two cases dealt with flight cancellations and delays at the beginning of the Covid-19 pandemic, when there was not yet a common position on whether this event would be considered an unforeseeable circumstances or a force majeure event, which excludes the possibility of compensation (similar to the extracted factor/attribute *adverse weather conditions*). Despite this, we decided to hold the hearings with the predicted values. As we emphasised, the agreement in this situation would be an autonomous decision by the parties and independent of the judge's subsequent evaluation. However, as shown, the judge did not recognise the immaterial damage occurrence in these two lawsuits.

In general, we understand that the model is susceptible to application. As seen, it tends to judge favourably the consumer (false positive), and this is preferable to the unfavourable judgement (false negative), since we are dealing with Brazilian Consumer Law, whereby the consumer is seen as the vulnerable party in the relationship.

5 CONCLUSION

5.1 FINAL REMARKS AND CONTRIBUTIONS

Doing this interdisciplinary work allowed us to view the advantages of legal research when combined with technology. Law has much to gain when the data produced by its operators is transformed into knowledge through instruments belonging to other knowledge areas, such as Computer Science and Engineering. From this perspective, the interaction with researchers from these areas and the group work were essential for our thesis proposal.

Part of the systemic and complex view is the recognition of Law (and the Judiciary) as an autopoietic organisation that adapts on its own to social needs. In the ideal world, there would be no conflict. Consequently, there would be no lawsuit. But once the real world is full of litigation, a possible structural change is using the caseload that disturbs the organisation as a tool to achieve non-litigation. From the litigiousness comes data, which can be processed, analysed and modelled, and then converted into a settlement option.

Concerning the data, we spend significant time and work understanding and structuring it with the least possible biases and variations, aiming not only for the outcome accuracy but also for the input quality. Efforts to give effectiveness to the judicial process by developing ML and NLP solutions should consider this step performed by Law professionals as essential. Hence, Law courses and institutions must start a tradition in empirical research.

Retrieving the research question, we can affirmatively conclude that it is possible to apply an ML-based model to predict judgement results in the Brazilian JECs. Furthermore, these predictions increase the probability of settlement, improve the conciliation hearings' quality, and are useful to the litigating parties. With this, we confirm the hypothesis and fulfil our general objective.

Therefore, we define the contributions of the thesis as both academic and practical. The academic contribution corresponds to the guide, the step-by-step a legal researcher should follow for helping in an ML-based solution development. This includes: (i) what kind of data and how to collect it (text data from court decisions, preferably in TXT format); (ii) how to extract information from data and organise it (judgement factors/attributes and judgement results/labels); and (iii) which techniques to apply to transform it into knowledge (using tools that can be understood by legal students - such as Orange 3 and Carrot[2]). Yet, (iv) how to apply this generated knowledge without violating the parties' autonomy and the legal procedure. The practical contribution refers to the products resulting from the research:

- A dataset both textual (TXT files) and structured (XLS file) specific to air transport service failures (Consumer Law) belonging to the JEC/UFSC, available at: https://bityli.com/XXkJT. The names of the plaintiff and defendant were removed.

- A ML-based model that provides accurate predictions about the JEC/UFSC judgement results. The model addresses four ML pipelines (clustering, association, classification and regression), whose files in software-processable format and Python codes are also included in the folder.

Although the proposed model can also be employed by judges, since the predictions generated concern their decisions, we understand that its use is more suitable in the ADR/ODR context, such as conciliation hearings or similar. The purpose is to eliminate the dispute at its origin, avoiding further procedural costs until the lawsuit's judgement. In addition, as we have highlighted, we aim to encourage a culture of pacification, providing subsidies for the parties to be able to decide for themselves (being independent of the judge's decision).

5.2 LIMITATIONS

While we achieved satisfactory results in the case study and experiments conducted, our work faced certain limitations.

Our dataset is only composed of judgements (the document representing the final 1st-degree decision in the lawsuit). When we started the research, these were the available and processed data. We did not have access to the other process documents (i.e., the claim and the defence). Even having ulterior access to them, these documents are scanned images or PDFs created with different headers and footers, which would require costly work of Optical Character Recognition (OCR) and manual document cleaning. To not delay our research schedule, we decided to focus on the judgements and get results from them.

If the data used – as well as data from other courts – were already organised and provided by the Brazilian Judiciary in a processable form, we would not have spent time on manual collection and extraction. Also, we could have performed more experiments, improving the proposed model. The efforts toward data governance in the Brazilian judiciary are still recent, and it will be essential to move forward in innovation.

Moreover, all the ML techniques used to predict an outcome are incapable of explaining it. We extracted the factors and search for a relation between them and the results in an attempt to provide an explanation for the parties. On a limited basis, this explanation was offered but through our intervention, it did not occur automatically.

Another problem we found was related to the position of some airlines in the conciliation hearings. In some cases, they send a hired lawyer who does not know the case just to attend the hearing and avoid default. Consequently, this lawyer does not present any offer and has no power to negotiate. This practice discourages any agreement attempt during the hearing, either with or without the model's assistance. In this regard, we would have applied the model in more cases to increase our validation sample. However, this requires previous scheduling and a time interval for inclusion in the conciliation agenda and parties' notification, which could delay the course completion.

Finally, due to our limited resources (human and funds), the model was applied through a simple presentation to the parties (see Appendix F). We cannot construct an ODR platform to the parties and conciliators easily and directly access the predictions and statistics about the lawsuits.

5.3 FUTURE WORK

As future work, in terms of experimenting, we intend to use XAI techniques, such as SHAP (LUNDBERG; LEE, S.-I., 2017) and LIME (RIBEIRO *et al.*, 2016), to assess the importance of an attribute/judgement factor in the prediction, as well as to detect relevant parts of the text that the classifier or regressor has considered.

In terms of an E-Justice project, the proposed model can be used by the Brazilian Judiciary, specifically in other JECs. To provide the model's autonomy and dispense our participation in conciliation hearings, we can construct a ODR system in which anyone (parties, conciliators, lawyers and judges) has easy and open access to the predictions and explanations. It would be possible by indicating, in a questionnaire, some factors related to a particular case (for example, if there was a flight delay, what the delay interval, if there was a temporary baggage loss, for how many days, etc.).

For this purpose, we need access to the judgements of other JECs and extract the same information. Also, it is necessary to involve other researchers in Law, Computer Science and Software Engineering. Hence, the parties should be instructed by the JECs to consult the system before the conciliation hearing or even before filing a lawsuit. This is a final suggestion for an arrangement between the University and the State Court.

5.4 PUBLICATIONS

It follows a list of the main papers published during the PhD program up to the present time:

1. SABO, I. C.; DAL PONT, T. R.; ROVER, A. J.; HUBNER, J. F. Classificação de sentenças de Juizado Especial Cível utilizando aprendizado de máquina. **Revista Democracia Digital e Governo Eletrônico**, v. 1, n. 18, p. 94-106, 2019.

2. SABO, I. C.; ROVER, A. J. Resolução de conflitos online e técnicas de Inteligência artificial: uma revisão sistemática da literatura. In: X ENCONTRO INTERNACIONAL

DO CONPEDI VALËNCIA/ESPANHA, 2020, Valência. Direito, Governança e Novas Tecnologias. **Anais...** Florianópolis: CONPEDI, 2020. p. 170-186.

3. SABO, I. C.; ROVER, A. J. Observância de precedentes e gestão de demandas repetitivas através do aprendizado de máquina. **Revista Opinião Jurídica (Fortaleza)** v. 18, n. 28, p. 69-93, 2020.
4. SABO, I. C.; KURTZ, L. P. ; REGINALDO, P. A. ; SANTOS, P. M. ; ROVER, A. J. Entraves ao governo aberto na Justiça Federal brasileira. **Revista Direito GV (online)**, v. 16, p. e1950, 2020.
5. SANTOS, P. M. ; KURTZ, L. P. ; SABO, I. C. ; REGINALDO, P. A. ; ROVER, A. J. **Ferramentas, funcionalidades e procedimentos para um processo eletrônico mais célere**. Erechim: Deviant, 2020.
6. DAL PONT, T. R.; SABO, I. C.; HUBNER, J. F.; ROVER, A. J. Impact of text specificity and size on word embeddings performance: an empirical evaluation in Brazilian legal domain. In: Cerri R.; Prati R.C. (Org.). **Lecture Notes in Computer Science**. Cham: Springer International Publishing, 2020, v. 12319, p. 521-535.
7. SABO, I. C.; ROVER, A. J. Audiência de conciliação e mediação inteligentes: aplicações de aprendizado de máquina e sistemas multiagentes. **Revista Meritum**, v. 16, n. 2, p. 177-200, 2021.
8. SABO, I. C.; DAL PONT, T. R.; WILTON, P. E. V.; ROVER, A. J.; HUBNER, J. F. Clustering of Brazilian legal judgments about failures in air transport service: an evaluation of different approaches. **Artificial Intelligence and Law**, v. 30, n. 1, p. 1-35, 2021.
9. DAL PONT, T. R.; SABO, I. C.; WILTON, P. E. V.; MENEZES, V. A.; COPETTI, R.; ZAMBROTA, L.; MARTINS, P. P.; COSTA, E. C.; SCHNITZLER, E. L.; SANTOS, P. M.; CUNHA, R. R.; KASTER, G. B.; ROVER, A. J.. Classification and association rules in Brazilian Supreme Court judgments on pre-trial detention. In: Kö A., Francesconi E., Kotsis G., Tjoa A.M., Khalil I. (Org.). **Lecture Notes in Computer Science**. Springer International Publishing, 2021, v. 12926, p. 131-142.
10. DAL PONT, T. R.; SABO, I. C.; HUBNER, J. F.; ROVER, A. J. Regression applied to legal judgments to predict compensation for immaterial damage. **PeerJ Computer Science** *(In press)*, 2022.
11. SABO, I. C.; BILLI, M.; LAGIOIA, F.; SARTOR, G.; ROVER, A. J. Unsupervised factor extraction from pretrial detention decisions by Italian and Brazilian Supreme Courts. In: 1ST INTERNATIONAL WORKSHOP ON DIGITAL JUSTICE, DIGITAL LAW, AND CONCEPTUAL MODELING, 2022, Hyderabad, India. **Proceedings...** Hyderabad: 2022 *(In press)*.

REFERENCES

ABRAHAMS, Brooke; BELLUCCI, Emilia; ZELEZNIKOW, John. Incorporating fairness into development of an integrated multi-agent online dispute resolution environment. **Group Decision and Negotiation**, Springer, V. 21, N. 1, p. 3–28, 2012.

ADADI, Amina; BERRADA, Mohammed. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). **IEEE Access**, IEEE, V. 6, p. 52138–52160, 2018.

AGGARWAL, Charu C. **Machine learning for text**. Cham: Springer, 2018.

AGGARWAL, Charu C; ZHAI, ChengXiang. A survey of text classification algorithms. *In*: AGGARWAL, Charu C; ZHAI, ChengXiang (Eds.). **Mining text data**. New York: Springer, 2012. p. 163–222.

ALETRAS, Nikolaos; TSARAPATSANIS, Dimitrios; PREOŢIUC-PIETRO, Daniel; LAMPOS, Vasileios. Predicting judicial decisions of the European Court of Human Rights: a natural language processing perspective. **PeerJ Computer Science**, PeerJ Inc., V. 2, e93, 2016.

ALPAYDIN, Ethem. **Introduction to Machine Learning**. 2. ed. Cambridge, MA: MIT Press, 2009.

ASHLEY, Kevin D. Prospects for Legal Analytics: Some Approaches to Extracting More Meaning from Legal Texts. **University of Cincinnati Law Review**, V. 90, N. 4, p. 5, 2022.

BELLUCCI, Emilia. Development of a mediation tool in Family Law: AssetDivider. *In*: CONFERENCE ON LEGAL KNOWLEDGE AND INFORMATION SYSTEMS (JURIX), 21, 2008, FLORENCE. **Proceedings...** Amsterdam, The Netherlands: IOS Press, 2008. p. 88–97. Available from: `https://dl.acm.org/doi/10.5555/1564008.1564019`. Visited on: 9 July 2021.

BELLUCCI, Emilia; ZELEZNIKOW, John. Developing Negotiation Decision Support Systems that support mediators: a case study of the Family_Winner system. **Artificial Intelligence and Law**, Springer, V. 13, N. 2, p. 233–271, 2005.

BENJAMIM, Antonio Herman V. O transporte aéreo e o código de defesa do consumidor. **Revista de Direito do Consumidor**, V. 100, N. 24, p. 23–37, 2015.

BOJANOWSKI, Piotr; GRAVE, Edouard; JOULIN, Armand; MIKOLOV, Tomas. Enriching word vectors with subword information. **Transactions of the Association for Computational Linguistics**, MIT Press, V. 5, p. 135–146, 2017.

BRASIL. Emenda Constitucional nº 45 de 30 de dezembro de 2004. Altera dispositivos dos arts. 5º, 36, 52, 92, 93, 95, 98, 99, 102, 103, 104, 105, 107, 109, 111, 112, 114, 115, 125, 126, 127, 128, 129, 134 e 168 da Constituição Federal, e acrescenta os arts. 103-A, 103B, 111-A e 130-A, e dá outras providências. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 30 Dec. 2004. Available from: `http://www.planalto.gov.br/ccivil_03/constituicao/emendas/emc/emc45.htm`. Visited on: 7 June 2021.

BRASIL. Lei nº 11.419, de 19 de dezembro de 2006. Dispõe sobre a informatização do processo judicial; altera a Lei nº 5.869, de 11 de janeiro de 1973 - Código de Processo Civil; e dá outras providências. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 19 Dec. 2006. Available from: `http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2006/lei/l11419.htm`. Visited on: 7 June 2021.

BRASIL. Lei nº 13.105, de 16 de março de 2015. Código de Processo Civil. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 16 Mar. 2015a. Available from: `http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13105.htm`. Visited on: 7 June 2021.

BRASIL. Lei nº 13.140, de 26 de junho de 2015. Dispõe sobre a mediação entre particulares como meio de solução de controvérsias e sobre a autocomposição de conflitos no âmbito da administração pública; altera a Lei nº 9.469, de 10 de julho de 1997, e o Decreto nº 70.235, de 6 de março de 1972; e revoga o § 2º do art. 6º da Lei nº 9.469, de 10 de julho de 1997. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 26 June 2015b. Available from: `http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13140.htm`. Visited on: 1 July 2021.

BRASIL. Lei nº 8.078, de 11 de setembro de 1990. Dispõe sobre a proteção do consumidor e dá outras providências. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 11 Sept. 1990. Available from: `http://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm`. Visited on: 7 June 2021.

BRASIL. Lei nº 9.099, de 26 de setembro de 1995. Dispõe sobre os Juizados Especiais Cíveis e Criminais e dá outras providências. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 26 Sept. 1995. Available from: `http://www.planalto.gov.br/ccivil_03/leis/l9099.htm`. Visited on: 7 June 2021.

BRASIL. CONSTITUIÇÃO FEDERAL (1988). **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República, 1988. Available from: `http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm`. Visited on: 23 July 2021.

BUSSAB, Wilton de O; MORETTIN, Pedro A. Estatística básica. *In*: São Paulo: Saraiva, 2010.

CALMON, Petrônio. **Fundamentos da mediação e da conciliação**. Rio de Janeiro: Forense, 2007.

CAMPBELL, Joshua Charles; HINDLE, Abram; STROULIA, Eleni. Latent Dirichlet allocation: extracting topics from software engineering data. *In*: BIRD, Christian; MENZIES, Tim; ZIMMERMANN, Thomas (Eds.). **The art and science of analyzing software data**. Massachusetts: Morgan Kaufmann, 2015. p. 139–159.

CAPUANO, Nicola; LONGHI, Andrea; SALERNO, Saverio; TOTI, Daniele. Ontology-driven generation of training paths in the legal domain. **International Journal of Emerging Technologies in Learning (iJET)**, V. 10, N. 7, p. 14–22, 2015.

CAPUANO, Nicola; TOTI, Daniele. Experimentation of a smart learning system for law based on knowledge discovery and cognitive computing. **Computers in Human Behavior**, V. 92, p. 459–467, 2019.

CARNEIRO, Davide; GOMES, Marco; COSTA, Ângelo; NOVAIS, Paulo; NEVES, José. Enriching conflict resolution environments with the provision of context information. **Expert Systems**, Wiley Online Library, V. 34, N. 5, e12049, 2017.

CARNEIRO, Davide; NOVAIS, Paulo; ANDRADE, Francisco; ZELEZNIKOW, John; NEVES, José. Online dispute resolution: an artificial intelligence perspective. **Artificial Intelligence Review**, Springer, V. 41, N. 2, p. 211–240, 2014.

CARNEIRO, Davide; NOVAIS, Paulo; ANDRADE, Francisco; ZELEZNIKOW, John; NEVES, José. Using case-based reasoning and principled negotiation to provide decision support for dispute resolution. **Knowledge and Information Systems**, Springer, V. 36, N. 3, p. 789–826, 2013a.

CARNEIRO, Davide; NOVAIS, Paulo; NEVES, José. An Agent-Based Architecture for Multifaceted Online Dispute Resolution Tools. *In*: MEHROTRA, Kishan G.; MOHAN, Chilukuri; OH, Jae C.; VARSHNEY, Pramod K.; ALI, Moonis (Eds.). **Developing Concepts in Applied Intelligence**. Berlin, Heidelberg: Springer, 2011. p. 89–94.

CARNEIRO, Davide; NOVAIS, Paulo; NEVES, José. Using genetic algorithms to create solutions for conflict resolution. **Neurocomputing**, Elsevier, V. 109, p. 16–26, 2013b.

CHAI, Tianfeng; DRAXLER, Roland R. Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. **Geoscientific model development**, Copernicus GmbH, V. 7, N. 3, p. 1247–1250, 2014.

CHAPPELL, David. Introducing azure machine learning: a guide for technical professionals. **Microsoft Corporation**, 2015.

CICHOSZ, Pawel. **Data mining algorithms: explained using R**. Chichester: John Wiley & Sons, 2015.

CONSELHO NACIONAL DE JUSTIÇA (CNJ). **Justiça em Números 2021**: ano-base 2020. Brasília, DF, 2021a. Available from: `https://www.cnj.jus.br/wp-content/uploads/2021/09/relatorio-justica-em-numeros2021-12.pdf`. Visited on: 7 July 2022.

CONSELHO NACIONAL DE JUSTIÇA (CNJ). Portaria nº 271 de 4 de dezembro de 2020. Regulamenta o uso de Inteligência Artificial no âmbito do Poder Judiciário. **Diário da Justiça eletrônico**, Brasília, DF, 4 Dec. 2020a. Available from: `https://atos.cnj.jus.br/atos/detalhar/3613`. Visited on: 16 June 2021.

CONSELHO NACIONAL DE JUSTIÇA (CNJ). Programa Resolve, 2018. Available from: `https://www.cnj.jus.br/programas-e-acoes/programa-resolve/`. Visited on: 7 June 2021.

CONSELHO NACIONAL DE JUSTIÇA (CNJ). Resolução nº 125 de 29 de novembro de 2010. Dispõe sobre a Política Judiciária Nacional de tratamento adequado dos conflitos de interesses no âmbito do Poder Judiciário e dá outras providências. **Diário da Justiça**

eletrônico, Brasília, DF, 29 Nov. 2010. Available from: `https://atos.cnj.jus.br/atos/detalhar/156`. Visited on: 7 June 2021.

CONSELHO NACIONAL DE JUSTIÇA (CNJ). Resolução nº 331 de 20 de agosto de 2020. Institui a Base Nacional de Dados do Poder Judiciário - DataJud como fonte primária de dados do Sistema de Estatística do Poder Judiciário - SIESPJ para os tribunais indicados nos incisos II a VII do art. 92 da Constituição Federal. **Diário da Justiça eletrônico**, Brasília, DF, 20 Aug. 2020b. Available from: `https://atos.cnj.jus.br/atos/detalhar/3428`. Visited on: 16 June 2021.

CONSELHO NACIONAL DE JUSTIÇA (CNJ). Resolução nº 332 de 21 de agosto de 2020. Dispõe sobre a ética, a transparência e agovernança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências. **Diário da Justiça eletrônico**, Brasília, DF, 21 Aug. 2020c. Available from: `https://atos.cnj.jus.br/atos/detalhar/3429`. Visited on: 16 June 2021.

CONSELHO NACIONAL DE JUSTIÇA (CNJ). Resolução nº 395 de 7 de junho de 2021. Institui a Política de Gestão da Inovação no âmbito do Poder Judiciário. **Diário da Justiça eletrônico**, Brasília, DF, 7 June 2021b. Available from: `https://atos.cnj.jus.br/atos/detalhar/3973`. Visited on: 16 June 2021.

COVER, Thomas M.; THOMAS, Joy A. Elements of information theory second edition solutions to problems. **Internet Access**, p. 19–20, 2006.

DAL PONT, Thiago Raulino. **Representation, classification and regression techniques applied to legal judgments about immaterial damage due to failures in air transport services**. 2021. 120p: Dissertação (Mestrado em Engenharia de Automação e Sistemas) - Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis, 2021.

DASGUPTA, Dipankar; MICHALEWICZ, Zbigniew. **Evolutionary algorithms in engineering applications**. [*S.l.*]: Springer Science & Business Media, 2013.

DEMŠAR, Janez *et al.* Orange: data mining toolbox in Python. **Journal of Machine Learning Research**, V. 14, N. 1, p. 2349–2353, 2013.

DEVORE, Jay L. **Probability and Statistics for Engineering and the Sciences**. 8. ed. Andover: Cengage Learning, 2011.

DINIZ, Maria Helena. Proteção jurídica da existencialidade. **Revista Eletrônica Direito e Sociedade (REDES)**, Centro Universitario La Salle (UNILASALLE), V. 8, N. 2, p. 181, 2020.

DONEFF, Andrea; ORDOVER, Abraham. **Alternatives to Litigation**: mediation, arbitration, and the art of dispute resolution. 3. ed. United States: Wolters Kluwer Law & Business, 2014.

DREISEITL, Stephan; OHNO-MACHADO, Lucila. Logistic regression and artificial neural network classification models: a methodology review. **Journal of biomedical informatics**, Elsevier, V. 35, N. 5-6, p. 352–359, 2002.

EL JELALI, Soufiane; FERSINI, Elisabetta; MESSINA, Enza. Legal retrieval as support to eMediation: matching disputant's case and court decisions. **Artificial Intelligence and Law**, Springer, V. 23, N. 1, p. 1–22, 2015.

FARIA, José Eduardo. **O Estado e o Direito depois da crise**. São Paulo: Saraiva, 2011.

FENANZO JR, A. Darwinian evolution as a paradigm for AI research. **ACM SIGART Bulletin**, ACM New York, NY, USA, N. 97, p. 22–23, 1986.

FERSINI, Elisabetta; MESSINA, Enza; MANENTI, Lorenza; BAGNARA, Giuliana; EL JELALI, Soufiane; AROSIO, Gaia. eMediation-Towards Smart Online Dispute Resolution. *In*: INTERNATIONAL CONFERENCE ON KNOWLEDGE MANAGEMENT AND INFORMATION SHARING (KMIS), 2014, ROME. **Proceedings...** Setúbal, Portugal: SciTePress, 2014. p. 228–236. Available from: `https://www.scitepress.org/Link.aspx?doi=10.5220/0005080002280236`. Visited on: 14 July 2021.

FREY, Brendan J; DUECK, Delbert. Clustering by passing messages between data points. **Science**, American Association for the Advancement of Science, V. 315, N. 5814, p. 972–976, 2007.

GARCÍA, Salvador; LUENGO, Julián; HERRERA, Francisco. **Data preprocessing in data mining**. Switzerland: Springer International Publishing, 2015. V. 72.

GLASER, Ingo; SCEPANKOVA, Elena; MATTHES, Florian. Classifying semantic types of legal sentences: Portability of machine learning models. *In*: PALMIRANI, Monica (Ed.).

Legal Knowledge and Information Systems (JURIX 2018). Amsterdam: IOS Press, 2018. p. 61–70.

GOMES, Marco; OLIVEIRA, Tiago; CARNEIRO, Davide; NOVAIS, Paulo; NEVES, José. Studying the effects of stress on negotiation behavior. **Cybernetics and Systems**, Taylor & Francis, V. 45, N. 3, p. 279–291, 2014.

GONÇALVES, Carlos Roberto. **Responsabilidade civil**. São Paulo: Saraiva, 2020.

GOZAÍNI, Osvaldo Alfredo. **Formas alternativas para la resolución de conflictos**: arbitraje, mediación, conciliación, ombudsman, procesos alternativos. Buenos Aires: Depalma, 1995.

GUIDOTTI, Riccardo; MONREALE, Anna; RUGGIERI, Salvatore; TURINI, Franco; GIANNOTTI, Fosca; PEDRESCHI, Dino. A survey of methods for explaining black box models. **ACM Computing Surveys (CSUR)**, ACM New York, NY, USA, V. 51, N. 5, p. 1–42, 2018.

HAIR, Joseph F.; BLACK, Bill; BLACK, William C.; BABIN, Barry J.; ANDERSON, Rolph E. **Multivariate data analysis**. Andover: Cengage Learning, 2019.

HAN, Jiawei; PEI, Jian; YIN, Yiwen. Mining frequent patterns without candidate generation. **ACM sigmod record**, ACM New York, NY, USA, V. 29, N. 2, p. 1–12, 2000.

HARRIS, Joshua D; QUATMAN, Carmen E; MANRING, MM; SISTON, Robert A; FLANIGAN, David C. How to write a systematic review. **The American journal of sports medicine**, SAGE Publications Sage CA: Los Angeles, CA, V. 42, N. 11, p. 2761–2768, 2014.

HAWKINS, Douglas M. The problem of overfitting. **Journal of chemical information and computer sciences**, ACS Publications, V. 44, N. 1, p. 1–12, 2004.

HENDRYCKS, Dan; BURNS, Collin; CHEN, Anya; BALL, Spencer. CUAD: An expert-annotated nlp dataset for legal contract review. **arXiv preprint arXiv:2103.06268**, 2021.

HENSLER, Deborah R. Our courts, ourselves: how the alternative dispute resolution movement is re-shaping our legal system. **Penn St. L. Rev.**, HeinOnline, V. 108, p. 165, 2003.

HU, Xia; LIU, Huan. Text analytics in social media. *In*: AGGARWAL, Charu C.; ZHAI, Cheng Xiang (Eds.). **Mining text data**. New York: Springer, 2012. P. 385–414.

HUSBANDS, Philip; HARVEY, I; CLIFF, D; MILLER, G. Artificial evolution: A new path for artificial intelligence? **Brain and cognition**, Elsevier, V. 34, N. 1, p. 130–159, 1997.

JIANG, Shengyi; PANG, Guansong; WU, Meiling; KUANG, Limin. An improved K-nearest-neighbor algorithm for text categorization. **Expert Systems with Applications**, Elsevier, V. 39, N. 1, p. 1503–1509, 2012.

JIVANI, Anjali Ganesh *et al.* A comparative study of stemming algorithms. **Int. J. Comp. Tech. Appl**, V. 2, N. 6, p. 1930–1938, 2011.

JOLLIFFE, Ian T.; CADIMA, Jorge. Principal component analysis: a review and recent developments. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, V. 374, N. 2065, p. 20150202, 2016.

JURAFSKY, Daniel; MARTIN, James H. **Speech and Language Processing: an introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition**. [*S.l.*]: Stanford University, 2019.

KARYSTINOS, George N; PADOS, Dimitrios A. On overfitting, generalization, and randomly expanded training sets. **IEEE Transactions on Neural Networks**, IEEE, V. 11, N. 5, p. 1050–1057, 2000.

KATSH, Ethan. ODR: a look at history - a few thoughts about the present and some speculation about the future. *In*: WAHAB, Mohamed S.; KATSH, Ethan; RAINEY, Daniel (Eds.). **Online Dispute Resolution: theory and practice: a treatise on technology and dispute resolution**. The Hague, Netherlands: Eleven International Publishing, 2011. p. 21–33.

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**, Public Library of Science San Francisco, Califórnia, V. 12, N. 4, e0174698, 2017.

KATZ, Daniel Martin; BOMMARITO, Michael J.; BLACKMAN, Josh. Predicting the behavior of the supreme court of the united states: A general approach. **arXiv preprint arXiv:1407.6333**, 2014.

KELLEHER, John D. **Deep Learning**. London: MIT Press, 2019.

KOHAVI, Ron; PROVOST, Foster. Glossary of terms. **Journal of Machine Learning**, Special Issue on Applications of Machine Learning and the Knowledge Discovery Process 1998.

KOTU, Vijay; DESHPANDE, Bala. **Data Science**: concepts and practice. 2. ed. Cambridge, MA: Morgan Kaufmann (Elsevier Science), 2019.

KOWSARI, Kamran; JAFARI MEIMANDI, Kiana; HEIDARYSAFA, Mojtaba; MENDU, Sanjana; BARNES, Laura; BROWN, Donald. Text classification algorithms: A survey. **Information**, Multidisciplinary Digital Publishing Institute, V. 10, N. 4, p. 150, 2019.

KUHN, Max; JOHNSON, Kjell, *et al.* **Applied predictive modeling**. New York: Springer, 2013. V. 26.

KWARTLER, Ted. **Text mining in practice with R**. Chichester: John Wiley & Sons, 2017.

LEE, Mong Li; LU, Hongjun; LING, Tok Wang; KO, Yee Teng. Cleansing data for mining and warehousing. *In*: INTERNATIONAL CONFERENCE ON DATABASE AND EXPERT SYSTEMS APPLICATIONS (DEXA). **Proceedings...** [*S.l.*: *s.n.*], 1999. p. 751–760. Available from: `https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.656&rep=rep1&type=pdf`. Visited on: 23 June 2021.

LIBERATI, Alessandro *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. **Journal of clinical epidemiology**, Elsevier, V. 62, N. 10, e1–e34, 2009.

LIEBERMAN, Jethro K; HENRY, James F. Lessons from the alternative dispute resolution movement. **The University of Chicago law review**, JSTOR, V. 53, N. 2, p. 424–439, 1986.

LIPPI, Marco; PAŁKA, Przemysław; CONTISSA, Giuseppe; LAGIOIA, Francesca; MICKLITZ, Hans-Wolfgang; SARTOR, Giovanni; TORRONI, Paolo. CLAUDETTE: an automated detector of potentially unfair clauses in online terms of service. **Artificial Intelligence and Law**, Springer, V. 27, N. 2, p. 117–139, 2019.

LIU, Fei Tony; TING, Kai Ming; ZHOU, Zhi-Hua. Isolation forest. *In*: IEEE INTERNATIONAL CONFERENCE ON DATA MINING (ICDM), 8, 2008, PISA. **Proceedings...** Pisa, Italy: [*s.n.*], 2008. p. 413–422. Available from: `https://ieeexplore.ieee.org/document/4781136`. Visited on: 28 Feb. 2022.

LIU, Raymond; GILLIES, Duncan F. Overfitting in linear feature extraction for classification of high-dimensional image data. **Pattern Recognition**, Elsevier, V. 53, p. 73–86, 2016.

LODDER, Arno R.; ZELEZNIKOW, John. Artificial intelligence and Online Dispute Resolution. *In*: WAHA, Mohamed S. Abdel; KATSH, Ethan; RAINEY, Daniel (Eds.). **Online Dispute Resolution: Theory and Practice: A Treatise on Technology and Dispute Resolution**. The Netherlands: Eleven International Publishing, 2012. p. 73–94.

LOPES, Saskya Miranda; LANIADO, Ruthy Nadia. Os Juizados Especiais Cíveis de Defesa do Consumidor: a ampliação dos direitos no Estado democrático. **Política & Sociedade**, V. 9, N. 17, p. 359–392, 2010.

LUHMANN, Niklas. **Sociologia do direito I**. Rio de Janeiro: Tempo Brasileiro, 1983.

LUNDBERG, Scott M; LEE, Su-In. A unified approach to interpreting model predictions. **Advances in neural information processing systems**, V. 30, 2017.

MANIA, Karolina. Online dispute resolution: the future of justice. **International Comparative Jurisprudence**, V. 1, N. 1, p. 76–86, 2015.

MANNING, Christopher; RAGHAVAN, Prabhakar; SCHÜTZE, Hinrich. Introduction to information retrieval. **Natural Language Engineering**, Cambridge university press, V. 16, N. 1, p. 100–103, 2010.

MATURANA, Humberto R.; VARELA, Francisco J. **A árvore do conhecimento**: as bases biológicas da compreensão humana. São Paulo: Palas Athena, 2011.

MENDES-MOREIRA, João; SOARES, Carlos; JORGE, Alípio Mário; SOUSA, Jorge Freire de. Ensemble approaches for regression: a survey. **Acm computing surveys (csur)**, New York, V. 45, N. 1, p. 1–40, 2012.

MEZZAROBA, Orides; MONTEIRO, Cláudia Servilha. **Manual de metodologia da pesquisa no Direito**. 5. ed. São Paulo: Saraiva, 2009.

MILLER, Tim. Explanation in artificial intelligence: Insights from the social sciences. **Artificial intelligence**, Elsevier, V. 267, p. 1–38, 2019.

MITCHELL, Tom M. **Machine Learning**. Redmond, WA/Ithaca, NY: McGraw-Hill Education, 1997.

MOHRI, Mehryar; ROSTAMIZADEH, Afshin; TALWALKAR, Ameet. **Foundations of Machine Learning**. 2. ed. Cambridge, MA: MIT Press, 2018.

MOORE, Christopher W. **O processo de mediação**: estratégias práticas para a resolução de conflitos. 2. ed. Porto Alegre: ArtMed, 1998.

MORAIS, José Luis Bolzan; SPENGLER, Fabiana Marion. **Mediação e arbitragem**: alternativas à jurisdição. Porto Alegre: Livraria do Advogado, 1999.

MORGAN, George A; LEECH, Nancy L; GLOECKNER, Gene W; BARRETT, Karen C. **SPSS for introductory statistics: Use and interpretation**. 2. ed. New York: Psychology Press, 2004.

OSIŃSKI, Stanisław; STEFANOWSKI, Jerzy; WEISS, Dawid. Lingo: Search results clustering algorithm based on singular value decomposition. *In*: KŁOPOTEK, Mieczysław A.; WIERZCHOŃ, Sławomir T.; TROJANOWSKI, Krzysztof (Eds.). **Intelligent information processing and web mining**. Berlin, Heidelberg: Springer, 2004. p. 359–368.

OSIŃSKI, Stanisław; WEISS, Dawid. **Carrot2**: User and Developer Manual for version 3.17.0-SNAPSHOT. [*S.l.*: *s.n.*], 2019.

PAPADIMITRIOU, Christos H; RAGHAVAN, Prabhakar; TAMAKI, Hisao; VEMPALA, Santosh. Latent semantic indexing: A probabilistic analysis. **Journal of Computer and System Sciences**, Elsevier, V. 61, N. 2, p. 217–235, 2000.

PENG, Hao; LI, Jianxin; HE, Yu; LIU, Yaopeng; BAO, Mengjiao; WANG, Lihong; SONG, Yangqiu; YANG, Qiang. Large-scale hierarchical text classification with recursively regularized deep graph-cnn. *In*: WORLD WIDE WEB CONFERENCE (WWW), 2018, LYON. **Proceedings...** Lyon, France: [*s.n.*], 2018. p. 1063–1072. Available from: `https://dl.acm.org/doi/pdf/10.1145/3178876.3186005`. Visited on: 1 Feb. 2022.

PETERS, Georg; CRESPO, Fernando; LINGRAS, Pawan; WEBER, Richard. Soft clustering–fuzzy and rough approaches and their extensions and derivatives. **International Journal of Approximate Reasoning**, Elsevier, V. 54, N. 2, p. 307–322, 2013.

POOLE, David L.; MACKWORTH, Alan K. **Artificial Intelligence**: foundations of computational agents. New York: Cambridge University Press, 2010.

REALE, Miguel. Moral Damages in Brazilian Law. **Panorama of Brazilian Law**, p. 121–127, 1992.

RIBEIRO, Marco Tulio; SINGH, Sameer; GUESTRIN, Carlos. " Why should i trust you?" Explaining the predictions of any classifier. *In*: PROCEEDINGS of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. [*S.l.*: *s.n.*], 2016. P. 1135–1144.

ROSENBERG, Joshua D.; FOLBERG, H. Jay. Alternative dispute resolution: an empirical analysis. **Stanford Law Review**, JSTOR, p. 1487–1551, 1994.

ROUSSEEUW, Peter J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. **Journal of computational and applied mathematics**, Elsevier, V. 20, p. 53–65, 1987.

ROVER, Aires José. **Informática no Direito**: inteligência artificial: introdução aos sistemas especialistas legais. Curitiba: Juruá, 2001.

RUGGERI, Federico; LAGIOIA, Francesca; LIPPI, Marco; TORRONI, Paolo. Detecting and explaining unfairness in consumer contracts through memory networks. **Artificial Intelligence and Law**, Springer, V. 30, N. 1, p. 59–92, 2022.

RUSSELL, Stuart J.; NORVIG, Peter. **Artificial Intelligence**: a modern approach. New Jersey: Prentice Hall (Pearson Education), 2010.

SABO, Isabela Cristina; DAL PONT, Thiago Raulino; ROVER, Aires José; HÜBNER, Jomi Fred. Classificação de sentenças de Juizado Especial Cível utilizando aprendizado de máquina. **Revista Democracia Digital e Governo Eletrônico**, V. 1, N. 18, p. 94–106, 2019.

SABO, Isabela Cristina; DAL PONT, Thiago Raulino; WILTON, Pablo Ernesto Vigneaux; ROVER, Aires José; HÜBNER, Jomi Fred. Clustering of Brazilian legal judgments about failures in air transport service: an evaluation of different approaches. **Artificial Intelligence and Law**, Springer, p. 1–37, 2021.

SABO, Isabela Cristina; ROVER, Aires José. Resolução de conflitos online e técnicas de inteligência artificial: uma revisão sistemática da literatura. *In*: ENCONTRO INTERNACIONAL DO CONPEDI - DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS, 10, 2019, VALÊNCIA. **Anais eletrônicos...** Florianópolis: CONPEDI, 2020. p. 170–186. Available from: `http://site.conpedi.org.br/publicacoes/150a22r2/ig3a9xr8/QT9AA7zkYjc422K5.pdf`. Visited on: 23 June 2021.

SADIKU, Asmir. Immaterial Damage and Some Types of its Compensation. **Prizren Social Science Journal**, Prizren Social Science Journal, V. 4, N. 1, p. 50–56, 2020.

SAMPAIO, Rosana Ferreira; MANCINI, Marisa Cotta. Estudos de revisão sistemática: um guia para síntese criteriosa da evidência científica. **Brazilian Journal of Physical Therapy**, V. 11, N. 1, p. 83–89, 2007.

SANTOS, Paloma Maria; KURTZ, Lahis Pasquali; SABO, Isabela Cristina; REGINALDO, Paula Adão; ROVER, Aires José. **Ferramentas, funcionalidades e procedimentos para um processo eletrônico mais célere**. Erechim: Deviant, 2020.

SARDETO, Patricia Eliane da Rosa. Contribuições da visão sistêmica de Maturana e Varela para o e-judiciário. *In*: CONGRESSO NACIONAL DO CONPEDI - POLÍTICA JUDICIÁRIA, GESTÃAO E ADMINISTRAÇÃO DA JUSTIÇA, 26, 2017, SÃO LUÍS. **Anais eletrônicos...** Florianópolis: CONPEDI, 2017. [recurso eletrônico on–line]. Available from: `https://ainfo.cnptia.embrapa.br/digital/bitstream/item/106823/1/1786.pdf`. Visited on: 23 June 2021.

SAYGILI, Murathan; MERT, Isa Emre; TOKDEMIR, Onur Behzat. A decentralized structure to reduce and resolve construction disputes in a hybrid blockchain network. **Automation in construction**, Elsevier, V. 134, p. 104056, 2022.

SHARMA, Sugam K; SHANDILYA, Ritu; SHARMA, Swadesh. Predicting Indian Supreme Court Judgments, Decisions, or Appeals: eLegalls Court Decision Predictor (eLegPredict). **Statute Law Review**, 2022.

SILVA, Ivan Nunes da; SPATTI, Danilo Hernane; FLAUZINO, Rogerio Andrade; LIBONI, Luisa Helena Bartocci; ALVES, Silas Franco dos Reis. **Artificial neural networks**: a practical course. Switzerland: Springer, 2018.

SILVA, Nilton Correia da *et al.* Document type classification for Brazil's Supreme Court using a convolutional neural network. *In*: INTERNATIONAL CONFERENCE ON FORENSIC COMPUTER SCIENCE AND CYBER LAW (ICOFCS), 10, 2018, SÃO PAULO, p. 29–30. Available from: https://cic.unb.br/~teodecampos/ViP/correiaDaSilva_etal_icofcs2018.pdf. Visited on: 21 July 2021.

SIMKOVA, Nikola; SMUTNY, Zdenek. Business E-NeGotiAtion: A method using a genetic algorithm for online dispute resolution in B2B relationships. **Journal of Theoretical and Applied Electronic Commerce Research**, MDPI, V. 16, N. 5, p. 1186–1216, 2021.

SMOLENSKY, Paul. Connectionist AI, symbolic AI, and the brain. **Artificial Intelligence Review**, Springer, V. 1, N. 2, p. 95–109, 1987.

ŞULEA, Octavia-Maria; ZAMPIERI, Marcos; VELA, Mihaela; VAN GENABITH, Josef. Predicting the law area and decisions of French Supreme Court cases. *In*: INTERNATIONAL CONFERENCE RECENT ADVANCES IN NATURAL LANGUAGE PROCESSING (RANLP), 2017, VARNA. **Proceedings...** Turkey: INCOMA Ltd., 2017. p. 716–722. Available from: https://doi.org/10.26615/978-954-452-049-6_092. Visited on: 21 July 2021.

SUN, Ron. Artificial intelligence: connectionist and symbolic approaches. **Citeseer**, 1999.

TAN, Pang-Ning; STEINBACH, Michael; KUMAR, Vipin. **Introduction to data mining**. London: Pearson Education, 2014.

THOMPSON, Darin. Creating new pathways to justice using simple artificial intelligence and online dispute resolution. **IJODR**, HeinOnline, V. 2, p. 4, 2015.

TRIVIÑOS, Augusto Nibaldo Silva. **Introdução à pesquisa em Ciências Sociais**: a pesquisa qualitativa em educação. São Paulo: Atlas, 1987.

TSUJII, Jun-ichi; SHIRAI, Yoshiaki. **Artificial Intelligence**: concepts, techniques, and applications. Singapura: Wiley, 1985.

TSUREL, David; DORON, Michael; NUS, Alexander; DAGAN, Arnon; GUY, Ido; SHAHAF, Dafna. E-commerce dispute resolution prediction. *In*: ACM INTERNATIONAL CONFERENCE ON INFORMATION KNOWLEDGE MANAGEMENT (CIKM), 29, 2020, IRELAND. **Proceedings...** Virtual Event, Ireland: [*s.n.*], 2020. P. 1465–1474. Available from: `https://dl.acm.org/doi/pdf/10.1145/3340531.3411906`. Visited on: 7 July 2022.

TUGGENER, Don; DÄNIKEN, Pius von; PEETZ, Thomas; CIELIEBAK, Mark. LEDGAR: a large-scale multi-label corpus for text classification of legal provisions in contracts. *In*: LANGUAGE RESOURCES AND EVALUATION CONFERENCE (LREC), 12, 2020, MARSEILLE. **Proceedings...** Luxemburgo: European Language Resources Association, 2020. p. 1228–1234. Available from: `https://aclanthology.org/2020.lrec-1.155`. Visited on: 21 July 2021.

VANDEGINSTE, B. G. M.; MASSART, D. L.; BUYDENS, L. M. C.; DE JONG, S.; LEWI, P. J.; J., Smeyers-Verbeke. Chapter 30 - Cluster analysis. *In*: VANDEGINSTE, B. G. M.; MASSART, D. L.; BUYDENS, L. M. C.; DE JONG, S.; LEWI, P. J.; J., Smeyers-Verbeke (Eds.). **Data Handling in Science and Technology**. [*S.l.*]: Elsevier, 1998. V. 20. p. 57–86.

VIRTUCIO, Michael Benedict L. *et al.* Predicting decisions of the Philippine Supreme Court using natural language processing and machine learning. *In*: ANNUAL COMPUTER SOFTWARE AND APPLICATIONS CONFERENCE (COMPSAC), 42, 2018, TOKYO. **Proceedings...** New York: IEEE, 2018. p. 130–135. Available from: `https://ieeexplore.ieee.org/document/8377844`. Visited on: 21 July 2021.

WATANABE, Kazuo. Filosofia e características básicas do Juizado Especial de Pequenas Causas. *In*: WATANABE, Kazuo (Ed.). **Juizado Especial de Pequenas Causas: lei n. 7.244/1984**. São Paulo: Revista dos Tribunais, 1985. p. 1–7.

WEISS, Gerhard. **Multiagent Systems**: a modern approach to Distributed Artificial Intelligence. Cambridge, MA: MIT Press, 2001.

WINSTON, Patrick Henry. **Artificial Intelligence**. 3. ed. Reading, MA: Addison-Wesley, 1992.

WONG, Tzu-Tsung. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. **Pattern Recognition**, Elsevier, V. 48, N. 9, p. 2839–2846, 2015.

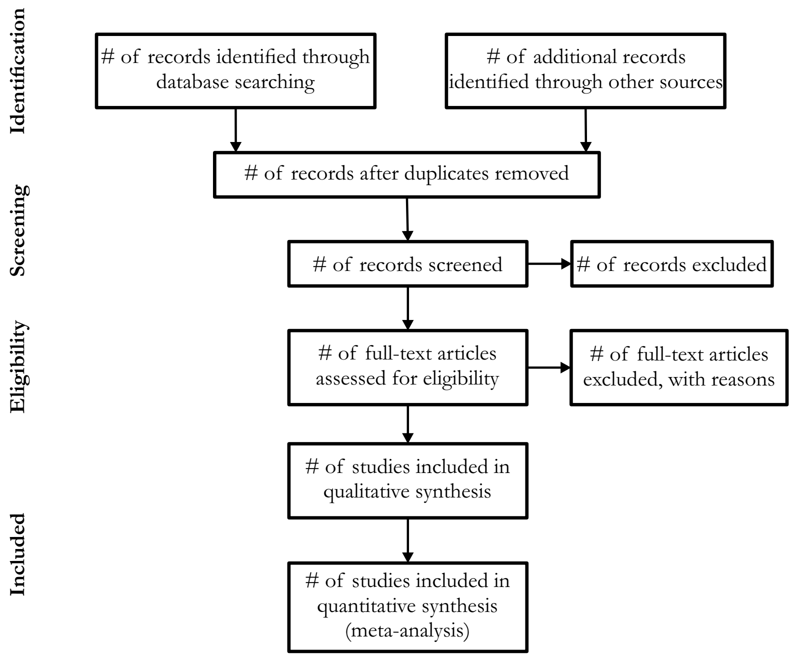
ZHANG, Yanru; HAGHANI, Ali. A gradient boosting method to improve travel time prediction. **Transportation Research Part C: Emerging Technologies**, Elsevier, V. 58, p. 308–324, 2015.

ZHENG, George G. China's grand design of people's smart courts. **Asian Journal of Law and Society**, Cambridge University Press, V. 7, N. 3, p. 561–582, 2020.

APPENDIX A – SYSTEMATIC LITERATURE REVIEW

A systematic review (SLR), as well as other literature reviews, is a type of search in which we use publications/records as a data source. This investigation provides the state-of-art on a given problem by applying explicit search methods, critically appraising of the results and then summarising the selected information. It becomes useful to identify topics that need new studies, helping to guide future research (SAMPAIO; MANCINI, 2007). Figure 36 illustrates the methodological path of a SLR.

Figure 36 – SLR path



Source: Liberati *et al.* (2009).

Based on that methodology, we structured this SLR into: (1) research question and search terms definition; (2) databases and filters; (3) inclusion and exclusion criteria; (4) results and papers selected; (5) quantitative analysis; and (6) qualitative analysis. We conducted searches during the period from 04/03/2019 to 04/05/2019.[11] Then, we repeated and updated the SLR during the period from 07/04/2022 to 07/06/2022.

1. Research question and search terms definition

Starting from the research question "What are the advances in science on the application of AI techniques in ADR?" we set research terms for each knowledge domain covered.

a. ADR/ODR: *"online dispute resolution"*, *"alternative dispute resolution"*, *emediation*.

[11] This appendix is adapted from the paper Sabo and Rover (2020).

b. AI: *"artificial intelligence"*, *"machine learning"*, *"multi-agent system?"*, *"learning system?"*, *"intelligent environment*"*, *"natural language processing"*.

c. Judiciary Branch: *law*, *legal*, *judicial*, *justice*, *court*, *process*.

2. Databases and filters

- *Databases*: Initially, we chosen the databases Scopus, Web of Science and SciELO as repositories of interest because of the interdisciplinary of the research and for indexing journals with high impact factor. We also added Google Scholar to search for any interesting records not seen in the first three databases. Later, in updating the SLR, we included IEEE Xplore and ACM Digital Library, due to further interest in specific literature covering computing and information technology.

- *Document fields*: We limited the document fields by title, abstract and keywords, expecting to obtain assertive papers regarding the cross-referencing of subjects. Exceptionally in Google Scholar, the fields are reduced by title due to limitations of this database.

- *Document type*: We filtered the type by journal articles, intending to return more robust and complete papers at the experimentation level (conference papers usually concern phases of ongoing research). We observed that in Google Scholar is not possible to choose this filter and for ACM we applied an equivalent, named by them as 'research article'.

- *Language, year and publisher*: We did not restrict the search by language, year or publisher. Nevertheless, due to the interdisciplinary aspect of the selected papers, consequently all of them are written in English.

3. Inclusion and exclusion criteria

- *Inclusion*: We selected documents that present experiments, case studies or ADR/ODR systems, which detail the techniques used, considering that we are searching for the AI application in ADR/ODR.

- *Exclusion*: Essentially theoretical articles, or that discuss trends in the area, or with regulatory proposals, or that simply state an ADR/ODR system without explaining it technically were discarded.

4. Results and papers selected

Table 30 indicates the search strategies used in each database and the results obtained.

Table 30 – SLR search strategies

Database	Research terms	Document fields	Document type	TR	DR	SR	NR	NS
Scopus	"online dispute resolution" AND ("artificial intelligence" OR "machine learning" OR "multi-agent system?" OR "learning system?" OR "intelligent environment*" OR "natural language processing") AND (law OR legal OR judicial OR justice OR court OR process)	Title, abstract and keywords	Journal article	14	0	0	13	2
Web of Science	"online dispute resolution" AND ("artificial intelligence" OR "machine learning" OR "multi-agent system?" OR "learning system?" OR "intelligent environment*" OR "natural language processing") AND (law OR legal OR judicial OR justice OR court OR process)	Title, abstract and keywords	Journal article	9	1	0	18	1
SciELO	"online dispute resolution" AND ("artificial intelligence" OR "machine learning" OR "multi-agent system?" OR "learning system?" OR "intelligent environment*" OR "natural language processing") AND (law OR legal OR judicial OR justice OR court OR process)	Title, abstract and keywords	Journal article	0	0	0	0	0
IEEE Xplore	"online dispute resolution" AND ("artificial intelligence" OR "machine learning" OR "multi-agent system?" OR "learning system?" OR "intelligent environment*" OR "natural language processing") AND (law OR legal OR judicial OR justice OR court OR process)	Title, abstract and keywords	Journal article	0	0	0	1	0
ACM Digital Library	"online dispute resolution" AND ("artificial intelligence" OR "machine learning" OR "multi-agent system?" OR "learning system?" OR "intelligent environment*" OR "natural language processing") AND (law OR legal OR judicial OR justice OR court OR process)	Title, abstract and keywords	Research article	0	0	0	10	1
Google Scholar	"online dispute resolution" AND "artificial" AND "intelligence"	Title	All	10	8	3	1	0
	"alternative dispute resolution" AND "artificial" AND "intelligence"			1	1	0	0	0
	"emediation"			6	1	1	1	0

*TR: total results; DR: divergent results; SR: selected results; NR: new results (SLR update); NS: new selected results (SLR update).

After applying the criteria and evaluating the methodological quality of each manuscript, we selected 19 documents, which we list in Table 31.

Table 31 – SLR selected papers

Title	**Authors**	**Affiliation**	**Year**	**Publisher**	**Cited by***	**Document type**
An agent-based architecture for multifaceted online dispute resolution tools	Carneiro, D.; Novais, P.; Neves, J.	University of Minho (Portugal)	2011	Springer International Publishing	3	Book chapter
Incorporating fairness into development of an integrated multi-agent online dispute resolution environment	Abrahams, B.; Bellucci, E.; Zeleznikow, J.	Victoria University (Australia)	2012	Group Decision and Negotiation	27	Journal article
Artificial intelligence and online dispute resolution	Lodder, A.; Zeleznikow, J.	Vrije Universiteit (Amsterdam) and Victoria University (Australia)	2012	Online Dispute Resolution Theory and Practice	33	Book chapter
Using case-based reasoning and principled negotiation to provide decision support for dispute resolution	Carneiro, D.; Novais, P.; Andrade, F.; Zeleznikow, J.; Neves, J.	University of Minho (Portugal) and Victoria University (Australia)	2013	Knowledge and Information Systems	80	Journal article
Using genetic algorithms to create solutions for conflict resolution	Carneiro, D.; Novais, P.; Neves, J.	University of Minho (Portugal)	2013	Neurocomputing	13	Journal article
Studying the effects of stress on negotiation behaviour	Gomes, M.; Oliveira, T.; Carneiro, D.; Novais, P.; Neves, J.	University of Minho (Portugal)	2014	Cybernetics and Systems	19	Journal article

Online dispute resolution: an artificial intelligence perspective	Carneiro, D.; Novais, P.; Andrade, F.; Zeleznikow, J.; Neves, J.	University of Minho (Portugal) and Victoria University (Australia)	2014	Artificial Intelligence Review	81	Journal article
Conflict resolution and its context: from the analysis of behavioural patterns to efficient decision-making	Carneiro, D.; Novais, P.; Neves, J.	University of Minho (Portugal)	2014	Springer International Publishing	40	Book
eMediation: towards smart online dispute resolution	Fersini, E.; Messina, E.; Manenti, L.; Bagnara, G.; El Jelali, S.; Arosio, G.	University of Milano-Bicocca (Italy)	2014	International Conference on Knowledge Management and Information Sharing	4	Proceedings article
Legal retrieval as support to eMediation: matching disputant's case and court decisions	El Jelali, S.; Fersini, E.; Messina, E.	University of Milano-Bicocca (Italy)	2015	Artificial Intelligence and Law	14	Journal article
Ontology-driven generation of training paths in the legal domain	Capuano, N.; Longhi, A.; Salerno, S.; Toti, D.	University of Salerno (Italy)	2015	International Journal of Emerging Technologies in Learning	14	Journal article
Creating new pathways to justice using simple artificial intelligence and online dispute resolution	Thompson, D.	York University (Canada)	2015	International Journal of Online Dispute Resolution	23	Journal article
Enriching conflict resolution environments with the provision of context information	Carneiro, D.; Gomes, M.; Costa, Â.; Novais, P.; Neves, J.	University of Minho (Portugal)	2017	Expert Systems	6	Journal article

Can artificial intelligence and online dispute resolution enhance efficiency and effectiveness in courts	Zeleznikow, J.	Victoria University (Australia)	2017	International Journal for Court Administration	47	Journal article
Experimentation of a smart learning system for law based on knowledge discovery and cognitive computing	Capuano, N.; Toti, D.	University of Salerno and Roma Tre University (Italy)	2019	Computers in Human Behavior	22	Journal article
E-commerce dispute resolution prediction	Tsurel, D.; Doron, M.; Nus, A.; Dagan, A.; Guy, I.; Shahaf, D.	Hebrew University of Jerusalem (Israel) and eBay Research	2020	International Conference on Information and Knowledge Management	2	Proceedings article
China's grand design of people's Smart Courts	Zheng, G.	Shanghai Jiaotong University (China)	2020	Asian Journal of Law and Society	5	Journal article
Business E-NeGotiAtion: a method using a genetic algorithm for online dispute resolution in B2B relationships	Simkova, N.; Smutny, Z.	Masaryk University and Prague University of Economics and Business (Czech Republic)	2021	Journal of Theoretical and Applied Electronic Commerce Research	1	Journal article
A decentralized structure to reduce and resolve construction disputes in a hybrid blockchain network	Murathan, S.; Mert, I.; Tokdemir, O.	Masaryk Middle East Technical University of Ankara (Turkey)	2022	Automation in Construction	1	Journal article

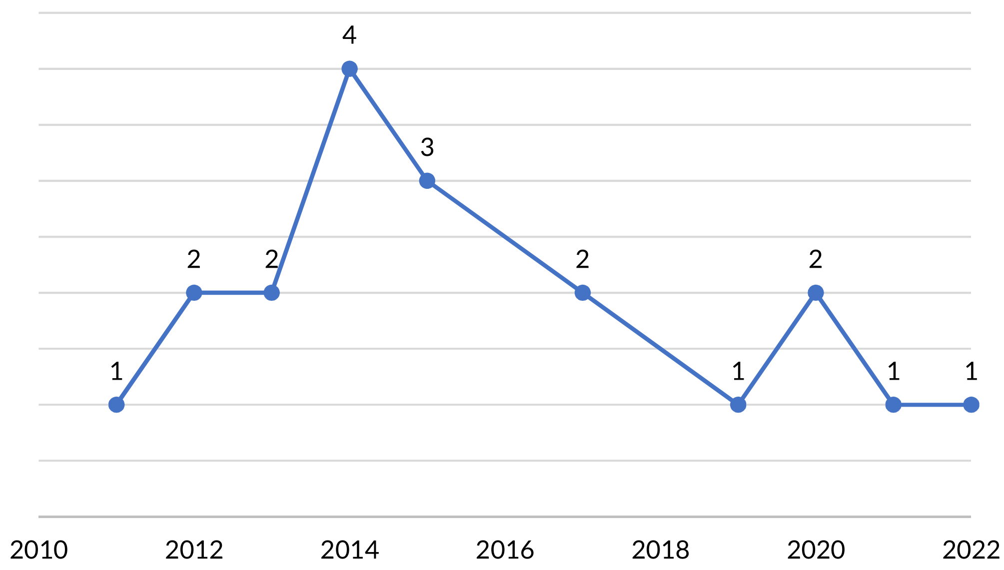
*Updated as counted in Google Scholar on 07/07/2022.

5. Quantitative analysis

The quantitative analysis of this review assesses the year with the highest incidence of papers, the authors, their respective universities and countries that most produced on the topic under discussion.

Figure 37 shows that in the last ten years, 2014 and 2015 had the highest number of publications. Although there are no papers by Brazilian authors, 2014 and 2015 were relevant years in the Brazilian Judiciary in terms of process computerisation and encouragement of ADR, as provisions of the new Code of Civil Procedure (Law n. 13.105/2015). It suggests that these movements are global in the judicial environment.

Figure 37 – (SLR) Number of documents per year



Figures 38 and 39, in turn, show a variety of researchers from different universities around the world who have worked on AI-based ADR/ODR systems. The group from the University of Minho (Portugal) is the one with the most scientific production on the subject, and its works are also the most cited, as can be seen in the Table 31.

Figure 38 – (SLR) Number of documents per author

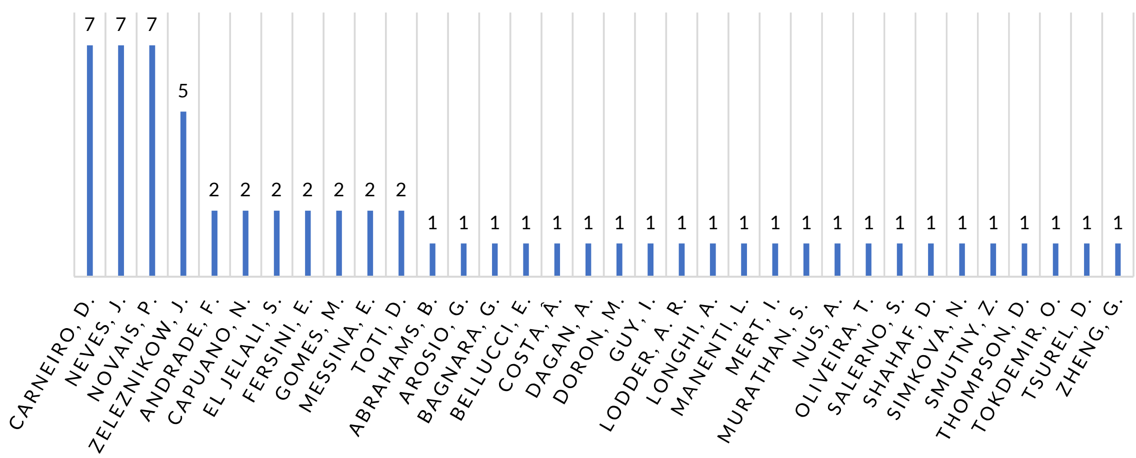
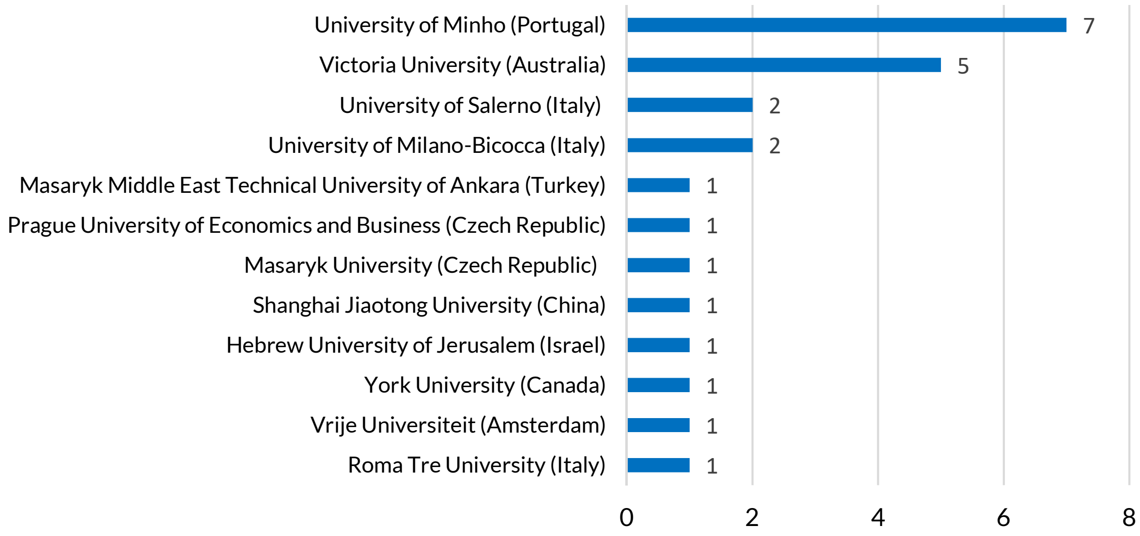


Figure 39 – (SLR) Number of documents per university/country



We note that this variation was widened after the SLR was updated and new articles were included. Due to the Covid-19 pandemic and the exponential growth of online transactions, we can see that research and investment in ODR was a global consequence.

6. Qualitative analysis

The qualitative analysis of this review presents an overview of the AI techniques used to advance the state-of-the-art in the ADR/ODR domain. Figure 40 shows a word cloud formed with the keywords of the articles, in which we can notice highlighted terms in its center.

After reading and evaluating the papers, we observed that in most of them, AI techniques serve to represent legal knowledge in ODR systems for better interaction with the parties in dispute (such as the terms Ontology, Expert Systems). Architecture with more than one agent is also widely used to obtain better agreement options (such as the term Multi-Agent Systems). Genetic algorithm is another approach adopted to select solutions. In addition, the idea of retrieving previous cases to guide the parties has also achieved success in research (see terms Information Retrieval). The details of each work were explained in section 2.2.1.

However, the use of ML and NLP techniques to obtain result estimates in potential or ongoing conflicts have been less explored. In the first SLR, we found only one paper that cursory addresses this topic, indicating a knowledge gap, i.e., a problem to be researched. Then, with the update two more were added, which may suggest a direction of the scientific community towards this field, especially with the further pandemic of Covid-19 and the increase of online negotiations. Consequently, there is a growing need to solve problems with technology support nowadays.

Figure 40 – (SLR) Word cloud

Anyway, even with this lapse of time and the addition of new papers, we still have not found in the context of the Brazilian Judiciary an ODR solution based on ML and NLP to convert data into useful knowledge and to guide the parties to a settlement. We believe that research for this purpose is necessary and original because, besides helping in the settlement culture, it will provide transparency about how the Judiciary decides its cases, encouraging the parties to search for better solutions by themselves.

APPENDIX B – CLUSTERS DESCRIPTIONS OF HIERARCHICAL CLUSTERING

Cluster descriptions (number of documents per cluster)	**Number of documents in disagreement with the descriptions**
C1. Permanent baggage loss / Well-founded or partly founded (6)	1
C2. Temporary baggage loss / Well-founded or partly founded (7)	1
C3. Permanent or temporary baggage loss / Well-founded or partly founded (19)	1
C4. Return flight cancellation due to no-show on the outbound flight / Well-founded or partly founded (7)	0
C5. Waiver of ticket by consumer / Discussion about abusive fine / Repayment problems / Not founded or partly founded (7)	2
C6. Right to regret / Well-founded or partly founded (20)	16
C7. Flight delay or cancellation / Well-founded or partly founded (15)	5
C8. Consumer unfair denied boarding / Well-founded (4)	0
C9. Flight delay or cancellation / Well-founded or partly founded (18)	10
C10. Promotional ticket offer not fulfilled by a specific airline / Well-founded or partly founded (6)	0
C11. Flight delay / Well-founded or partly founded (18)	0
C12. Late flight check-in (consumer fault) / Proven bad weather / Not founded (8)	3
C13. Flight delay by technical problems / Well-founded or partly founded (17)	6
C14. Unjustified flight change or cancellation / Return flight cancellation due to no-show on the outbound flight / Well-founded or partly founded (14)	5
C15. No reply from airline in the time required by law / Well-founded or partly founded (5)	0
C16. Theft of baggage items / Well-founded or partly founded (2)	2
C17. Permanent or temporary baggage loss / International flight / Well-founded or partly founded (17)	3
C18. Flight delay or cancellation / International flight / Well-founded or partly founded (17)	3
C19. Flight delay or cancellation / International flight / Cases subject to the Montreal and Warsaw Convention / Well-founded or partly founded (28)	11
C20. Flight delay or cancellation / International flight / Cases subject to the Montreal and Warsaw Convention / Well-founded or partly founded (35)	1
C21. Permanent or temporary baggage loss / Baggage damaged / International flight / Cases subject to the Montreal and Warsaw Convention / Well-founded or partly founded (21)	4
C22. Flight delay / Judgement by assistant judge / Well-founded or partly founded (5)	0
C23. Flight delay / Judgement by assistant judge / Well-founded or partly founded (4)	0
C24. Flight delay / Judgement by assistant judge / Well-founded or partly founded (10)	0
C25. Flight delay / Judgement by assistant judge / Well-founded or partly founded (4)	1
C26. Flight delay / Well-founded or partly founded (6)	0
C27. Temporary baggage loss / Well-founded or partly founded (4)	0
C28. Theft of baggage items / Well-founded (4)	2
C29. Ticket issued with incorrect personal data (5)	3
C30. Notice about the existence of a previous lawsuit that is identical or similar to the current one (*lis pendens* or connection) (2)	0
C31. No consumer relation / Not founded (1)	0
C32. Illegal act not proven / Not founded (3)	0
C33. Flight delay by technical problems / Lawsuits filed by members of the same family / Well-founded (2)	0
C34. Flight delay / Judgement by assistant judge / Well-founded (2)	0
C35. Flight delay / Judgement by voluntary judge / Well-founded (2)	0
C36. Under-four-hour flight delay / Not founded (2)	0
C37. Flight delay by technical problems / Well-founded or partly founded (6)	1

C38. Waiver of ticket by consumer / Repayment problems / Judgement by voluntary judge / Well-founded or partly founded (2)	0
C39. Flight cancellation / Repayment claim / Well-founded or partly founded (7)	3
C40. Problems with ticketing with loyalty program / Repayment claim / Well-founded or partly founded (4)	2
C41. Temporary baggage loss / Well-founded or partly founded (4)	1
C42. Temporary baggage loss / Well-founded or partly founded (4)	0
C43. No reply from airline in the time required by law / Well-founded or partly founded (4)	1
C44. Flight delay / Not proven bad weather / Well-founded or partly founded (7)	2
C45. Flight delay / Loss of workday or professional engagement / Well-founded (6)	2
C46. Booking cancellation due to airline system failure / Well-founded or partly founded (4)	1
C47. Under-four-hour flight delay / Not founded (8)	1
C48. Passenger who did not follow flight security standards / Not founded (4)	1
C49. Flight delay / Proven bad weather / Not founded (5)	2
C50. Late flight check-in (consumer fault) / Not founded (10)	5
C51. Permanent baggage loss / Well-founded or partly founded (4)	1
C52. Temporary baggage loss / Theft of baggage items / Partly founded (4)	2
C53. Theft of baggage items / Well-founded or partly founded (7)	0
C54. Temporary baggage loss / Well-founded or partly founded (6)	0
C55. Temporary baggage loss / Well-founded or partly founded (9)	1
C56. Waiver of ticket by consumer / Right to regret denied / Not founded (3)	0
C57. Flight cancellation or change at least seventy-two hours in advance / Not founded (7)	4
C58. Overbooking / Partly founded (7)	3
C59. Problems with ticketing with loyalty program / Well-founded or partly founded (4)	2
C60. Incorrect charging for overweight baggage / Partly founded (6)	2
C61. Preliminary judicial order granted (5)	2
C62. Waiver of ticket by consumer / Discussion about abusive fine / Repayment problems / Well-founded or partly founded (8)	1
C63. Problems with travel company and tour packages / Partly founded (2)	1
C64. Waiver of ticket by consumer / Discussion about abusive fine / Repayment problems / Well-founded or partly founded (7)	2
C65. Flight delay or cancellation / Well-founded or partly founded (8)	2
C66. Temporary baggage loss / Baggage damaged / Well-founded or partly founded (11)	4
C67. Return flight cancellation due to no-show on the outbound flight / Well-founded or partly founded (5)	0
C68. Flight change without notice / Well-founded or partly founded (16)	11
C69. Flight delay / Not proven bad weather / Well-founded or partly founded (18)	1
C70. Flight delay / Not proven bad weather / Well-founded or partly founded (24)	8
C71. Flight delay by technical problems / Well-founded or partly founded (8)	0
C72. Flight delay or cancellation by technical problems / Well-founded or partly founded (14)	0
C73. Flight delay or cancellation by airplane maintenance / Well-founded or partly founded (9)	4
C74. Flight delay or cancellation / No airline assistance or inadequate assistance / Well-founded or partly founded (20)	6
C75. Flight delay or cancellation / Airline assistance / Well-founded or partly founded (15)	1
C76. Flight delay or cancellation by airline employees' strike / Well-founded or partly founded (6)	2

APPENDIX C – CLUSTERS DESCRIPTIONS OF K-MEANS

Cluster descriptions (number of documents per cluster)	**Number of documents in disagreement with the descriptions**
C1. Different law events(243) • Flight delay / Flight cancellation / Flight change / Overbooking / Return flight cancellation due to no-show / Well-founded or partly founded • Baggage irregularities (permanent and temporary loss, damage, theft) / Well-founded or partly founded • Incorrect charging for cancellation fine / Incorrect charging for overweight baggage / Well-founded or partly founded • Under-four-hour flight delay / Not founded	34
C2. Different law events (60) • Baggage irregularities (permanent and temporary loss, damage, theft) / International flight / Cases subject to the Montreal and Warsaw Convention / Well-founded or partly founded • Flight delay / Flight cancellation / Flight change / International flight / Cases subject to the Montreal and Warsaw Convention / Well-founded or partly founded	6
C3. Different law events (275) • Flight delay / Flight cancellation / Flight change / Overbooking / Return flight cancellation due to no-show / Well-founded or partly founded • Baggage irregularities (permanent and temporary loss, damage, theft) / Well-founded or partly founded • Waiver of ticket by consumer / Discussion about abusive fine / Repayment problems / Well-founded or partly founded • Problems with ticketing / Booking not completed due to error in airline system • Promotional ticket offer not fulfilled / Well-founded or partly founded • Flight delay / Proven bad weather / Airport closures / Not founded • Late flight check-in (consumer fault) / Not founded	25
C4. Flight delay or cancellation / Judgement by assistant judge / Well-founded or partly founded (12)	0
C5. Flight delay or cancellation / International flight / Cases subject to the Montreal and Warsaw Convention / Well-founded or partly founded (31)	11
C6. Different law events (44) • Baggage irregularities (permanent and temporary loss, damage, theft) / International flight / Cases subject to the Montreal and Warsaw Convention / Well-founded or partly founded • Flight delay / Flight cancellation / Flight change / International flight / Cases subject to the Montreal and Warsaw Convention / Well-founded or partly founded	3

APPENDIX D – CLUSTERS DESCRIPTIONS OF LINGO

To make these results understandable, we organise the table as follows: the first column contains the exact descriptions generated by Lingo (kept in Portuguese); the second contains the translation or the expert's explanation of what it means; the third contains the score of each cluster, by which the table is ordered. Personal names are suppressed and replaced with "***".

Cluster descriptions in Portuguese (number of documents per cluster)	What it means in English or to which law event it refers	Lingo Score
C1. Ré *** (61)	Airline "A" as defendant	51.18
C2. Três dias (62)	The baggage took three days to be delivered to the passenger	50.48
C3. Empresas rés (66)	Two or more airlines as defendant	47.98
C4. Novos horários voos (65)	Flight change	45.06
C5. Centavos procedente (69)	Cases well-founded	44.94
C6. Realizado *** (64)	Airline "B" as defendant	41.80
C7. Reserva passagens (67)	Ticket booking	40.22
C8. *** (51)	Airline "C" as defendant	39.90
C9. Empresa Air (83)	Airline with the prefix "air" in their name as defendant	38.70
C10. Quatro dias viagem (68)	Four-day trip	38.46
C11. *** (119)	Airline "D" as defendant	37.85
C12. Entrega malas (65)	Baggage delivery	37.19
C13. Manutenção aeronave fl (88)	Airplane maintenance	36.40
C14. Período embarque (66)	Boarding time	34.88
C15. Referida alteração (70)	Flight change	34.70
C16. Cancelamento pela (67)	Flight cancellation	33.41
C17. Havia comprado passagem (115)	Ticket purchase	32.97
C18. Requerida reembolsar (87)	Repayment claim	32.55
C19. Empresa *** (90)	Airline "E" as defendant	30.98
C20. Voltou operar (64)	The airline returned to operation (flight delay)	29.87
C21. Sessão plenária (44)	Plenary session (Superior Court)	29.87
C22. Vinte cinco (68)	Twenty-five (refers to a compensation value)	29.66
C23. Dias após extravio (67)	Temporary baggage loss	29.48
C24. Peso bagagem (69)	Baggage weight	29.06
C25. Técnicos manutenção aeronave (91)	Airplane maintenance and technical problems	27.79
C26. Duas horas atraso (80)	Two-hour flight delay	27.40
C27. Aeroporto Rio de Janeiro (68)	Rio de Janeiro airport	27.27
C28. Horário embarque alterado (39)	Boarding time changed	25.67
C29. Decolagem aeronave (59)	Airplane take-off	24.76
C30. Extravio mala (102)	Baggage loss	24.73
C31. Estabelecido Convenção (124)	Montreal and Warsaw Convention rules	24.79
C32. Forneceu assistência (61)	The airline provided assistance to the consumer during the period of flight delay	24.26
C33. Seis horas atraso (99)	Six-hour flight delay	24.07
C34. Privado pertences (66)	Baggage loss on the outbound flight	23.94
C35. Precedente TJSC Recurso Inominado (68)	A State Court precedent cited in the judgment	22.80
C36. Produção outras provas além (78)	Evidence presented in the case	22.36
C37. Tanto viagem ida (70)	Outward and return flights	21.75
C38. Fls 29 (68)	Page 29 (refers to procedural file)	20.99
C39. Alteração destino final (63)	Change of final destination	20.73
C40. Sistema requerida (69)	Airline system failure	20.58

C41. Bens extraviados (71)	Baggage loss	19.94
C42. Cancelamento voo questão (75)	Flight cancellation	19.60
C43. Ida volta respectivamente (80)	Outward and return flights	18.62
C44. Cinco centavos assim (97)	Five cents (refers to a compensation value)	18.55
C45. Sentença de improcedência (66)	Cases not founded	17.91
C46. Repercussão geral (143)	A Superior Court precedent cited in the judgement	17.88
C47. Ida volta considerando (134)	Outward and return flights	16.20
C48. Produção provas audiência inicialmente (83)	Evidence presented in the case	15.48
C49. Desprovido sentença mantida (71)	Appeal dismissed and judgement upheld	13.56
C50. Fl 17 ademais (53)	Page 17 (refers to procedural file)	13.11
C51. Itens bagagem (65)	Theft of baggage items	12.82
C52. Trinta quatro (64)	Thirty-four (refers to a compensation value)	12.71
C53. Limitando se argumentar (71)	The defendant has not proven its allegations	12.35
C54. Limite peso (72)	Baggage weight limit	12.30
C55. Produção outras provas (162)	Evidence presented in the case	11.99
C56. Sete horas após (79)	Seven-hour flight delay	11.99
C57. Bagagem deixou comprovar (64)	Allegations not proven about a baggage irregularity	11.39
C58. Condições climáticas aeroporto (68)	Weather conditions	10.98
C59. Presente Convenção (71)	Montreal and Warsaw Convention rules	10.83
C60. Provimento ao recurso (60)	Appeal granted	8.67
C61. Atrasou cerca hora (43)	Flight delay	8.01
C62. *** (67)	Refers to a common name (e.g., Bob) that can be party, lawyer, judge	4.84
C63. Other topics (11)		0

APPENDIX E – CLUSTERS DESCRIPTIONS OF AFFINITY PROPAGATION

Cluster descriptions (number of documents per cluster)	Number of documents in disagreement with the descriptions
C1. Rerouting / Downgrade (change to an inferior class) / Well-founded (1)	0
C2. Return flight cancellation due to no-show on the outbound flight / Judgement by assistant judge / Partly founded (1)	0
C3. Return flight cancellation due to no-show on the outbound flight / Judgement by assistant judge / Partly founded (1)	0
C4. Permanent baggage loss / Judgement by assistant judge / Partly founded (1)	0
C5. Permanent baggage loss / Judgement by assistant judge / Partly founded (1)	0
C6. Permanent baggage loss / Judgement by assistant judge / Well-founded (1)	0
C7. Long flight delay / International flight / Cases subject to the Montreal and Warsaw Convention / Judgement by assistant judge / Partly founded (1)	0
C8. Long flight delay / Judgement by assistant judge / Well-founded or partly founded (2)	0
C9. Flight delay / Passenger stayed locked inside the airplane / Judgement by assistant judge / Well-founded (1)	0
C10. Flight cancellation / International flight / Cases subject to the Montreal and Warsaw Convention / Judgement by assistant judge / Partly founded (1)	0
C11. Temporary baggage loss / International flight / Cases subject to the Montreal and Warsaw Convention / Judgement by assistant judge / Partly founded (2)	0
C12. Flight delay or cancellation by technical problems / Partly founded (4)	0
C13. Theft of baggage items / Partly founded (3)	0
C14. Flight delay / International flight / Cases subject to the Montreal and Warsaw Convention / Well-founded or partly founded (17)	3
C15. Long delays on both outbound and return flights / International flight / Cases subject to the Montreal and Warsaw Convention / Partly founded (1)	0
C16. Flight cancellation / Passenger lost public contest exam / Partly founded (2)	0
C17. Flight delay by air traffic / Judgement by assistant judge / Partly founded (6)	0
C18. Theft of baggage items / Judgement by assistant judge / Well-founded (1)	0
C19. Not founded / Dismissed without prejudice (36)	19
C20. Flight delay / Loss of a family engagement / International flight / Cases subject to the Montreal and Warsaw Convention / Judgement by assistant judge / Well-founded (1)	0
C21. Flight delay / Loss of a family engagement / Judgement by assistant judge / Well-founded (2)	0
C22. Flight delay / Loss of a family engagement / International flight / Cases subject to the Montreal and Warsaw Convention / Judgement by assistant judge / Partly founded (1)	0
C23. Flight delay / Loss of a family engagement / Judgement by assistant judge / Partly founded (1)	0
C24. Flight delay by technical problems / Lawsuits filed by members of the same family / Well-founded (2)	0
C25. Temporary baggage loss / International flight / Cases subject to the Montreal and Warsaw Convention / Partly founded (2)	0
C26. Flight delay / Loss of a family engagement / Judgement by assistant judge / Well-founded (2)	0
C27. Overbooking / Temporary baggage loss / Judgement by assistant judge / Partly founded (1)	0
C28. Flight delay or cancellation / Well-founded or partly founded (22)	6
C29. Baggage damaged / International flight / Cases subject to the Montreal and Warsaw Convention / Partly founded (2)	0
C30. Problems with ticketing with loyalty program / Failed trip / Partly founded (1)	0

C31. Long flight delay / International flight / Cases subject to the Montreal and Warsaw Convention / Partly founded (1)	0
C32. Waiver of ticket by consumer/ Discussion about abusive fine / Partly founded (3)	0
C33. Flight delay or cancellation / International flight / Cases subject to the Montreal and Warsaw Convention / Well-founded or partly founded (23)	0
C34. Flight delay or cancellation / International flight / Cases subject to the Montreal and Warsaw Convention / Well-founded or partly founded (10)	0
C35. Right to regret / Well-founded (1)	0
C36. Change of final destination / Partly founded (3)	1
C37. Flight delay or cancellation / Bad weather not proven / International flight / Cases subject to the Montreal and Warsaw Convention / Well-founded or partly founded (10)	0
C38. Not founded (39)	16
C39. Flight delay by technical problems / Well-founded or partly founded (11)	0
C40. Flight cancellation by crew rescheduling / Partly founded (3)	1
C41. Change of departure airport / Passenger failed to board / Partly founded (2)	0
C42. Flight cancellation by airline network restructuring / Partly founded (3)	0
C43. Flight delay or cancellation / International flight / Partly founded (9)	1
C44. Temporary baggage loss/ Baggage damaged / Partly founded (7)	2
C45. Waiver of ticket by consumer/ Discussion about abusive fine / Partly founded (2)	0
C46. Temporary loss of children's baggage / International flight / Cases subject to the Montreal and Warsaw Convention / Partly founded (1)	0
C47. Temporary baggage loss/ International flight / Well-founded or partly founded (15)	4
C48. Right to regret / Well-founded (1)	0
C49. Flight delay / International flight / Cases subject to the Montreal and Warsaw Convention / Well-founded or partly founded (8)	0
C50. Flight delay or cancellation by technical problems / Well-founded or partly founded (8)	0
C51. Flight delay or cancellation / Bad weather not proven / Well-founded or partly founded (25)	5
C52. Right to regret denied / Not founded (3)	0
C53. Ticket cancellation due to suspected registration fraud / Consumer unfair denied boarding / Partly founded (1)	0
C54. Incorrect charge for ticket cancelled by consumer / Partly founded (1)	0
C55. Baggage irregularities (permanent and temporary loss, damage, stolen) / Well-founded or partly founded (21)	8
C56. Incorrect charging for overweight baggage / Judgement by assistant judge / Partly founded (1)	0
C57. Long delay on both outbound and return flights / International flight / Partly founded (1)	0
C58. Misguided passenger who missed his connecting flight / Judgement by assistant judge / Partly founded (1)	0
C59. Flight delay / No airline assistance / Judgement by assistant judge / Well-founded (2)	0
C60. Flight delay or cancellation / Overbooking / Return flight cancellation due to no-show on the outbound flight / Well-founded or partly founded (52)	11
C61. Promotional ticket offer not fulfilled / Well-founded or partly founded (6)	0
C62. Return flight cancellation due to no-show on the outbound flight / Well-founded or partly founded (7)	0
C63. Flight delay or cancellation / Well-founded or partly founded (14)	0
C64. Temporary baggage loss with medicines / Passenger with diabetes and heart disease / Partly founded (1)	0
C65. Flight delay / Flight cancellation / Flight change / Well-founded or partly founded (29)	6
C66. Permanent or temporary baggage loss / Well-founded or partly founded (21)	0
C67. Permanent baggage loss / Well-founded or partly founded (6)	1
C68. Flight delay or cancellation by airplane maintenance / Well-founded or partly founded (8)	3
C69. Temporary baggage loss/ International flight / Cases subject to the Montreal and Warsaw Convention / Partly founded (4)	0

C70. Temporary baggage loss/ Judgement by voluntary judge / Partly founded (1)	0
C71. Different law events (49) • Not founded • Tour package problems / Well-founded or partly founded • Baggage irregularities (permanent and temporary loss, damage, stolen) / Well-founded or partly founded • Waiver of ticket by consumer/ Right to regret / Discussion about abusive fine / Well-founded or partly founded • Flight delay or cancellation / Overbooking / Well-founded or partly founded	4
C72. Temporary baggage loss / Well-founded or partly founded (5)	0
C73. Flight delay / Well-founded (6)	0
C74. Ticketing error / Booking not completed / Well-founded or partly founded (16)	10
C75. Flight cancellation / Well-founded or partly founded (8)	0
C76. Temporary baggage loss / International flight/ Well-founded or partly founded (3)	0
C77. Ticket cancellation not requested by consumer / International flight / Well-founded (1)	0
C78. Flight delay / Well-founded or partly founded (24)	0
C79. Flight delay / Well-founded or partly founded (22)	0
C80. Theft of baggage items / Well-founded or partly founded (4)	0
C81. Temporary baggage loss/ Partly founded (6)	1
C82. Promotional ticket offer not fulfilled / Partly founded (1)	0
C83. Flight cancellation without notice / Partly founded (3)	0
C84. Flight delay / Consumer responsible for a group of passengers / International flight / Partly founded (1)	0
C85. Flight cancellation / Consumer had to purchase ticket from another airline / International flight / Cases subject to the Montreal and Warsaw Convention / Well-founded or partly founded (2)	0
C86. Return flight cancellation due to no-show on the outbound flight / Partly founded (2)	0
C87. Daughter of consumer prevented from boarding / Improper document requirement / Partly founded (1)	0
C88. Flight rebooking by consumer / Discussion about abusive fine / Partly founded (1)	0
C89. Flight rebooking by consumer / Discussion about abusive fine / Baggage damaged / Partly founded (1)	0
C90. Flight delay / No airline assistance / International flight / Cases subject to the Montreal and Warsaw Convention / Judgement by assistant judge/ Partly founded (1)	0
C91. Waiver of ticket by consumer/ Public tender cancelled due to truckers' strike in Brazil / Discussion about abusive fine / Judgement by assistant judge / Partly founded (1)	0
C92. Flight rebooking by consumer / Incorrect charging for overweight baggage / Judgement by assistant judge / Partly founded (1)	0
C93. Flight change without notice / International flight / Cases subject to the Montreal and Warsaw Convention / Judgement by assistant judge/ Partly founded (1)	0
C94. Flight delay or cancellation / Well-founded or partly founded (13)	0
C95. Temporary baggage loss / International flight / Cases subject to the Montreal and Warsaw Convention / Judgement by assistant judge/ Well-founded (1)	0

APPENDIX F – (JEC/UFSC) EXPLANATORY MATERIAL PRESENTED TO THE PARTIES DURING THE PARTICIPANT OBSERVATION OF CONCILIATION HEARINGS

Projeto *Conciliação Inteligente*

Juizado Especial Cível do Fórum da Universidade Federal de Santa Catarina
JEC/UFSC
Grupo de Pesquisa *Governo Eletrônico, Inclusão Digital e Sociedade do Conhecimento*
EGOV/UFSC

Tema: Transporte Aéreo - Direito do Consumidor

Objetivo

- Fornecer às partes conhecimento claro e objetivo sobre as sentenças locais, incluindo previsões de indenização por danos morais.
- Incentivar às partes, com segurança e informação baseada em dados, a alcançarem um acordo.

Como?

- Base de dados formada por +- 1.200 sentenças proferidas pelo JEC/UFSC sobre problemas vivenciados por consumidores com o serviço de transporte aéreo.
- Técnicas de Mineração de Texto e Inteligência Artificial (Aprendizado de Máquina).

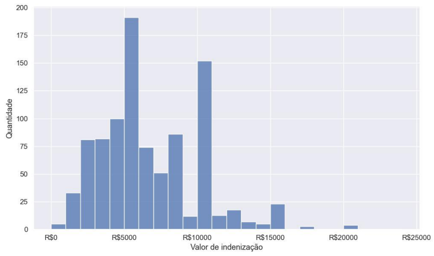
Vantagens do acordo

- Celeridade na prestação jurisdicional.
- Informalidade e simplicidade.
- Possibilidade de formular decisões diversas daquelas que venham a ser dadas pela Justiça.

Fatores de influência na decisão judicial

- **Fatores objetivos:** intervalo de espera do consumidor quando há atraso/cancelamento de voo, intervalo de espera para entrega de bagagem quando há extravio.
- **Fatores subjetivos:** assistência ou não da companhia aérea, perda de compromisso profissional, viagem com significado emocional, vulnerabilidade do consumidor (problemas de saúde, idoso, menor).

Panorama das indenizações por danos morais no JEC/UFSC



Casos extremos de indenização

- R$ 20.000,00 - R$ 25.000,00 (5 casos).
- Intervalos muito largos (acima de 24h), tanto na ida quanto na volta da viagem.
- Perda de velório e enterro de ente querido.
- Várias ocorrências de atraso/cancelamento pleiteadas em uma única ação.

Processo n. 5007767-84.2020.8.24.0090

AUTOR
x
RÉU

Audiência 13/09/2021 - 16:00

Previsões de Indenização por Danos Morais

- **Fator objetivo:** antecipação de voo / intervalo de aproximadamente 3h55min.
- **Fator subjetivo:** cancelamento de jantar de comemoração do aniversário de casamento dos autores.

Médias (em caso de procedência da ação - por autor):

- **R$ 4.000,00 - R$ 5.900,00**

Ponto controvertido (improcedência da ação ou redução na indenização):

- **Recebimento da comunicação dentro de 72h.**

APPENDIX G – (JEC/UFSC) SURVEY

Projeto "Conciliação Inteligente"

Juizado Especial Cível da UFSC e Grupo EGOV/UFSC

Tema: Transporte Aéreo (Direito do Consumidor).

Caro(s) autor(es), acionado(s) e advogado(s).

Visando avaliar as nossas sessões de conciliação realizadas no âmbito do projeto, pedimos a gentileza, se possível, de responderem o breve questionário a seguir. Seu feedback é essencial para aprimorarmos a iniciativa e avançarmos em busca de uma Justiça moderna, conciliadora e eficiente para a sociedade.

* Não estamos coletando dados pessoais no questionário, de modo que o seu preenchimento é anônimo.

*Obrigatório

1. Qual a sua posição no processo submetido à audiência? *

   Marcar apenas uma oval.

   - ◯ Parte Autora
   - ◯ Parte Acionada (Preposto/Representante)
   - ◯ Advogado(a) da Parte Autora
   - ◯ Advogado(a) da Parte Acionada
   - ◯ Outro

2. Foi realizado acordo durante a audiência? *

   Marcar apenas uma oval.

   - ◯ Sim
   - ◯ Não
   - ◯ Parcial (somente sobre parte do pedido)

3. Caso a resposta anterior tenha sido "não", qual(is) motivo(s) você atribuiria ao insucesso da conciliação?

   Marque todas que se aplicam.

   - ☐ Não comparecimento de uma ou de ambas as Partes
   - ☐ Desinteresse da Parte Autora
   - ☐ Desinteresse da Parte Acionada (ausência de proposta)
   - ☐ O valor proposto pela Parte Acionada não correspondeu às expectativas da Parte Autora

4. Caso a resposta anterior tenha sido "sim" ou "parcial", qual(is) motivo(s) você atribuiria ao sucesso da conciliação?

   Marque todas que se aplicam.

   - ☐ Interesse das Partes
   - ☐ Apresentação de informações sobre as sentenças locais e médias de indenizações por danos morais pelo conciliador(a)
   - ☐ Apresentação de proposta pela Parte Acionada
   - ☐ As Partes estavam dispostas à ceder

5. As informações apresentadas pelo conciliador(a) sobre as sentenças locais (fatores que influenciam a decisão, gráfico de indenizações, etc.) lhe foram úteis? *

   Marcar apenas uma oval.

   - Extremamente úteis
   - Muito úteis
   - Mais ou menos úteis
   - Um pouco úteis
   - Nem um pouco úteis

6. As médias de indenização por danos morais previstas para o seu processo (para fins de negociação) lhe foram úteis? *

   Marcar apenas uma oval.

   - Extremamente úteis
   - Muito úteis
   - Mais ou menos úteis
   - Um pouco úteis
   - Nem um pouco úteis

7. Avalie seu nível de satisfação com a atividade prestada durante a conciliação, de modo geral:

   Marcar apenas uma oval.

   - Muito satisfeito
   - Mais ou menos satisfeito
   - Nem satisfeito, nem insatisfeito
   - Mais ou menos insatisfeito
   - Muito insatisfeito

8. Se desejar, deixe aqui um comentário, crítica ou elogio:

ANNEX A – (JEC/UFSC) RESEARCHER'S AUTHORISATION RELATED TO THE NON-PARTICIPANT OBSERVATION OF THE CONCILIATION HEARINGS

PODER JUDICIÁRIO
de Santa Catarina

AUTORIZAÇÃO PARA A REALIZAÇÃO DE PESQUISA EM TESE DE DOUTORADO RELATIVA A AUDIÊNCIAS DE PROCESSOS DO FÓRUM DA UNIVERSIDADE FEDERAL DE SANTA CATARINA (COMARCA DA CAPITAL – NORTE DA ILHA) E TERMO DE COMPROMISSO

Por meio da presente, a MM. Juíza de Direito titular do Fórum da Universidade Federal de Santa Catarina, Dra. Vânia Petermann, autoriza a doutoranda em Direito da Universidade Federal de Santa Catarina Isabela Cristina Sabo, CPF n. 075.836.359-18, a desenvolver neste Fórum pesquisa relativa às audiências que aqui ocorrem, para compor seu levantamento de dados em sua tese de doutorado intitulada "Modelagem inteligente de audiência de conciliação ou mediação".

Na qualidade de observadora das audiências, na forma acima, a doutoranda compromete-se a manter sigilo acerca dos processos judiciais sobre os quais recaia segredo de justiça, a que venha a ter acesso por conta das atividades de pesquisa.

Florianópolis, 26 de março de 2019

Vânia Petermann
Juíza de Direito

Isabela Cristina Sabo
Doutoranda

Aires José Rover
Orientador

ANNEX B – (JEC/UFSC) RESEARCHER'S ASSIGNMENT AS CONCILIATOR RELATED TO THE PARTICIPANT OBSERVATION OF THE CONCILIATION HEARINGS

Poder Judiciário
JUSTIÇA ESTADUAL
Tribunal de Justiça do Estado de Santa Catarina
Juizado Especial Cível e Criminal da Universidade Federal de Santa Catarina

Av. Des. Vitor Lima, 183, fundos- Campus da UFSC - Bairro: Serrinha - CEP: 88040-400 - Fone: (48)3287-5019 - Email: nortedailha.juizado1@tjsc.jus.br

PROCEDIMENTO DO JUIZADO ESPECIAL CÍVEL Nº 5007767-84.2020.8.24.0090/SC

AUTOR:

AUTOR:

RÉU:

DESPACHO/DECISÃO

Avoco os autos.

Vistos, em mutirão de conciliação para ações de consumo (envolvendo Cias Aéreas).

Preocupa a este juízo o volume de ações que se encontram em andamento, e a necessidade de dar celeridade e efetividade que são do Sistema dos Juizados, expressos também na Constituição.

Durante a pandemia, além de se avolumarem as ações, esta Unidade contou com perdas de força de trabalho, e suas competências vão muito além dos Juizados Cíveis, ou seja, envolvem: Juizados Cíveis, Juizados Criminais, ações do Cível Comum, inclusive Família, além da coordenação de Centro Judiciário de Solução de Conflitos.

O nome da Unidade "Juizados" leva à ideia de que somente nos dedicamos às ações de massa, de simples solução; porém, além da competência alargada, há muito nos Juizados de complexidade, ainda mais pelos efeitos de novas situações e conflitos em tempos de pandemia.

Ademais, percebeu-se, neste tempo incomum, uma dificuldade de concretizar a grande missão dos Juizados: autocomposição.

É da Política Judiciária Nacional, editada pelo Conselho Nacional de Justiça, a conciliação e a mediação como instrumentos essenciais para o acesso à justiça, determinando aos órgãos judiciários a responsabilidade por oferecer mecanismos alternativos de solução de controvérsias como a mediação e a conciliação (Resolução n. 125/2010, e aditamentos posteriores).

Assim, considerando que:

1) a autocomposição pode dar solução a processos específicos, com atenção especial às situações concretas, enquanto se aguarda o julgamento;

2) os processos selecionados não terão prejuízo na ordem de preferência de julgamento, que será mantida; há tratamento igualitário neste primeiro mutirão do ano de 2021 pela matéria;

3) a tramitação de ações em massa de atrasos em voos, e os efeitos da pandemia que podem prolongar a discussão sobre a incidência, ou não, da lei, que alterou diversos direitos dos consumidores, como moratória em pagamentos, e dever de comprovar dano moral (ainda não analisados pela Jurisprudência de forma estável), dentre outros;

4) a disposição de conciliadora e profissional do Direito, com saberes sobre a matéria, voltada a atender a mesma Cia Área num mesmo período de dias, por concentração, e especialização;

5) a existência, acaso aceitem as partes e advogados, de recurso de Inteligência Artificial para prever a média de indenizações, e em que situações foram dadas, em ações dessa natureza (e o que será bem explicado na audiência, sempre resguardada a proteção de dados; além de ressaltar a importância das partes e dos advogados, de serem ouvidos a respeito do que vem a ser um futuro esperado das ações na justiça);

6) ao Juiz é dado instituir mutirões; e que a medida não trará prejuízo algum às partes, que, enquanto aguardam o julgamento, terão a oportunidade, via *online*, com segurança de saúde, de composições em mutirão exclusivo para a matéria; não tendo condições de fazer via *online,* nos casos de não ter assistência de advogado e não ter condição de usar meio tecnológico, comparecer ao Fórum da UFSC, onde será designada sala especial, atendendo-se às medidas sanitárias para participar do ato, **o que deverá ser comunicado, sem nova vista ao juízo, em até 10 (dez) dias antes do ato para nossa organização**, e evitar acúmulo de pessoas;

7) ainda que não desejem conhecer o processo de Inteligência Artificial, a participação no mutirão é da política de tratamento dos conflitos nesta Unidade, que se alinha às diretrizes do Conselho Nacional de Justiça. Recordo que, ao eleger os Juizados, as partes concordam que o Sistema tem como base participar de aproximações para a solução alternativa do conflito, submetendo-se às audiências de conciliação, salvo impedimentos legais;

8) eventual audiência anterior não impede marcar mutirões, o que é, aliás, praxe anual dos Tribunais; que ora se antecipa pela situação contextualizada nesta deliberação;

9) a ausência importará à parte autora desistência, e da ré, revelia, salvo os aspectos de ordem pública;

10) para os processos sem contestação ou réplica, o lapso para tanto começará a contar do dia útil seguinte do ato, salvo se as partes concordarem de modo diferente na audiência;

11) observo que não serão canceladas audiências ao argumento de não interesse em compor, do que este juízo deixa claro o quanto é importante - além de dever de quem escolhe os Juizados, como outrora lembrado - de estar presente na oportunidade de autocomposição, e que a cooperação e a compreensão dos envolvidos é muito importante para a Justiça, aos quais, desde já, registra-se, sobretudo, o respeito e agradecimento deste juízo;

12) visando a dar efetividade às audiências, com a reserva de dados, a conciliadora poderá entrar em contato prévio com as partes e/ou advogados para esclarecer a importância do ato (esta medida se chama pré-conciliação e, na prática deste juízo, mostrou-se valorosa para o ato em si, dando a atenção aos Jurisdicionados e esclarecendo aspectos do mutirão)

Diante do exposto, por cooperação que se observa às partes e advogados, considerando os itens acima, encaminho os autos à conciliação *online,* sem prejuízo na data de conclusão, ou seja, a ordem de antiguidade será respeitada, não havendo acordo.

Intimem-se as partes e procuradores para fornecerem seus endereços de e-mail para remessa do *link* da audiência, no prazo de 5 (cinco) dias, cientes de que, no silêncio, o *link* será encaminhado aos e-mails que já constam nos autos.

Esclareço, por fim, que as audiências serão realizadas por duplas, compostas em revezamento por um membro da equipe, e sempre com a presença da conciliadora designada e destacada para conduzir exclusivamente o mutirão, a Doutoranda em Direito pela UFSC, Isabela Sabo. Informe-se o TSI atuante no Fórum para auxílio tecnológico.

Documento eletrônico assinado por **VANIA PETERMANN, Juíza de Direito**, na forma do artigo 1º, inciso III, da Lei 11.419, de 19 de dezembro de 2006. A conferência da **autenticidade do documento** está disponível no endereço eletrônico https://eproc1g.tjsc.jus.br/eproc/externo_controlador.php?acao=consulta_autenticidade_documentos, mediante o preenchimento do código verificador **310017539546v1** e do código CRC **b00611ac**.

Informações adicionais da assinatura:
Signatário (a): VANIA PETERMANN
Data e Hora: 6/8/2021, às 12:58:16