

UTILIZAÇÃO DE REDES NEURAIIS PARA PREVISÃO DA DURAÇÃO DE
TRAMITAÇÃO DE PROCESSOS JUDICIAIS NO SUPERIOR TRIBUNAL DE JUSTIÇA
(STJ)

Djacy Lopes do Nascimento Júnior

Brasília, DF

2025



UTILIZAÇÃO DE REDES NEURAIS PARA PREVISÃO DA DURAÇÃO DE
TRAMITAÇÃO DE PROCESSOS JUDICIAIS NO SUPERIOR TRIBUNAL DE JUSTIÇA
(STJ)

Djacy Lopes do Nascimento Júnior

Trabalho de Conclusão de Curso apresentado ao Programa de Pós-Graduação em Governança e Inovação em Políticas Públicas (PPG-GIPP), da Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas, como requisito para obtenção do título de Mestre em Governança e Inovação em Políticas Públicas.

Orientador: Sérgio Ricardo Miranda Nazaré

Brasília, DF

2025



Djacy Lopes do Nascimento Júnior

**Utilização de Redes Neurais para Previsão da Duração de Tramitação de Processos
Judiciais no Superior Tribunal de Justiça (STJ)**

Trabalho de Conclusão de Curso apresentado ao Programa de Pós-Graduação em Governança e Inovação em Políticas Públicas (PPG-GIPP), da Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas, como requisito para obtenção do título de Mestre em Governança e Inovação em Políticas Públicas.

Data da defesa: 26/03/2025

Comissão Examinadora:

Professor Doutor Sérgio Ricardo Miranda Nazaré
Orientador
PPG-GIPP/UnB

Professor Doutor Eduardo Tadeu Vieira
Examinador Interno
PPG-GIPP/UnB

Doutor Montgomery Wellington Muniz
Examinador Externo
Doutor em Direito Constitucional
Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (IDP)



RESUMO

A morosidade na tramitação dos processos judiciais representa um desafio significativo para o Poder Judiciário brasileiro, impactando negativamente a eficiência institucional e a percepção pública sobre a justiça. Este estudo teve por objetivo desenvolver e validar um modelo preditivo baseado em redes neurais artificiais, especificamente *Multi-Layer Perceptron* (MLP), visando aprimorar a precisão na previsão do tempo integral de tramitação dos processos judiciais no Superior Tribunal de Justiça (STJ). Utilizou-se uma metodologia quantitativa exploratória e explicativa, com destaque para técnicas avançadas de aprendizado de máquina e otimização de hiperparâmetros por *RandomizedSearchCV*. A base de dados analisada foi extraída do Painel Justiça em Números, disponibilizado pelo Conselho Nacional de Justiça (CNJ), abrangendo mais de 600 mil processos baixados pelo STJ entre janeiro de 2023 e junho de 2024. O algoritmo apresentou resultados significativos, com coeficiente de determinação (R^2) de 0,9243 e erro absoluto médio (MAE) de 71 dias. Adicionalmente, a análise de *SHapley Additive exPlanations* (SHAP Values) identificou o "Órgão Julgador" como a variável mais influente na previsão da duração dos processos, demonstrando sua relevância no contexto estudado. Como contribuição prática, o modelo desenvolvido pode ser integrado ao sistema de gestão processual do STJ, otimizando a alocação estratégica de recursos e proporcionando intervenções antecipadas em casos potencialmente morosos. Este estudo amplia o campo acadêmico sobre aplicações de inteligência artificial no Judiciário.

Palavras-chave: Redes Neurais Artificiais; Previsão Processual; SHAP Values; Aprendizado de Máquina; Inteligência Artificial; STJ.



ABSTRACT

The slowness in judicial proceedings represents a critical challenge for the Brazilian Judiciary, negatively impacting operational efficiency and public perception of justice. This study aimed to develop and validate a predictive model using artificial neural networks, specifically the Multi-Layer Perceptron (MLP), to enhance the accuracy in predicting the overall processing duration of judicial cases at the Superior Tribunal de Justiça (STJ). A quantitative exploratory and explanatory methodology was applied, emphasizing advanced machine learning techniques, including RandomizedSearchCV for hyperparameter optimization. The dataset comprised over 600,000 judicial cases completed by the STJ from January 2023 to June 2024, extracted from the Conselho Nacional de Justiça (CNJ) "Justice in Numbers" panel. The developed algorithm demonstrated high predictive performance, achieving a determination coefficient (R^2) of 0.9243 and a Mean Absolute Error (MAE) of 71 days. Additionally, the SHapley Additive exPlanations (SHAP Values) analysis identified "Judging Body" (Órgão Julgador) as the most influential variable affecting case duration predictions. Practically, this model can be integrated into the STJ's case management system, enabling strategic resource optimization and proactive management of lengthy judicial processes. Academically, this research contributes to expanding knowledge on artificial intelligence applications in the Judiciary.

Keywords: Artificial Neural Networks; Judicial Prediction; SHAP Values; Machine Learning; Artificial Intelligence; STJ.

LISTA DE FIGURAS

Figura 2.1 - BOXPLOT Distribuição da Duração dos Processos Judiciais	311
Figura 2.2 - Gráfico de Dispersão dos Resultados	35
Figura 2.3 - Gráfico de Distribuição de Erros	366
Figura 2.4 - Desempenho do Modelo por Intervalo de Duração dos Processos Judiciais.....	37
Figura 2.5 - Gráfico SHapley Additive exPlanations (SHAP)	38
Figura 2.6 - Evolução do Impacto das Variáveis ao Longo do Tempo	39
Figura 3.1 - Preparação do Ambiente de Produção	53
Figura 3.2 - Capacidade de Processamento	54
Figura 3.3 - Estrutura dos Dados	55
Figura 3.4 - Expansão dos Códigos de Assuntos e de Classes	57
Figura 3.5 - Variáveis de Sazonalidade	57
Figura 3.6 - Criação das Colunas “Código órgão_Procedimento” e “Frequência_Tribunal”	58
Figura 3.7 - Definição das Variáveis X e Y	58
Figura 3.8 - Escolha e Divisão da Amostra	59
Figura 3.9 - Definição do Pipeline.....	59
Figura 3.10 - Pré-Processamento.....	60
Figura 3.11 - Criação do Pipeline	61
Figura 3.12 - Parametrização do RandomizedSearchCV	64
Figura 3.13 - Treinamento do Modelo com os Hiperparâmetros mais precisos.....	64
Figura 3.14 - Avaliação do Modelo.....	65
Figura 3.15 - BOXPLOT Distribuição da Duração dos Processos Judiciais.....	66
Figura 3.16 - Gráfico de Dispersão dos Resultados	70
Figura 3.17 - Gráfico de Distribuição de Erros	71
Figura 3.18 - Desempenho do Modelo por Intervalo de Duração dos Processos Judiciais.....	72
Figura 3.19 - Gráfico SHapley Additive exPlanations (SHAP)	75
Figura 3.20 - Evolução do Impacto das Variáveis ao Longo do Tempo	76
Figura 3.21 - Levantamento Bibliométrico	79
Figura 3.22 - Publicações por Área do Conhecimento	80
Figura 3.23 - Diagrama Funcionamento do Modelo MLP	84
Figura 3.24 - MLP de uma camada oculta.....	86
Figura 3.25 - Série Histórica da Despesa com Informática e com Capital.....	90
Figura 3.26 - Série Histórica dos Casos Novos e Processos Baixados.....	90
Figura 3.27 - Série histórica dos casos pendentes	91



Figura 3.28 - Tempo de Giro do Acervo por Tribunal	92
Figura 3.29 - Série Histórica dos Casos Pendentes	93
Figura 3.30 - Série Histórica dos Casos Novos e Baixados	93



LISTA DE TABELAS

Tabela 2.1 - Desempenho das Combinações de Hiperparâmetros.....	32
Tabela 3.1 – Dados Categóricos Originais.....	60
Tabela 3.2 – Aplicação do One-Hot Encoding.....	60
Tabela 3.3 – Definição de Hiperparâmetros do MLPRegressor.....	61
Tabela 3.4 – Seleção de Hiperparâmetros do MLPRegressor com RandomizedSearchCV.....	63
Tabela 3.5 – Validação do Desempenho do Modelo.....	64
Tabela 3.6 – Desempenho das Combinações de Hiperparâmetros.....	67
Tabela 3.7 – 50 Melhores Resultados	72
Tabela 3.8 – 50 Piores Resultados	74
Tabela 3.9 – Eixos Temáticos	80



LISTA DE ABREVIATURAS E SIGLAS

API RESTful – Application Programming Interface Representational State Transfer
CNJ – Conselho Nacional de Justiça
CSV – Comma-Separated Values
DF – Distrito Federal
IA – Inteligência Artificial
IQR – Intervalo Interquartil
LIODS – Laboratório de Inovação, Inteligência e ODS
MAE – Erro Absoluto Médio
MLP – Multi-Layer Perceptron
MLPRegressor – Multi-Layer Perceptron Regressor
ONU – Organização das Nações Unidas
ODS – Objetivos de Desenvolvimento Sustentável
PJe – Processo Judicial Eletrônico
PNUD – Programa das Nações Unidas para o Desenvolvimento
PPG-GIPP – Programa de Pós-Graduação em Governança e Inovação em Políticas Públicas
PPT – Produto Técnico-Tecnológico
SHAP – SHapley Additive exPlanations
STJ – Superior Tribunal de Justiça
VLSI – Very-Large-Scale Integration

SUMÁRIO

1. INTRODUÇÃO	12
2. ARTIGO: PREVISÃO DE TRAMITAÇÃO JUDICIAL NO STJ - MODELO COM REDES NEURAIIS.....	18
2.1. Introdução	18
2.2. Quadro Teórico Conceitual	20
2.2.1 Fundamentos das Redes Neurais e Deep Learning	20
2.2.2 Inovação e Inteligência Artificial na Gestão Processual	23
2.3. Metodologia e Técnicas	26
2.4. Resultados e Discussão	29
2.4.1 Resultados	30
2.4.2 Discussão dos Resultados	39
2.5. Conclusão	42
3. PTT: MODELO PREDITIVO PARA PRAZOS PROCESSUAIS NO SUPERIOR TRIBUNAL DE JUSTIÇA.....	48
3.1. Introdução	48
3.2. Descrição geral do produto	49
3.2.1 Metodologia de Pesquisa	49
3.2.2 Coleta de Dados.....	51
3.2.3 Preparação do ambiente de produção.....	52
3.2.4 Carregamento, limpeza e preparação dos dados.....	54
3.2.5 Pré-processamento dos Dados.....	59
3.2.6 Criação do Pipeline e Treinamento do Modelo.....	61
3.2.7 Ajuste de Hiperparâmetros do Modelo MLP	61
3.2.8 Avaliação do Modelo	64
3.2.9 Resultados	65
3.3. Limitações da Pesquisa	76
3.4. Produto Técnico Tecnológico	77
3.5. Base teórica utilizada.....	78
3.5.1 Eixo 1: Inteligência Artificial	81
3.5.2 Eixo 2: Dados do Poder Judiciário Brasileiro	89
3.5.3 Eixo 3: Inovação.....	94
3.5.4 Eixo 4: Utilização de soluções de IA pelo Sistema Jurídico.....	98
3.6. Relevância do Produto	100
3.6.1 Complexidade e Aderência.....	100
3.6.2 Potencial inovador.....	102
3.6.3 Aplicabilidade	102

3.6.4	Impacto Potencial.....	102
3.6.5	Documentos comprobatórios e evidências.....	103
REFERÊNCIAS.....		105

1. INTRODUÇÃO

Remodelado pela Constituição Federal de 1988, o Poder Judiciário brasileiro abrange uma diversidade de órgãos que operam em vários níveis governamentais - União, Estados, Distrito Federal e Territórios. Esse período marcou o início de uma série de reformas administrativas, com ênfase na transparência e na busca por resultados, em contraste com o modelo anterior, centrado em processos burocráticos (BRASIL, 1988).

Nesse cenário de mudança, surgiram novas abordagens organizacionais, inspiradas em modelos orgânicos, que consideram a instituição como um ser vivo, interagindo dinamicamente com seu ambiente. Essas abordagens enfrentaram desafios de inovação e ruptura do *status quo*, que reforçam a necessidade de uma evolução contínua. A título de exemplificação, pode-se mencionar a metáfora das organizações como organismos vivos — proposta por Morgan (2006, p. 53) — que destaca a importância da adaptação ao ambiente como condição para a sobrevivência e o desenvolvimento organizacional.

Essa lógica adaptativa encontra-se também no setor público, no qual a articulação harmoniosa entre os diversos órgãos e setores da administração torna-se essencial para o sucesso das organizações estatais e para o desenvolvimento de políticas públicas mais eficazes, refletindo a dinâmica e as exigências do contexto atual (CAVALCANTE; CAMÕES, 2017).

Em decorrência disso e alinhando-se aos esforços globais para um futuro sustentável, o Brasil, como signatário da Agenda 2030 da Organização das Nações Unidas (ONU), adotou os 17 Objetivos de Desenvolvimento Sustentável (ODS) para erradicar a pobreza, proteger o planeta e assegurar a paz e a prosperidade (ONU, 2015).

Essa orientação global reflete-se no Poder Judiciário brasileiro, particularmente através da criação do Laboratório de Inovação, Inteligência e Objetivos de Desenvolvimento Sustentável (LIODS) pelo Conselho Nacional de Justiça (CNJ) em 2019. Em consonância com essa iniciativa, no XIII Encontro Nacional do Poder Judiciário em 2019, foi estabelecida a **Meta 9**, visando integrar os ODS na estratégia do Poder Judiciário, com um **foco especial na prevenção de litígios e na promoção da sustentabilidade** (CNJ, 2020).

Dessa integração do Poder Judiciário aos ODS, deu-se espaço para a adoção de novas tecnologias como inteligência artificial, internet das coisas e computação em nuvem, que utilizam soluções algorítmicas e têm aprimorado significativamente as capacidades de captura, armazenamento e processamento de dados. Esse avanço possibilitou um novo nível de visibilidade e controle sobre aspectos variados da vida cotidiana, tornando-os mais mensuráveis.

O modo como tais tecnologias são implementadas pela Administração Pública faz com que essa nova estratégia venha acompanhada de uma série de automações nas decisões e no uso de novas tecnologias. Dessa forma, essa mudança não é uma simples incorporação de técnicas mais modernas, e sim uma aceitação tácita de atualização, que traga mais eficiência, neutralidade e velocidade oferecidas pela correlação algorítmica de dados e aplicação de seus métodos matemáticos.

Essa racionalização de processos e a análise de grandes volumes de dados contribuem para identificação de padrões de demandas e permitem uma atuação preventiva na origem dos conflitos e, por conseguinte, fomenta práticas mais sustentáveis no uso de recursos e na gestão administrativa.

Isso possibilitou um olhar mais atento aos altos índices de litigiosidade do sistema judiciário brasileiro. Segundo o relatório Justiça em Números, do CNJ, “o Poder Judiciário finalizou o ano de 2022 com 81,4 milhões de processos em tramitação, aguardando alguma solução definitiva. Desses, 17,7 milhões, ou seja, 21,7%, estavam suspensos, sobrestados ou em arquivo provisório, aguardando alguma situação jurídica futura. Dessa forma, desconsiderados tais processos, tem-se que, em andamento, ao final do ano de 2022, existiam 63 milhões de ações judiciais” (CNJ, 2023, p. 92).

Em complementação a esse cenário, o indicador de “tempo de giro do acervo” representa o tempo para zerar o acervo de processos, mesmo que não houvesse ingresso de novas demandas e fosse mantida a produtividade dos magistrados e dos servidores. O dado evidencia o impacto do acúmulo processual na capacidade de resolução do judiciário. Nos Tribunais Superiores, ainda conforme o relatório, seriam necessários aproximadamente 1 ano e 2 meses, na Justiça Estadual, o resultado é de 2 anos e 11 meses; na Justiça Federal, é de 2 anos e 11 meses; na Justiça do Trabalho, é de 1 ano e 7 meses e na Justiça Militar Estadual, é de 1 ano (SALOMÃO *et al.* 2023).

A conjuntura apresentada torna imperativo buscar soluções inovadoras para gerenciar e otimizar o sistema judiciário, considerando que os produtos dos órgãos jurisdicionais não servem para satisfazer desejos e necessidades individuais, mas sociais.

Assim, no contexto das instituições públicas, a inovação tem sido uma das principais estratégias para melhorar a eficiência e eficácia dos serviços prestados e, conforme abordado por Cavalcante e Camões (2017), é necessário considerar uma série de fatores de influência, como: ambientais (cultura organizacional, legislação e pressão da sociedade), organizacionais (estrutura e processos internos da organização, como a liderança, a comunicação e a capacidade de aprendizado), características de inovação e níveis dos indivíduos/funcionários.

Por isso, a inovação pública não se trata apenas de novas tecnologias ou produtos, mas também de novas formas de gestão e organização do setor público, sendo positivamente influenciada pela capacidade administrativa e aprendizagem organizacional. A geração de inovação é tipicamente mais lenta do que a sua adoção, ao mesmo tempo em que a adoção tende a ser mais fácil que a geração (CAVALCANTE; CAMÕES, 2017).

Consequentemente, a inteligência na gestão pública, aqui compreendida como uma inovação que envolve o uso de tecnologia para melhorar a tomada de decisão e planejar atividades públicas, possibilita visualizar a multidimensionalidade e a complexidade das soluções tecnológicas que podem ser empregadas no setor público (MELATI; JANISSEK-MUNIZ, 2022).

Isso acarreta a otimização de recursos, a melhoria da tomada de decisões e a aceleração de processos, sendo importante ressaltar a existência de barreiras de transição para um governo inteligente, as quais podem ser mais institucionais do que tecnológicas (MELATI; JANISSEK-MUNIZ, 2022). Já Castro e Guimarães (2019) abordaram as dimensões institucionais da inovação no Judiciário, o que possibilitar entender que a adoção de IA depende de fatores como liderança, cooperação interinstitucional e cultura organizacional.

Atualmente, o CNJ utiliza indicadores estatísticos consolidados, como o tempo médio de tramitação, a taxa de congestionamento e o índice de atendimento à demanda,

para fins de monitoramento e avaliação do desempenho dos tribunais (CONSELHO NACIONAL DE JUSTIÇA, 2024).

Embora úteis para o acompanhamento da atividade judicial, tais indicadores são descritivos e não oferecem capacidade preditiva individualizada. Assim, este estudo propõe o uso de redes neurais, com base no aprendizado profundo, como uma alternativa metodológica voltada à antecipação mais precisa da duração dos processos judiciais no STJ, contribuindo para o aperfeiçoamento da gestão do tempo processual.

Nesse contexto, a presente pesquisa busca responder à seguinte questão: **de que maneira as redes neurais, especialmente por meio do *Deep Learning*, podem ser aplicadas à previsão da duração da tramitação de processos judiciais no Superior Tribunal de Justiça (STJ)?**

As metodologias tradicionais, baseadas em médias históricas, apresentam limitações em sua capacidade de generalização, o que pode levar a estimativas imprecisas e prejudicar a gestão eficiente dos tribunais. Com o intuito de responder a essa questão, este trabalho teve como **objetivo geral** desenvolver um modelo preditivo baseado em redes neurais, por meio de técnicas de aprendizado profundo (*Deep Learning*), com o objetivo de estimar a duração da tramitação de processos judiciais no Superior Tribunal de Justiça (STJ), em contraste com os métodos estatísticos atualmente utilizados, com os seguintes **objetivos específicos**:

- I. Analisar os padrões históricos de tramitação de processos no STJ para identificar variáveis-chave que influenciam a duração dos processos; e
- II. Desenvolver um algoritmo de rede neural que incorpore essas variáveis para prever a duração dos processos.

Em resposta a esses desafios, esta pesquisa apresenta o desenvolvimento de um **Produto Técnico-Tecnológico (PPT)**, que consiste num modelo baseado em redes neurais com o objetivo de melhorar a previsão da duração dos processos judiciais no STJ. Esse produto foi construído utilizando técnicas de *deep learning* para analisar dados históricos de tramitação de processos, extraídos do painel Justiça em Números do CNJ.

No campo teórico, a pesquisa é fundamentada em estudos recentes que exploram

a aplicação de redes neurais no setor público, enfatizando a capacidade desses algoritmos de modelar relações complexas e de identificar padrões latentes em grandes volumes de dados. Além disso, a revisão da literatura destaca o impacto da inteligência artificial na modernização das operações e na tomada de decisões, oferecendo uma visão geral de como essas tecnologias podem transformar a atuação das instituições públicas (OLWONOMI et al., 2021; KAPLAN e HAENLEIN, 2019; ALMEIDA NETO et al., 2018; FUX et al., 2022; LACERDA; D'AMORIM; ALVIM, 2021; MAIA FILHO; JUNQUILHO, 2018; CAVALCANTE & CAMÕES, 2017; MELATI, JANISSEK-MUNIZ, 2022).

A estrutura metodológica, baseada em abordagens quantitativas exploratórias e explicativas, utilizando ferramentas de aprendizado de máquina para tratar desafios reais enfrentados pelo sistema judiciário, alinha-se aos recentes estudos como SABO (2020), MACOHIN (2019), LACERDA; D'AMORIM; ALVIM (2021), FUX et al. (2022). O uso de redes neurais, em particular, destaca-se por sua capacidade de capturar padrões complexos que métodos tradicionais dificilmente conseguiriam abordar.

Essa abordagem permite traduzir dados históricos em previsões úteis, contribuindo diretamente para a gestão de prazos e a alocação de recursos no Superior Tribunal de Justiça (STJ), corroborando com o caráter empírico da pesquisa em um Mestrado Profissional, que possui uma natureza mais exógena, e estabelece, portanto, maiores relações extramuros, seguindo a lógica da produção do conhecimento compartilhado, aplicado a problemas reais (GIACOMAZZO; LEITE, 2014).

O conjunto de dados utilizado contém informações sobre mais de 600.000 (seiscentos mil) processos baixados no STJ, coletados no mês de agosto de 2024, abrangendo os processos com *status* de “baixados”, entre os meses de janeiro de 2023 e junho de 2024, contendo variáveis como datas de início e término, códigos de assunto e classe processual, tribunal responsável e duração da tramitação.

Para a análise, os dados foram tratados para remover inconsistências, como valores ausentes e duração negativa de processos, e em seguida foram transformados em variáveis adequadas para serem processadas por algoritmos de aprendizado de máquina.

A construção do modelo preditivo utilizou redes neurais do tipo *Multi-Layer*

Perceptron (MLP), na linguagem de programação *Python* e bibliotecas especializadas, como a do *scikit-learn*, com ajuste de hiperparâmetros via *RandomizedSearchCV*, a fim de maximizar a precisão do modelo. O desempenho do modelo foi avaliado por meio de métricas como o coeficiente de determinação (R^2) e o erro absoluto médio (*MAE*), aplicando técnicas de validação cruzada para garantir a robustez e generalização dos resultados, cujo resultado foi um **R^2 de 0,9243 e um MAE de 71 dias**.

Uma aplicação prática do modelo inclui sua **integração aos sistemas de gestão processual** do STJ, permitindo que os magistrados e gestores tenham acesso a estimativas de tempo mais precisas, o que pode contribuir para uma alocação mais eficiente dos recursos e para a tomada de decisões mais assertivas.

Sobre esse enfoque, a inteligência na gestão pública, considerada como uma inovação que envolve o uso de tecnologia para apoiar e aprimorar a tomada de decisão (MELATI; JANISSEK-MUNIZ, 2022) e a explosão de informações digitais em toda a sociedade, possibilitam o surgimento de um governo mais eficiente, transparente e eficaz (GIL-GARCIA; DAWES; PARDO, 2018).

O uso desse tipo de tecnologia para melhorar a eficiência da tramitação processual oferece um novo caminho para o desenvolvimento de ferramentas tecnológicas de apoio à tomada de decisão no setor público (CAVALCANTE; CAMÕES, 2017; SCHWAB, 2016). O presente trabalho propõe uma abordagem inovadora que pode impactar diretamente a eficiência operacional do STJ, com alto potencial de replicabilidade aos demais órgãos do Poder Judiciário, considerando a **flexibilidade e adaptabilidade do treinamento empregado para treinamento da rede neural**, o qual pode ser ajustado de acordo com as particularidades dos dados e dos sistemas de gestão processual de cada tribunal.

Assim, esta pesquisa contribui para a aplicação de **inteligência artificial na previsão da tramitação processual no STJ**, um campo ainda pouco explorado na literatura. Diferente de estudos que analisaram tribunais regionais (Oliveira et al., 2022; Silva et al., 2023), este estudo foca em um tribunal superior, identificando **quais fatores impactam a tramitação dos processos**. Além disso, a pesquisa adota **técnicas de validação estatística** para demonstrar o grau de acurácia das previsões.

2. ARTIGO: PREVISÃO DE TRAMITAÇÃO JUDICIAL NO STJ - MODELO COM REDES NEURAIAS

2.1. Introdução

O Poder Judiciário brasileiro passou por uma série de transformações a partir da promulgação da Constituição Federal de 1988, através de reformas administrativas com ênfase na transparência, na eficiência e na busca por resultados. Surgiram novas abordagens organizacionais baseadas em modelos orgânicos, onde as instituições são concebidas como sistemas dinâmicos que interagem com seu ambiente, enfrentando desafios de inovação (CAVALCANTE; CAMÕES, 2017).

Como decorrência natural desse processo de transformação, em linha com os esforços globais por um futuro sustentável, o Brasil adotou os 17 Objetivos de Desenvolvimento Sustentável (ODS) da Agenda 2030 da ONU, o que se refletiu no Poder Judiciário com a criação do Laboratório de Inovação, Inteligência e ODS (LIODS) pelo Conselho Nacional de Justiça em 2019.

A partir dessa iniciativa, no XIII Encontro Nacional do Poder Judiciário em 2019, foi estabelecida a Meta 9, visando integrar os ODS na estratégia do Poder Judiciário, com um foco especial na prevenção de litígios e na promoção da sustentabilidade, o que exige a adoção de práticas para melhorar a eficácia da prestação jurisdicional.

Essa iniciativa tem impulsionado a adoção de novas tecnologias – como inteligência artificial, internet das coisas e computação em nuvem – que aprimoram significativamente as capacidades de captura, armazenamento e processamento de dados judiciais, oferecendo maior visibilidade e controle sobre os processos. Entretanto, com relação a previsão da duração dos processos judiciais, fundamental para a alocação eficiente de recursos e para a melhoria da celeridade processual, a gestão processual ainda se apoia majoritariamente em métodos tradicionais baseados em médias históricas.

Dados do painel Justiça em Números do CNJ indicam que, em 2022, 63 milhões de ações judiciais permaneciam em tramitação e que o “tempo de giro do acervo” variava significativamente entre as instâncias – desde aproximadamente 1 ano nos Tribunais Superiores até quase 3 anos na Justiça Estadual (SALOMÃO et al., 2023).

Neste cenário, a aplicação de técnicas de inteligência artificial, em especial das redes neurais do tipo *Multi-Layer Perceptron (MLP)*, mostra-se promissora para a previsão da duração dos processos judiciais. Estudos recentes apontam que algoritmos de *deep learning* e soluções de inteligência artificial têm a capacidade de modelar relações complexas e identificar padrões latentes em grandes volumes de dados (OLOWONOMI et al., 2021; KAPLAN e HAENLEIN, 2019; ALMEIDA NETO et al., 2018; FUX et al., 2022; LACERDA; D'AMORIM; ALVIM, 2021; MAIA FILHO; JUNQUILHO, 2018).

Assim, o presente estudo propõe o desenvolvimento e a validação de um modelo preditivo baseado em MLP, implementado em *python* com o uso de bibliotecas especializadas, como a *scikit-learn*, e ajustado por meio do *RandomizedSearchCV* para maximizar sua precisão.

O algoritmo proposto foi avaliado utilizando métricas estatísticas, alcançando um coeficiente de determinação (R^2) de 0,9243 e um erro absoluto médio (MAE) de 71 dias. Tais resultados demonstram a possibilidade da adoção de uma abordagem inovadora para a previsão da tramitação processual em um tribunal superior – uma área ainda pouco explorada na literatura (OLIVEIRA et al., 2022; SILVA et al., 2023).

Portanto, esta pesquisa contribui para o avanço do conhecimento ao explorar, de forma inovadora, a aplicação de redes neurais para a previsão da duração dos processos judiciais, propondo uma ferramenta com potencial futuro de acelerar a prestação jurisdicional, reduzir custos operacionais e atender, de maneira mais célere, às demandas sociais por uma justiça eficaz.

Diante deste cenário, a questão de pesquisa que orientou o presente estudo é: **"De que maneira as redes neurais, especialmente por meio de técnicas de aprendizado profundo (Deep Learning), podem ser aplicadas à previsão da duração da tramitação de processos judiciais no Superior Tribunal de Justiça (STJ)? "**

Para responder a essa pergunta, o objetivo geral desta pesquisa foi desenvolver e validar um modelo preditivo baseado em redes neurais artificiais, especificamente do tipo *Multi-Layer Perceptron (MLP)*, utilizando técnicas de aprendizado profundo (*Deep Learning*) e ajuste de hiperparâmetros com *RandomizedSearchCV*, com o propósito de estimar com maior exatidão a duração da tramitação de processos judiciais no STJ.

A escolha pelo MLP se justifica por sua capacidade de capturar relações complexas e não lineares nos dados processuais, além de permitir ajustes estruturais e de desempenho por meio de otimização de hiperparâmetros. Embora outros algoritmos supervisionados, como *Random Forest* ou *XGBoost*, possam apresentar desempenho comparável, optou-se por delimitar o estudo ao uso de redes neurais, foco central da proposta.

A relevância deste estudo reside em diversos aspectos. Primeiramente, ao superar as limitações dos métodos tradicionais baseados em médias históricas, o modelo preditivo proposto oferece uma ferramenta inovadora para a administração judiciária. Em segundo lugar, a precisão na previsão dos prazos processuais pode auxiliar na identificação precoce de processos que demandem intervenções preventivas, promovendo maior celeridade e eficácia na prestação jurisdicional.

Dessa forma, o presente artigo está estruturado da seguinte forma: após esta Introdução, que contextualiza o tema e apresenta a problemática estudada, segue o quadro teórico conceitual, que aborda os fundamentos teóricos sobre redes neurais, *Deep Learning* e inovação tecnológica no setor público. Posteriormente, detalha-se a metodologia adotada, descrevendo os procedimentos para coleta e tratamento dos dados, treinamento e validação do modelo preditivo e técnicas estatísticas utilizadas. Em seguida, os resultados obtidos são apresentados e discutidos à luz da literatura acadêmica pertinente. Por fim, o estudo encerra-se com as considerações finais, que destacam as contribuições acadêmicas e práticas, limitações identificadas e sugestões para pesquisas futuras.

2.2. Quadro Teórico Conceitual

2.2.1 Fundamentos das Redes Neurais e Deep Learning

2.2.1.1. Conceito de Redes Neurais

As redes neurais artificiais são sistemas computacionais inspirados no funcionamento do cérebro humano, cuja origem remonta aos estudos pioneiros de McCulloch e Pitts (1943), que propuseram o primeiro modelo matemático de um neurônio. Esse modelo, embora simplificado, estabeleceu as bases para a construção de sistemas capazes de processar informações de maneira semelhante aos processos biológicos. Segundo os autores, “*the activity of any neuron may be represented as a proposition*” (McCULLOCH; PITTS, 1943, p. 100), o que significa que cada disparo

neuronal pode ser interpretado como uma operação lógica, fundamentando a ideia de que o cérebro realiza computações.

Posteriormente, o desenvolvimento do Perceptron por Rosenblatt (1958) introduziu um modelo de aprendizado baseado em reforço, permitindo que os parâmetros – os pesos das conexões – fossem ajustados de acordo com estímulos reforçadores. Essa abordagem possibilitou a classificação de padrões por meio de um treinamento iterativo, marcando um avanço significativo na área. Segundo o autor, "the perceptron is expected to learn, or to form associations, as a result of the changes in value that occur as a result of the activity of the association cells" [espera-se que o perceptron aprenda, ou forme associações, como resultado das mudanças de valor que ocorrem em função da atividade das células de associação] (ROSENBLATT, 1958, p. 395).

A evolução natural desses modelos culminou no surgimento dos *Multilayer Perceptrons (MLPs)*, que incorporam uma ou mais camadas ocultas de neurônios entre a camada de entrada e a de saída. Essa estrutura em camadas permite que os MLPs capturem relações não lineares e complexas, tornando-os eficazes na análise de dados com alta dimensionalidade e variabilidade. O treinamento desses modelos é realizado por meio do algoritmo de retropropagação, que ajusta iterativamente os pesos com o objetivo de minimizar o erro de previsão, conforme demonstrado por Haykin (2017).

Com o avanço das técnicas de *deep learning*, tornou-se possível implementar modelos com maior profundidade e capacidade de generalização, ampliando as aplicações das redes neurais para problemas complexos e de grande escala, como a previsão da duração dos processos judiciais.

2.2.1.2. Multi-Layer Perceptron (MLP)

O *Multi-Layer Perceptron (MLP)* é uma arquitetura de rede neural supervisionada composta por uma camada de neurônios artificiais de entrada, uma ou mais camadas ocultas e uma camada de saída.

Essa estrutura permite a modelagem de relações complexas entre as variáveis, ao realizar transformações não lineares dos dados por meio das camadas ocultas. Em cada camada, os neurônios recebem entradas que são multiplicadas por pesos sinápticos, somadas e, em seguida, submetidas a uma função de ativação, que introduz a não

linearidade essencial para a aprendizagem dos padrões subjacentes.

O treinamento do MLP é realizado por meio do algoritmo de *backpropagation*, que propaga o erro da camada de saída de volta às camadas anteriores, ajustando iterativamente os pesos de maneira a minimizar a discrepância entre as saídas previstas e os valores desejados. Esse método permite a otimização dos parâmetros do modelo, aprimorando sua capacidade de generalização e resultando em previsões mais precisas (HAYKIN, 2017).

As funções de ativação desempenham um papel crucial no funcionamento do MLP, pois são responsáveis por introduzir a não linearidade necessária à modelagem dos dados. Entre as funções mais utilizadas estão a *ReLU* (*Rectified Linear Unit*), que retorna o máximo entre zero e o valor de entrada, e a *Tanh* (hiperbólica), que mapeia os valores de entrada para o intervalo $(-1, 1)$. Enquanto a *ReLU* é valorizada por sua eficiência computacional e por mitigar o problema do gradiente desaparecido em redes profundas, a *Tanh* é preferida em situações nas quais a centralização dos dados pode melhorar o desempenho do modelo (HAYKIN, 2017; BRAGA et al., 2011).

Dessa forma, o MLP, através do emprego do algoritmo de *backpropagation* e da utilização de funções de ativação, constitui-se numa ferramenta para a modelagem de problemas complexos.

2.2.1.3. Deep Learning e sua aplicação em dados complexos

Deep Learning, ou aprendizado profundo, consiste em um conjunto de técnicas que utilizam redes neurais com múltiplas camadas para modelar relações não lineares complexas e extrair, de forma autônoma, representações hierárquicas dos dados. Essa abordagem apresenta a vantagem de transformar dados brutos em características abstratas e relevantes, dispensando a necessidade de seleção manual de variáveis e possibilitando a identificação de padrões latentes mesmo em conjuntos com alta dimensionalidade e variabilidade (HAYKIN, 2017).

Entre as vantagens do aprendizado profundo destaca-se a sua capacidade de capturar relações não lineares que métodos tradicionais, baseados em médias históricas, não conseguem modelar adequadamente. Essa propriedade é especialmente útil em cenários onde o volume de dados é muito elevado e os padrões de comportamento são

complexos, como no caso dos dados processuais judiciais.

Mayer-Schönberger e Cukier (2013, p. 66) ressaltam que a "dataficação" transforma o mundo em um universo predominantemente composto por informações, o que exige técnicas capazes de processar e interpretar grandes bases de dados. A imensa produção de dados impulsiona o avanço desses métodos, pois a capacidade de processar e analisar sequências temporais e conjuntos de dados multidimensionais é crucial para a obtenção de previsões precisas.

Nesse contexto, *Deep Learning* se destaca por oferecer propriedades como não linearidade, mapeamento de entrada e saída – onde os pesos sinápticos são ajustados a partir de conjuntos de treinamento rotulados – adaptabilidade, resposta à evidência, informação contextual, tolerância a falhas, e uniformidade de análise e projeto, além de uma analogia neurobiológica que aproxima seu funcionamento dos processos cognitivos humanos (HAYKIN, 2017; KARAR et al., 2020).

Adicionalmente, a literatura demonstra que as técnicas de *Deep Learning* têm sido aplicadas com sucesso em diversas áreas que lidam com volumes massivos de dados. Por exemplo, estudos na previsão de demandas energéticas, análise de séries temporais e classificação de documentos evidenciam que esses métodos possibilitam a obtenção de predições mais precisas e robustas, contribuindo para a tomada de decisões estratégicas (KAPLAN; HAENLEIN, 2019).

Embora o presente estudo se concentre na previsão da duração dos processos judiciais, tais técnicas já foram empregadas em setores diversos, o que reforça a versatilidade e a eficácia do *Deep Learning* na modelagem de relações não lineares em ambientes com grande volume de dados.

2.2.2 Inovação e Inteligência Artificial na Gestão Processual

2.2.2.1. Conceito de Inovação no Setor Público

A inovação no setor público pode ser entendida como a introdução de novas ideias, métodos, produtos ou processos que visam aprimorar a eficiência e a qualidade dos serviços oferecidos à sociedade. Nesse contexto, Bekkers *et al.* (2011) distinguem a inovação incremental – caracterizada por melhorias contínuas e graduais nos processos já existentes – da inovação radical, que implica mudanças significativas na forma de

produção e oferta de serviços, transformando de maneira profunda a estrutura organizacional.

Conforme enfatizam Cavalcante e Camões (2017), a inovação no setor público não se restringe à simples incorporação de tecnologias modernas, mas envolve a reestruturação dos processos internos e a adoção de práticas que promovam maior eficiência, transparência e celeridade na prestação jurisdicional.

Além disso, a integração de soluções tecnológicas – como aquelas baseadas em inteligência artificial – no contexto da gestão processual reforça a importância da inovação para otimizar a alocação de recursos e melhorar a tomada de decisões. Dessa forma, a inovação no setor público se configura como um processo multifacetado que demanda a conjugação de avanços tecnológicos com mudanças estruturais, possibilitando que o sistema judiciário se adapte continuamente às exigências e desafios contemporâneos.

2.2.2.2. O Uso de Inteligência Artificial como Ferramenta Inovadora

O emprego de modelos de aprendizado de máquina, em especial os baseados em redes neurais, vem sendo explorado como alternativa para aprimorar o sistema judiciário brasileiro, contribuindo para a automação de tarefas, a redução de gargalos e a otimização da alocação de recursos.

De acordo com a Pesquisa sobre o Uso de Inteligência Artificial no Poder Judiciário (CNJ, 2023), 62 tribunais brasileiros já possuem iniciativas envolvendo inteligência artificial, totalizando 144 projetos em diferentes estágios de maturidade. Desses, 63 estão em produção e 46 em desenvolvimento, o que representa 109 projetos ativos ou em processo de implantação. Os 35 projetos restantes encontram-se na fase de ideação, ainda em nível conceitual ou de planejamento.

Esses dados revelam que a adoção de tecnologias de IA – como a classificação de texto, busca semântica e detecção de similaridade textual – é considerada crucial para a automatização de processos e para a melhoria da eficiência administrativa. Tais iniciativas não apenas visam reduzir o tempo de análise dos processos, mas também proporcionar uma melhor integração dos dados e a identificação de padrões que auxiliem na tomada de decisões por magistrados e servidores.

Nesse contexto, a utilização de redes neurais para a previsão da duração dos processos judiciais destaca-se como uma ferramenta inovadora, justamente por sua capacidade de modelar relações não lineares e de lidar com grandes volumes de dados heterogêneos.

Enquanto as abordagens baseadas em médias históricas se restringem à utilização de parâmetros agregados, ignorando a variabilidade intrínseca dos processos, os algoritmos de *deep learning* podem extrair padrões latentes e ajustar as previsões com base em múltiplas variáveis. Estudos conduzidos por Oliveira et al. (2022) e Silva et al. (2023) demonstram que o uso de técnicas de aprendizado profundo permite alcançar coeficientes de determinação elevados, refletindo uma acurácia superior na previsão dos prazos processuais.

A relevância dessa abordagem inovadora é corroborada por diversos estudos na área, os quais apontam que a aplicação de IA em sistemas judiciais não só melhora a eficiência na tramitação dos processos, mas também oferece subsídios para a tomada de decisão estratégica, conforme ilustrado nos exemplos apresentados a seguir.

Por exemplo, o Projeto Victor, implementado pelo Supremo Tribunal Federal (STF) em parceria com a Universidade de Brasília (UnB), demonstrou que a utilização de redes neurais para identificar processos com repercussão geral pode reduzir drasticamente o tempo de análise de recursos extraordinários, passando de 40 a 60 minutos, no caso da triagem manual, para apenas 5 segundos com o uso da IA (ANDRADE et al., 2021; DAL PONT et al., 2021; MAIA FILHO; JUNQUILHO, 2018).

Outros estudos, como o de Castro Júnior et al. (2020), evidenciam a aplicabilidade de modelos baseados em Processamento de Linguagem Natural (PLN) para a identificação de petições iniciais com teses semelhantes, demonstrando uma precisão de 96% na vinculação de processos repetitivos e na gestão de precedentes.

Da mesma forma, a ferramenta RAFA 2030, analisada por Fux et al. (2022), utiliza técnicas de aprendizado supervisionado e PLN para classificar processos judiciais segundo os Objetivos de Desenvolvimento Sustentável (ODS) da ONU, contribuindo para a priorização de julgamentos e para a eficiência na gestão processual.

Ademais, estudos que investigaram o uso de redes neurais convolucionais multimodais, como os de Mota et al. (2020), evidenciam que a combinação de informações textuais e visuais pode ser empregada para segmentação automática de documentos, o que potencializa a automação da gestão documental no Judiciário. Esses exemplos demonstram que, independentemente da área de aplicação, a adoção de técnicas de IA tem um impacto direto na capacidade de processamento e análise de dados massivos, proporcionando ferramentas que oferecem maior precisão e robustez na previsão de prazos e na gestão dos processos.

Portanto, a incorporação de soluções de IA, fundamentadas em técnicas avançadas de *deep learning*, possibilita uma análise mais detalhada e dinâmica dos dados processuais, criando condições para uma gestão mais eficiente e estratégica do sistema judiciário, para a redução de atrasos e para a melhoria da qualidade dos serviços prestados à sociedade.

2.3. Metodologia e Técnicas

A presente pesquisa adota uma abordagem quantitativa exploratória e explicativa, com o objetivo de analisar padrões e prever o tempo de tramitação de processos judiciais no Superior Tribunal de Justiça (STJ) por meio de aprendizado de máquina e redes neurais, o que permite identificar padrões de tramitação processual e prever a duração de processos judiciais, por meio de correlações não lineares, melhorar a precisão das previsões sobre a duração da tramitação dos processos judiciais no Superior Tribunal de Justiça (STJ), em comparação aos métodos tradicionais atualmente utilizados.

Sampieri et al. (2014) e Creswell (2021) afirmam que pesquisas quantitativas podem ser classificadas em descritivas, correlacionais, experimentais e exploratórias, dependendo da forma como os dados são utilizados para responder às perguntas da pesquisa.

No contexto deste estudo, a abordagem exploratória permite mapear padrões ocultos nos dados processuais, enquanto a abordagem explicativa busca evidenciar o grau de influência entre diferentes variáveis judiciais e a duração dos processos, para garantir clareza na formulação das variáveis, replicabilidade do estudo e validade estatística dos resultados. Essa metodologia possibilita que o algoritmo da rede neural treinada possa ser reaplicado e adaptado para diferentes tribunais, consolidando a utilização da inteligência

artificial como ferramenta auxiliar na modernização do sistema judiciário.

As variáveis foram extraídas do Painel Justiça em Números do CNJ, disponível para acesso público. A variável dependente do estudo é o tempo de tramitação do processo (medido em dias), enquanto as variáveis independentes incluem código da classe processual, código do órgão responsável, frequência de decisões de cada magistrado no tribunal e características processuais associadas, não sendo consideradas variáveis mediadoras ou moderadoras, pois a análise se concentra na modelagem preditiva direta das variáveis que podem influenciar a duração dos processos.

O procedimento metodológico empregado neste estudo está alinhado com pesquisas anteriores que utilizaram aprendizado de máquina para previsão de tempos processuais e modelos estatísticos, como os trabalhos de Oliveira et al. (2022), Silva et al. (2023) e Araújo et al. (2021). Os quais demonstram que modelos estatísticos aplicados ao aprendizado de máquina são capazes de prever a tramitação processual com alto grau de precisão, fornecendo subsídios para a gestão dos tribunais.

No mesmo sentido, a Pesquisa sobre o Uso de Inteligência Artificial no Poder Judiciário (CNJ, 2023) revelou que 62 tribunais brasileiros já utilizam aprendizado de máquina para previsão processual, sendo que 87% desses projetos aplicam abordagens supervisionadas, como a utilizada nesta pesquisa. Isso reforça a validade da abordagem quantitativa exploratória e explicativa e demonstra que o aprendizado de máquina é um método aceito e aplicado no sistema judiciário brasileiro.

Embora a literatura recente tenha explorado o uso de técnicas de inteligência artificial para previsão de tempos processuais no Judiciário, como no estudo realizado por Oliveira et al. (2022), é importante destacar que a abordagem adotada apresenta originalidade clara e objetiva. Enquanto Oliveira et al. (2022) se concentram especificamente na previsão do tempo para emissão da primeira decisão judicial após a disponibilização do processo ao magistrado, este estudo avança ao considerar a duração integral do processo judicial desde o ingresso até a sua baixa definitiva no Superior Tribunal de Justiça (STJ), ampliando assim o escopo e o impacto potencial das previsões geradas.

Assim, esta pesquisa oferece uma contribuição inédita ao aplicar redes neurais

artificiais, especificamente *Multi-Layer Perceptron (MLP)*, para prever todo o período de tramitação do processo judicial no STJ. Tal abordagem expande o conhecimento atual sobre aplicação da inteligência artificial na gestão judiciária, permitindo previsões mais abrangentes, o que facilita a alocação estratégica de recursos administrativos e judiciais ao longo de todo o ciclo processual.

Para responder às questões de pesquisa propostas, a metodologia foi estruturada em três etapas principais: **coleta e tratamento dos dados, treinamento e validação da rede neural e análise e interpretação dos resultados.**

A primeira etapa consistiu na coleta e tratamento dos dados, extraídos diretamente do Painel Justiça em Números do CNJ, utilizando o conjunto de bibliotecas específicas da linguagem de programação python. Seguindo a metodologia similar a utilizada por Cruz et al. (2024), foram analisados 631.437 processos baixados entre janeiro de 2023 e junho de 2024, independentemente da data de entrada desse processo no STJ, garantindo um conjunto de dados numericamente significativo para o treinamento da rede neural. Essa etapa teve como principal objetivo mitigar o risco de que o modelo fosse treinado com dados inconsistente, não padronizados ou insuficientes.

Na segunda etapa, foi realizado o treinamento e validação do modelo preditivo, utilizando uma rede neural do tipo *Multi-Layer Perceptron (MLP)*, treinada por meio de validação cruzada e *RandomizedSearchCV* para ajuste de hiperparâmetros. A escolha da rede neural MLP baseia-se na literatura especializada, que destaca sua eficiência na modelagem de dados complexos e na previsão de séries temporais, como observado nos estudos de Gruginskie & Vaccaro (2018) e Mota et al. (2020).

A técnica de otimização de hiperparâmetros via *RandomizedSearchCV*, além de maximizar o desempenho, também representa um elemento de inovação metodológica, permitindo que a aplicação do modelo seja ajustada com maior precisão às especificidades do conjunto de dados trabalhados.

O processo de validação teve como objetivo permitir, com certo grau de confiabilidade, que o modelo tivesse capacidade de generalização e evitar problemas de sobreajuste (*overfitting*), de forma a possibilitar que as previsões fossem aplicáveis a novos processos judiciais.

Na terceira etapa, foi realizada a análise e interpretação dos resultados, utilizando as métricas estatísticas coeficiente de determinação (R^2) e erro absoluto médio (MAE). O R^2 indica o percentual da variação dos dados explicada pelo modelo, enquanto o MAE mede a diferença média entre as previsões e os valores reais.

Além dessa análise estatística, foi aplicada a técnica de *SHAP values* (*SHapley Additive exPlanations*) para interpretar a influência de cada variável na previsão da tramitação processual. Os *SHAP values* permitem quantificar a contribuição individual de cada variável na previsão gerada pelo modelo.

Dessa forma, a abordagem segue a metodologia utilizada por Lundberg e Lee (2017), que propuseram os *SHAP values* como ferramenta para interpretar modelos de aprendizado de máquina de forma compreensível e explicável. No contexto deste estudo, essa análise possibilita entender quais fatores mais impactaram a previsão do tempo de tramitação dos processos judiciais no STJ, dentre as variáveis utilizados para treinamento da rede neural.

Segundo Pedregosa *et al.* (2011), a biblioteca *scikit-learn* torna o *machine learning* acessível para não especialistas, enfatizando a facilidade de sua utilização, desempenho e consistência, com poucas dependências e distribuição sob a licença de código aberto, o que incentiva sua utilização tanto em ambientes acadêmicos quanto comerciais, bem como promove uma consistência e uma padronização nas práticas de modelagem, contribuindo significativamente para o avanço da pesquisa e desenvolvimento em inteligência artificial.

Conforme destacado por Pedregosa *et al.* (2011), a escolha estratégica do *scikit-learn* em utilizar *arrays* do *numpy* para representação de dados de entrada, garante eficiência e integração com o ecossistema de ciência de dados em *python*.

Essa integração facilita a implementação de *pipelines* de *machine learning*, desde a fase de pré-processamento de dados até a modelagem e avaliação de algoritmos, em um ambiente propício à experimentação e à inovação acadêmica (PEDREGOSA *et al.* 2011).

2.4. Resultados e Discussão

2.4.1 Resultados

Os resultados obtidos com o treinamento do algoritmo indicam um alto grau de eficácia e assertividade, no que diz respeito à previsão da duração dos processos judiciais no Superior Tribunal de Justiça (STJ), considerando o conjunto de dados testados. A seguir, são apresentadas as principais métricas de desempenho e os resultados das combinações de hiperparâmetros testadas.

2.4.1.1. Descrição da Estrutura do Conjunto de Dados

Para treinamento do modelo preditivo, foi selecionada uma amostra aleatória simples, utilizando o método *sample(frac=0.10, random_state=42)* da biblioteca *pandas do python*. A amostra final foi composta por 50.515 processos judiciais baixados, abrangendo 3.293 características (colunas) associadas a cada processo.

A escolha da amostra correspondente a **aproximadamente 10% do conjunto original** (cerca de 600.000 processos) foi feita com base no equilíbrio entre a necessidade de representatividade estatística e as limitações técnicas do ambiente de processamento utilizado (*Google Colaboratory*, com 334 GB de RAM disponíveis).

Após o processo de limpeza e preparação dos dados, a amostra foi dividida em dois subconjuntos: **80% dos dados (40.412 processos)** foram utilizados para **treinamento da rede neural**, e **20% (10.103 processos)** foram destinados à **validação (teste) do modelo**. Essa divisão foi realizada com o auxílio da função *train_test_split*, utilizando *random_state=42* para garantir reprodutibilidade.

O tamanho e a qualidade da amostra foram validados a posteriori, considerando o desempenho obtido no treinamento do modelo preditivo ($R^2 \approx 92\%$), o que reforça que os dados selecionados foram suficientemente representativos para permitir generalização confiável dos resultados.

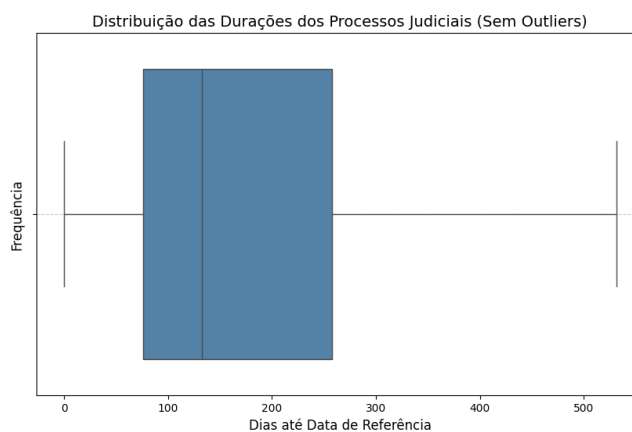
As principais estatísticas descritivas da duração dos processos judiciais em dias são apresentadas a seguir:

- **Quantidade de Processos Analisados:** 631.437;
- **Média:** 251,43 dias;

- **Desvio Padrão:** 407,06 dias;
- **25° Percentil:** 76 dias;
- **Mediana (50° Percentil):** 133 dias;
- **75° Percentil:** 259 dias;
- **Mínimo:** 0 dias; e
- **Máximo:** 11.221 dias.

Para ilustrar a análise, os dados foram aplicados à Figura 2.1, constituindo um *boxplot* da distribuição das durações dos processos judiciais, excluindo-se os *outliers*, podendo ser interpretada à luz das estatísticas descritivas apresentadas para a amostra de 631.437 processos.

Figura 2.1 - BOXPLOT Distribuição da Duração dos Processos Judiciais



Fonte: elaborado pelo autor

Observa-se que a **mediana**, situada em **133 dias**, está deslocada para a parte inferior da caixa, confirmando que a maior parte dos processos possui uma tramitação relativamente curta em relação à média de **251,43 dias**. Essa diferença entre a mediana e a média sugere a presença de assimetria positiva, ou seja, um conjunto de processos com durações muito superiores à mediana está impactando a média, tornando-a mais elevada.

O **intervalo interquartil (IQR)**, delimitado pelo **25° percentil (76 dias)** e **75° percentil (259 dias)**, representa a faixa onde estão concentrados **50% dos processos** analisados. Isso significa que, na maior parte dos casos, a tramitação processual varia entre cerca de **dois meses e meio a oito meses e meio**. Essa informação é relevante para

modelos preditivos, pois indica que previsões dentro desse intervalo possuem maior probabilidade de acerto.

O **desvio padrão de 407,06 dias** sugere uma dispersão elevada na duração dos processos, indicando que há uma grande variabilidade nos tempos de tramitação. Essa dispersão é evidenciada pelos valores **mínimos de 0 dias** (provavelmente relacionado a processos que foram resolvidos imediatamente) e **máximo de 11.221 dias** (o equivalente a mais de 30 anos), reforçando a existência de um subconjunto de processos com duração extremamente longa.

A presença dessa **assimetria positiva**, com casos de tramitação excessivamente longa, pode ser explicada por fatores como **complexidade processual, recursos interpostos e variação na carga de trabalho dos magistrados e tribunais**. Esse comportamento é fundamental para a interpretação dos resultados do modelo preditivo, pois sugere que as previsões devem considerar diferentes perfis de tramitação e não apenas a média dos dados.

Além disso, a remoção de outliers foi essencial para obter um retrato mais fiel da distribuição da duração processual sem a interferência de casos extremos, permitindo uma melhor calibração do modelo de previsão e garantindo maior precisão na estimativa do tempo de tramitação dos processos no **Superior Tribunal de Justiça (STJ)**.

2.4.1.2. Desempenho de Hiperparâmetros

A Tabela 2.1 resume o desempenho de cada combinação de hiperparâmetros testada durante o processo de otimização, ordenadas conforme o seu grau de assertividade:

Tabela 2.1 – Desempenho das Combinações de Hiperparâmetros

Combinação de Hiperparâmetros	Média do R^2	Desvio Padrão do R^2
{'model__max_iter': 200, 'model__learning_rate_init': 0.001, 'model__hidden_layer_sizes': (50,), 'model__alpha': 0.001}	0.927150	0.003116
{'model__max_iter': 200, 'model__learning_rate_init': 0.01, 'model__hidden_layer_sizes': (50, 100), 'model__alpha': 0.001}	0.922999	0.003186

{'model__max_iter': 200 , 'model__learning_rate_init': 0.01 , 'model__hidden_layer_sizes': (100 ,), 'model__alpha': 0.001 }	0.922152	0.004866
{'model__max_iter': 500 , 'model__learning_rate_init': 0.001 , 'model__hidden_layer_sizes': (100 ,), 'model__alpha': 0.001 }	0.915588	0.001064
{'model__max_iter': 500 , 'model__learning_rate_init': 0.001 , 'model__hidden_layer_sizes': (50 , 50), 'model__alpha': 0.001 }	0.896269	0.004947
{'model__max_iter': 500 , 'model__learning_rate_init': 0.001 , 'model__hidden_layer_sizes': (50 , 100), 'model__alpha': 0.001 }	0.882373	0.005589
{'model__max_iter': 1000 , 'model__learning_rate_init': 0.01 , 'model__hidden_layer_sizes': (50 , 100), 'model__alpha': 0.001 }	0.882373	0.005589
{'model__max_iter': 500 , 'model__learning_rate_init': 0.01 , 'model__hidden_layer_sizes': (50 , 100), 'model__alpha': 0.001 }	0.882373	0.005589
{'model__max_iter': 500 , 'model__learning_rate_init': 0.01 , 'model__hidden_layer_sizes': (50 , 50), 'model__alpha': 0.001 }	0.880978	0.005060
{'model__max_iter': 500 , 'model__learning_rate_init': 0.01 , 'model__hidden_layer_sizes': (50 , 100), 'model__alpha': 0.01 }	0.878742	0.006214

Fonte: elaborado pelo autor

Após a realização da busca por hiperparâmetros com o método *RandomizedSearchCV*, conforme Tabela 2.1, os melhores hiperparâmetros encontrados foram:

- Número de iterações máximas (***max_iter***): 200;
- Taxa de aprendizado inicial (***learning_rate_init***): 0.001;
- Tamanho das camadas ocultas (***hidden_layer_sizes***): (50,); e
- Regularização L2 (***alpha***): 0.001.

Esses parâmetros foram escolhidos com base em sua capacidade de maximizar o desempenho do modelo, refletindo uma otimização adequada para a complexidade dos dados.

2.4.1.3. Descrição das Variáveis (Features)

Para estimar a duração dos processos judiciais no STJ, diversas *features* (variáveis) foram utilizadas. A seguir, detalha-se cada uma delas, considerando sua relevância e possível impacto no desempenho do modelo.

A variável “*Data de Início do Processo*” captura padrões sazonais e tendências temporais que podem influenciar a duração dos processos. As variáveis derivadas da data de início, como “*Ano_Inicio*”, “*Mes_Inicio*” e “*Dia_Semana_Inicio*”, adicionam camadas de informação temporal relevantes para capturar tendências específicas. Por exemplo, processos iniciados em dias próximos a feriados ou em determinados dias da semana podem ter durações diferentes devido a fatores operacionais. Como essas informações estão disponíveis no início do processo, seu uso é apropriado.

Já as variáveis “*Códigos de Assuntos*” e “*Códigos de Classes Processuais*” representam os temas jurídicos abordados nos processos. Ao converter esses códigos em variáveis *dummy* por meio da codificação *one-hot*, o modelo captura a influência específica de cada assunto e natureza na duração do processo. Essa granularidade é importante, pois diferentes assuntos e naturezas processuais podem ter complexidades e procedimentos distintos.

A variável “*Código do Órgão*” identifica a unidade judiciária responsável pelo processo. Diferentes órgãos podem ter cargas de trabalho, eficiência e práticas processuais variadas, influenciando o tempo de tramitação. Da mesma forma, a variável “*Procedimento*” permite que o modelo diferencie entre processos que, embora possam ser da mesma classe ou assunto, seguem fluxos processuais diferentes. A codificação adequada dessa variável assegura que o modelo capture essas nuances sem introduzir ruído desnecessário.

A criação da variável combinada “*Código do Órgão_Procedimento*” busca capturar interações específicas entre o órgão e o procedimento adotado, com intuito de revelar padrões que não seriam evidentes ao considerar as variáveis isoladamente. Com lógica similar, a variável “*Frequência do Tribunal*”, calculada como a proporção de processos associados a cada órgão julgador, tem o potencial de refletir a carga de trabalho e a eficiência de cada unidade judiciária.

2.4.1.4. Desempenho do Modelo

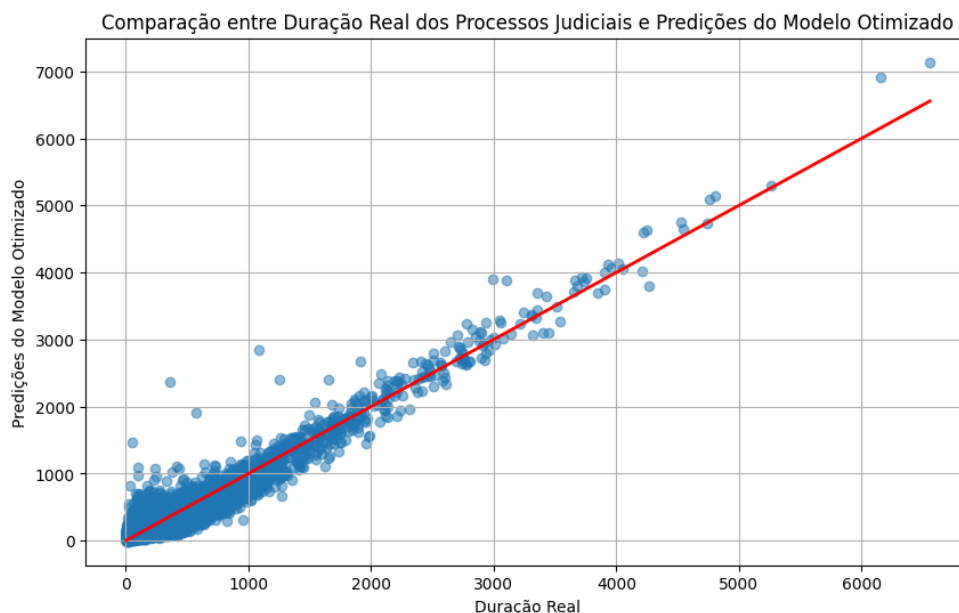
O desempenho do modelo foi avaliado por meio de duas métricas principais:

- **R²**: 0.9243; e
- **Erro Absoluto Médio (MAE)**: 71.61.

O valor do **R²** indica que aproximadamente **92.43% da variabilidade** na duração dos processos judiciais **pode ser explicada pelas variáveis preditivas** utilizadas, o que demonstra uma excelente capacidade de previsão do modelo. O **MAE**, por sua vez, sugere que, em média, **as previsões do modelo estão a cerca de 71 dias da duração real dos processos**, o que deve ser ponderado com cautela uma vez que a média de duração dos processos encontrada no conjunto de dados foi de 259 dias.

A partir dos resultados, foi gerado um gráfico de dispersão contendo como variáveis a previsão do modelo com as configurações otimizadas e a duração real dos processos judiciais, conforme Figura 2.2.

Figura 2.2 - Gráfico de Dispersão dos Resultados



Fonte: elaborado pelo autor

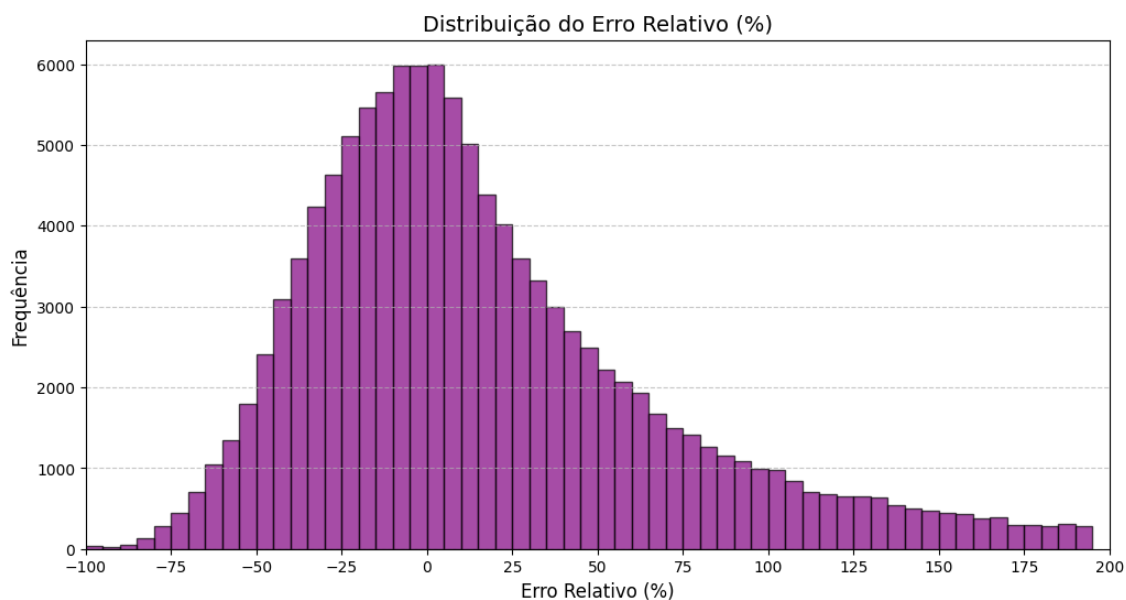
Ao analisar da Figura 2.2, percebe-se que a maioria dos pontos se aproxima da linha vermelha (duração real do processo), isso indica que o algoritmo tem uma boa capacidade de prever a duração dos processos na maior parte dos casos. Apesar disso, ao

observar os dados em longas durações - especialmente acima de 3000 dias, nota-se uma dispersão maior dos pontos, sugerindo que o modelo pode ter dificuldades em prever com precisão nesses casos.

O mesmo pode ser percebido na Figura 2.3, ao analisar a distribuição do erro relativo (%) percebe-se uma distribuição assimétrica à direita, indicando que, na maioria dos casos, o modelo apresenta erros pequenos ou moderados, concentrados em torno do zero, o que significa que as previsões foram, em grande parte, próximas dos valores reais.

A distribuição tem um pico próximo de 0%, evidenciando que o modelo tem uma boa acurácia na maioria das previsões.

Figura 2.3 - Gráfico de Distribuição de Erros



Fonte: elaborado pelo autor

No entanto, observa-se uma cauda longa à direita, sugerindo que há uma quantidade menor, mas significativa, de casos com erros elevados. Isso indica que, para alguns processos, o modelo superestima a duração da tramitação, possivelmente devido a variáveis não capturadas adequadamente ou eventos atípicos no fluxo processual.

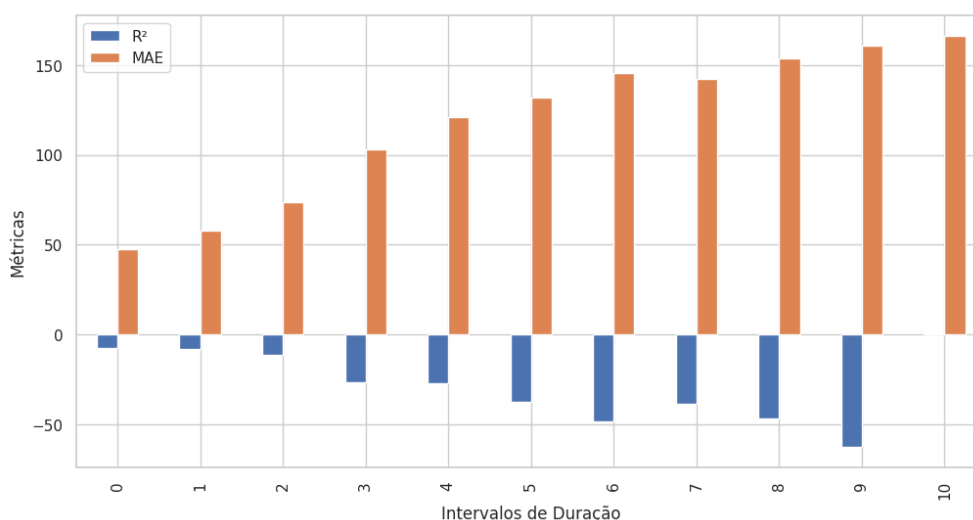
Essa assimetria pode estar relacionada, por exemplo, a decisões monocráticas, acordos prévios ou liminares, que resultam em tramitações mais rápidas e não estão explicitamente representadas nas variáveis do modelo. Além disso, a ausência de informações sobre movimentações internas — como redistribuições, prioridades de julgamento ou atuação de órgãos específicos — pode influenciar esses desvios.

Por outro lado, a assimetria menor à esquerda indica que o modelo também subestima a duração de alguns processos, mas em menor grau. Isso pode estar relacionado a características específicas de determinados processos ou à variabilidade inerente ao tempo de tramitação.

Isso também pode ser constatado na Figura 2.4, na qual exemplifica-se a análise do desempenho do modelo em diferentes intervalos de duração dos processos judiciais, indicando que a eficácia da previsão varia conforme a extensão temporal dos casos analisados.

O coeficiente de determinação (R^2), representado pelas barras azuis, apresenta valores negativos em praticamente todos os intervalos, o que sugere que o modelo não consegue explicar bem a variabilidade dos dados nesses grupos. Esse comportamento indica que, em determinados casos, as previsões podem ser menos precisas do que o uso da média histórica dos dados.

Figura 2.4 - Desempenho do Modelo por Intervalo de Duração dos Processos Judiciais



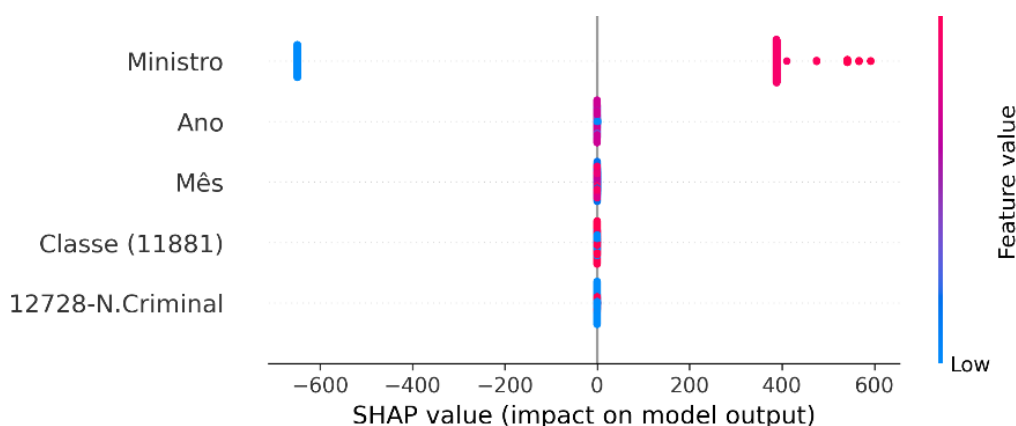
Fonte: elaborado pelo autor

O erro absoluto médio (MAE), representado pelas barras laranjas, cresce conforme aumenta a duração do processo, sugerindo que o modelo apresenta um erro médio mais baixo para os processos de curta duração e maior dificuldade para prever corretamente os casos que se estendem por períodos mais longos. Nos primeiros intervalos de duração, o MAE se mantém em níveis reduzidos, indicando que as previsões são mais confiáveis nesses cenários. No entanto, à medida que a duração dos processos aumenta, o MAE cresce significativamente, refletindo um desempenho menos eficiente

do modelo preditivo para processos com durações mais longas.

Da mesma forma, a análise dos resultados utilizando *SHAP* (*SHapley Additive exPlanations*) permitiu visualizar as 5 variáveis utilizadas que mais impactaram as previsões do modelo, conforme Figura 2.5.

Figura 2.5 - Gráfico *SHapley Additive exPlanations* (*SHAP*)



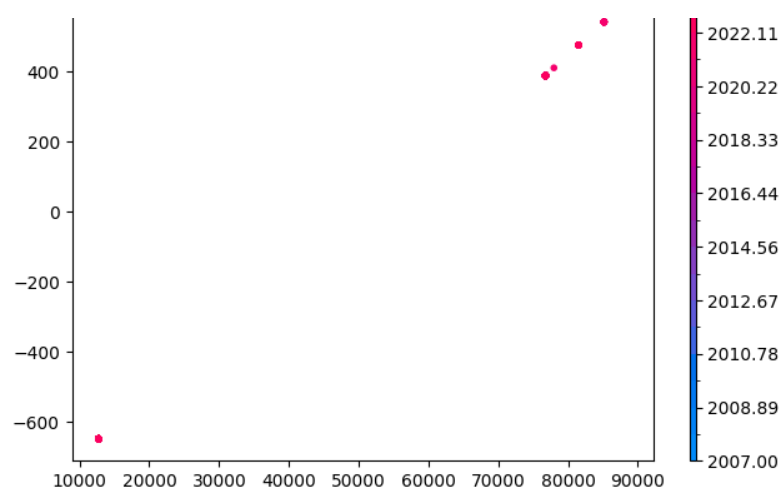
Fonte: elaborado pelo autor

A avaliação da importância das variáveis indicou que o fator "Ministro", aqui entendido como "órgão julgador", podendo ser por exemplo 1ª Turma ou Presidência, exerce uma influência significativa na previsão da duração dos processos, com valores *SHAP* elevados e bem distribuídos, sugerindo que o órgão responsável pelo julgamento é um elemento determinante para a celeridade processual.

Em contrapartida, variáveis temporais, como "Ano" e "Mês", demonstraram menor impacto direto nas previsões do modelo, apresentando valores *SHAP* próximos de zero. Isso sugere que esses fatores, isoladamente, não são preditores primários da duração dos processos, podendo estar mais associados a outras variáveis interligadas no sistema judiciário.

A partir desse resultado, considerando as 5 variáveis mais relevantes apontadas na Figura 2.5, o gráfico da Figura 2.6 mostra a relação entre o impacto das variáveis no modelo preditivo e a evolução temporal dos processos judiciais. O eixo horizontal representa uma métrica associada ao tempo, enquanto o eixo vertical indica a magnitude do impacto das variáveis no resultado do modelo. A escala de cores, variando do azul ao rosa, representa os anos dos processos, evidenciando padrões temporais.

Figura 2.6 - Evolução do Impacto das Variáveis ao Longo do Tempo



Fonte: elaborado pelo autor

Isso sugere que, ao longo do tempo, os fatores que influenciam a tramitação processual podem ter sofrido mudanças, com um aumento do impacto das variáveis mais recentes. Isso pode indicar a modernização do Judiciário, mudanças na composição dos tribunais ou alterações na legislação e nos procedimentos internos do STJ. Além disso, a presença de um ponto discrepante na parte inferior esquerda sugere um caso atípico, que pode ter sido influenciado por uma variável externa ou um erro na base de dados.

2.4.2 Discussão dos Resultados

Os resultados obtidos na pesquisa evidenciam a capacidade das redes MLP de capturar padrões complexos relacionados à duração dos processos judiciais no Superior Tribunal de Justiça (STJ). Com um coeficiente de determinação (R^2) de **0,9243** e um erro absoluto médio (**MAE**) de **71,61** dias, o modelo demonstrou desempenho preditivo significativo.

Esse resultado evidencia uma melhora expressiva na precisão das previsões em relação ao uso da média histórica como estimativa única, que, no caso da base analisada, é de aproximadamente 251 dias, com um desvio padrão elevado (407 dias). Ou seja, enquanto a média não considera a variabilidade e complexidade dos casos individuais, o modelo preditivo baseado em rede neural foi capaz de adaptar-se a essas diferenças, entregando estimativas mais personalizadas e precisas.

Portanto, esse resultado vai ao encontro dos argumentos apresentados por Kaplan

e Haenlein (2019) e Braga et al. (2011), que destacam a alta capacidade das redes neurais e técnicas de *deep learning* em lidar com conjuntos de dados complexos e não lineares, como a base de dados utilizada nesta pesquisa.

Da mesma forma, conforme destacado por Mayer-Schönberger e Cukier (2013), essa diversidade e volume maciço de informações reforça a importância do emprego de técnicas avançadas como *deep learning*, capazes de extrair padrões latentes de dados altamente heterogêneos.

Por outro lado, com relação aos achados desta pesquisa e a literatura relacionada, o presente trabalho tem por contribuição central sua originalidade metodológica. Diferentemente de estudos anteriores, como Oliveira *et al.* (2022) – o qual possui maior proximidade com o escopo deste projeto, cujo foco foi a previsão do tempo até a primeira decisão judicial, a presente pesquisa inovou ao abordar a duração total dos processos judiciais no STJ, desde sua entrada até sua conclusão.

Ambos os estudos, Oliveira et al. (2022) e o modelo desenvolvido neste trabalho, utilizaram técnicas de aprendizado supervisionado com o objetivo de prever a duração de processos judiciais. Contudo, apresentam diferenças significativas quanto à base de dados, escopo da previsão, técnicas aplicadas e desempenho dos modelos.

No que diz respeito à base de dados, Oliveira et al. (2022) trabalharam com informações fornecidas pelo Tribunal Regional do Trabalho, totalizando 8 milhões de registros, dos quais foram identificados 333.890 processos distintos, com dados de 94 órgãos julgadores, 375.995 partes, 341.860 advogados e 208 magistrados. A variável-alvo foi a duração entre a data de início da deliberação e a data da decisão, ou seja, o tempo até a primeira decisão.

Já o modelo proposto neste trabalho utilizou 631.437 processos judiciais baixados do STJ entre janeiro de 2023 e junho de 2024, focando na duração total da tramitação processual, desde o ingresso até o arquivamento do processo. Foram utilizados 50.515 processos na amostra de treinamento e validação, com 3.293 variáveis por processo.

Em termos metodológicos, Oliveira *et al.* aplicaram algoritmos de regressão, entre eles *MLP*, *Adaboost*, *Gradient Boosting*, *SVR* e empilhamentos com diferentes modelos. Os dados foram submetidos a duas técnicas de pré-processamento e divididos igualmente

entre treino/validação e teste. O *MLP*, nesse estudo, obteve R^2 de 0,664 e *RMSE* de 79,145 (com pré-processamento original), enquanto o *Adaboost* foi o melhor modelo, com R^2 de 0,819 e *RMSE* de 65,152.

Por outro lado, o presente estudo empregou a técnica de rede neural ***Multi-Layer Perceptron (MLP)***, com ajuste de hiperparâmetros via ***RandomizedSearchCV*** e validação cruzada. O modelo atingiu um desempenho de $R^2 = 0,9243$ e $MAE = 71,61$ dias, o que indica excelente aderência aos dados e capacidade preditiva robusta, especialmente considerando a complexidade envolvida na previsão de todo o ciclo processual.

No que tange à interpretabilidade, ambos os estudos utilizaram a técnica de ***SHAP values*** para compreender o impacto das variáveis nas previsões. Oliveira *et al.* destacou que “magistrado” e “órgão julgador” foram as variáveis com maior influência na previsão do tempo de decisão. No presente trabalho, resultados semelhantes foram observados, com o fator “Ministro” (órgão julgador no STJ) aparecendo entre os mais relevantes, enquanto variáveis como “Ano” e “Mês” apresentaram menor impacto.

Esses resultados sugerem que, no contexto específico do STJ, fatores institucionais e processuais são mais determinantes na duração dos processos do que as sazonalidades temporais isoladas. Isso está alinhado com as conclusões encontradas na literatura sobre a aplicação de inteligência artificial em gestão judiciária, em que fatores contextuais, organizacionais e institucionais geralmente superam elementos temporais simples (MAIA FILHO; JUNQUILHO, 2018; FUX et al., 2022).

Dessa forma, a abordagem metodológica aqui utilizada foi substancialmente mais complexa, pois o modelo precisa capturar uma série de eventos intermediários, decisões administrativas e fatores processuais que extrapolam a etapa inicial de julgamento. Portanto, a originalidade deste estudo reside justamente na sua amplitude, oferecendo uma visão global sobre a duração da tramitação processual, algo ainda não suficientemente explorado pela literatura acadêmica recente no Brasil.

A distribuição assimétrica observada no gráfico de dispersão e na distribuição de erros, com concentração próxima a valores baixos, mas com significativa presença de erros extremos, aponta um ponto crítico para melhorias futuras. Esse aspecto sugere que, embora o modelo apresente boa precisão geral, ainda há dificuldades em lidar com casos

atípicos ou extremos. Oliveira et al. (2022) já alertavam sobre a necessidade de aprofundar a compreensão dos processos mais complexos ou atípicos, pois esses casos frequentemente contêm elementos informativos importantes que, se bem integrados ao modelo, poderiam melhorar ainda mais seu desempenho preditivo.

Adicionalmente, os resultados demonstraram uma relação inversa entre o desempenho preditivo do modelo e a duração do processo. Conforme ilustrado pelos gráficos apresentados, a eficácia do modelo diminui à medida que o tempo de tramitação dos processos aumenta, revelando dificuldades na previsão de casos mais complexos e longos.

Essa situação reforça a importância da abordagem proposta por Haykin (2017), ao enfatizar que a eficácia das redes neurais depende não apenas da quantidade, mas também da qualidade e representatividade das variáveis utilizadas para treinar o modelo. Isso sugere que futuras pesquisas poderiam explorar técnicas adicionais de seleção ou redução de dimensionalidade.

Por fim, esses resultados também destacam a importância das estratégias de inovação incremental no Judiciário brasileiro. Como ressaltado por Bekkers *et al.* (2011), inovações incrementais, como o modelo aqui desenvolvido, têm maior facilidade de aceitação institucional, além de serem mais facilmente assimiladas pelas equipes administrativas e jurídicas.

2.5. Conclusão

Este estudo teve como objetivo desenvolver e validar um modelo preditivo baseado em redes neurais artificiais, especificamente por meio da técnica *Multi-Layer Perceptron (MLP)*, para aprimorar a previsão da duração integral da tramitação de processos judiciais no Superior Tribunal de Justiça (STJ). Para tanto, empregou-se uma metodologia quantitativa exploratória e explicativa, com otimização de hiperparâmetros via *RandomizedSearchCV* e validação estatística.

Os resultados encontrados demonstraram eficácia significativa do modelo desenvolvido, com coeficiente de determinação (R^2) otimizado de 0,9243 e um erro absoluto médio (MAE) de aproximadamente 71 dias, indicando alta precisão na maioria das previsões realizadas. As análises com *SHAP values* revelaram que fatores

institucionais, em especial a identidade dos magistrados responsáveis pelo julgamento dos processos, desempenharam um papel preponderante na determinação do tempo de tramitação, superando o impacto de variáveis puramente temporais, como mês ou ano de início do processo.

Dessa forma, este estudo respondeu afirmativamente à questão de pesquisa proposta, evidenciando que as técnicas de redes neurais, especialmente por meio do Deep Learning, representam avanços substanciais em relação aos métodos tradicionais baseados em médias históricas, destacando-se por capturar relações não lineares e complexas, típicas do contexto judiciário brasileiro.

No que concerne às contribuições acadêmicas e práticas, este trabalho trouxe avanços, primeiramente, em nível teórico, ampliou-se o escopo dos estudos existentes ao realizar previsões para a duração integral dos processos judiciais, indo além das abordagens focadas exclusivamente na previsão da primeira decisão judicial. Metodologicamente, o estudo demonstrou a aplicabilidade e eficácia de técnicas avançadas de aprendizado de máquina (*MLP* otimizado via *RandomizedSearchCV* e análise com *SHAP*) no contexto judicial.

Em termos práticos e institucionais, o modelo proposto tem potencial para ser integrado aos sistemas de gestão processual do STJ, oferecendo subsídios concretos para decisões estratégicas, melhor alocação de recursos e intervenção preventiva em processos com alta probabilidade de duração prolongada.

Entretanto, reconhecem-se algumas limitações na pesquisa. Destacam-se especialmente a restrição quanto à capacidade computacional que limitou o uso integral do conjunto de dados disponível (apenas 10% da base original utilizada), e a dificuldade enfrentada pelo modelo em prever com alta precisão os casos extremos e de duração excepcionalmente longa, cujas variáveis específicas podem não ter sido plenamente capturadas pelo conjunto de dados empregado.

Diante dessas limitações, sugere-se que pesquisas futuras explorem técnicas híbridas ou a segmentação prévia dos processos (por exemplo, *clustering*) para lidar melhor com processos atípicos ou excepcionalmente longos. Recomenda-se ainda a exploração de bases maiores de dados, com recursos computacionais mais avançados, bem como a inclusão de variáveis adicionais que possam capturar melhor a complexidade

processual e institucional presente nos casos mais extremos.

REFERÊNCIAS

- ALMEIDA NETO, A. et al. **Um modelo de inferência para a classificação de resultados processuais da Justiça Estadual**. Revista de Engenharia e Pesquisa Aplicada, v. 3, n. 3, ago. 2018.
- ANDRADE, M. D. et al. **Inteligência Artificial na gestão processual: Projeto Victor e seus impactos na classificação e triagem de processos**. Revista Opinião Jurídica, v. 21, n. 1, p. 312–335, 2021.
- ARAÚJO, R. M. et al. **Modelos estatísticos aplicados à previsão de prazos judiciais**. Revista de Administração Judiciária, v. 5, n. 1, p. 45-60, 2021.
- BEKKERS, V.; EDELENBOS, J.; STEIJN, B. **Innovation in the Public Sector: Linking Capacity and Leadership**. New York: Palgrave Macmillan, 2011.
- BRAGA, A. P.; LUDERMIR, T. B.; CARVALHO, A. **Redes Neurais Artificiais: Teoria e Aplicações**. Rio de Janeiro: LTC Editora, 2011.
- CAVALCANTE, P.; CAMÕES, M. **Inovação Pública no Brasil: teoria, tendências e casos no Brasil**. Brasília: ENAP, 2017.
- CASTRO JÚNIOR, R. et al. **Aplicação de PLN para identificação automatizada de petições com teses semelhantes**. Revista Jurídica Brasileira, v. 22, n. 3, p. 210-230, 2020.
- CASTRO, M. P.; GUIMARÃES, T. A. **Dimensões da inovação em organizações da justiça: proposição de um modelo teórico-metodológico**. Cadernos EBAPE.BR, v. 17, n. 1, p. 173–184, mar. 2019.
- CONSELHO NACIONAL DE JUSTIÇA. **Pesquisa sobre o Uso de Inteligência Artificial no Poder Judiciário**. Brasília: CNJ, 2023.
- CRESWELL, J. W.; CRESWELL, J. D. **Projeto de pesquisa: métodos qualitativos, quantitativos e mistos**. 5. ed. Porto Alegre: Penso, 2021.

CRUZ, A. et al. **Análise de bases de dados processuais judiciais utilizando Python.** Revista Brasileira de Ciência de Dados, v. 7, n. 1, p. 15-30, 2024.

DAL PONT, E. et al. **Automação e IA no Supremo Tribunal Federal: Análise do Projeto Victor.** Revista Brasileira de Tecnologia Jurídica, v. 4, n. 2, p. 120-140, 2021.

FUX, L. et al. **Classificação de processos judiciais usando aprendizado de máquina segundo os Objetivos de Desenvolvimento Sustentável.** Revista da Escola Nacional de Magistratura, v. 3, n. 2, p. 50-68, 2022.

GIACOMAZZO, G.; LEITE, D. **O mestrado profissional no âmbito das políticas públicas na educação: reflexões sobre a produção do conhecimento científico.** p. 475–493, 2014.

GIL-GARCIA, J. R.; DAWES, S. S.; PARDO, T. A. **Digital government and public management research: finding the crossroads.** Public Management Review, v. 20, n. 5, p. 633–646, 4 maio 2018.

GRUGINSKIE, L.; VACCARO, G. **Previsão de séries temporais utilizando redes neurais MLP.** Revista Brasileira de Computação Aplicada, v. 10, n. 2, p. 85-102, 2018.

HAYKIN, S. **Redes neurais: princípios e prática.** 2. ed. Porto Alegre: Bookman, 2017.

KAPLAN, A.; HAENLEIN, M. **Siri, Siri in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence.** Business Horizons, v. 62, n. 1, p. 15-25, 2019.

KARAR, M. E. et al. **Deep learning applications for predicting big data analytics outcomes.** Journal of Big Data Analytics, v. 7, n. 3, p. 67-78, 2020.

LACERDA, R. T.; D'AMORIM, M.; ALVIM, L. **O uso da Inteligência Artificial na Justiça Brasileira: uma análise exploratória.** Revista de Gestão Pública, v. 10, n. 3, p. 65-80, 2021.

LUNDBERG, S. M.; LEE, S.-I. **A Unified Approach to Interpreting Model Predictions.** Advances in Neural Information Processing Systems, v. 30, p. 4765-4774, 2017.

MACOHIN, A. **Análise da Jurisprudência do Conselho Nacional de Justiça através**

de Redes Complexas. Revista CNJ, v. v. 3, n. 2, p. 19–26, 2019.

MAIA FILHO, A.; JUNQUILHO, G. S. **Gestão de processos judiciais com inteligência artificial: experiências e desafios.** Revista Brasileira de Administração Pública, v. 52, n. 4, p. 218-237, 2018.

MAYER-SCHÖNBERGER, V.; CUKIER, K. **Big Data: Como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana.** Rio de Janeiro: Elsevier, 2013.

MCCARTHY, J. **What is Artificial Intelligence.** Stanford: Stanford University, 2007.

McCULLOCH, W. S.; PITTS, W. **A logical calculus of the ideas immanent in nervous activity.** The bulletin of mathematical biophysics, v. 5, n. 4, p. 115–133, 1943. Reimpresso em Bulletin of Mathematical Biology, v. 52, p. 99–115, 1990.

MELATI, C.; JANISSEK-MUNIZ, R. **Inovação e inteligência artificial na gestão pública brasileira.** Revista Brasileira de Gestão Pública, v. 14, n. 2, p. 233-256, 2022.

MOTA, A. et al. **Uso de Redes Neurais Convolucionais para Segmentação Automática de Documentos Jurídicos.** Brazilian Symposium on Information Systems, p. 345-356, 2020.

OLOWONOMI, A. et al. **Application of machine learning algorithms for the prediction of clinical outcomes.** Information Sciences Letters, v. 9, n. 2, p. 141-148, 2021.

OLIVEIRA, R. S. et al. **Predicting the number of days in court cases using artificial intelligence.** PLoS ONE, v. 17, n. 3, e0266574, 2022.

OLIVEIRA, M. P.; CUNHA, J. A. **Tratamento de outliers em modelos preditivos.** Revista Brasileira de Estatística Aplicada, v. 7, n. 2, p. 150-165, 2020.

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

ROSENBLATT, F. **The perceptron: a probabilistic model for information storage and organization in the brain.** Psychological Review, v. 65, n. 6, p. 386-408, 1958.

SABO, I. C.; ROVER, A. J. **Compliance to legal precedents and repetitive legal**

cases management with machine learning. Revista Opiniao Juridica, v. 18, n. 28, p. 69–93, 1 maio 2020.

SALOMÃO, L. F. et al. **Justiça em Números 2023.** Brasília: Conselho Nacional de Justiça, 2023.

SAMPIERI, R. H.; COLLADO, C. F.; LUCIO, M. P. B. **Metodologia de pesquisa.** 5. ed. Porto Alegre: Penso, 2014.

SCHWAB, K. **A Quarta Revolução Industrial.** São Paulo: Edipro, 2016.

SILVA, R. O. et al. **Predição de tempo restante para conclusão de processos de negócio utilizando aprendizado profundo.** Anais do XXXVIII Simpósio Brasileiro de Bancos de Dados, p. 125-136, 2023.

TURING, A. M. **Computing machinery and intelligence.** Mind, v. 59, p. 433-460, 1950.

WANG, P.; HAFSHEJANI, B. A.; WANG, D. **An improved multilayer perceptron approach for detecting sugarcane yield production in IoT based smart agriculture.** Microprocessors and Microsystems, v. 82, pp. 1-7, 2021.

3. PTT: MODELO PREDITIVO PARA PRAZOS PROCESSUAIS NO SUPERIOR TRIBUNAL DE JUSTIÇA

3.1. Introdução

O produto decorrente deste projeto é um algoritmo, construído com a utilização de técnicas de modelagem de redes neurais: o *Multi-Layer Perceptron (MLP)*, implementado através da biblioteca *scikit-learn*, na linguagem de programação *Python*, para prever a duração de um processo judicial assim que ele ingressar no **Superior Tribunal de Justiça (STJ)**, de forma originária ou recursal.

Institucionalmente no STJ, bem como no âmbito do Poder Judiciário Brasileiro, não existem soluções em uso para a predição do tempo de tramitação de processos judiciais com técnicas de inteligência artificial, conforme se observa na Pesquisa uso de inteligência artificial (IA) no Poder Judiciário (CNJ; PNUD, 2023), em que pese as recentes pesquisas que abordaram o tema.

Segundo a literatura, o processo de inovação pode ser incremental, com pequenas mudanças graduais; radical, com novos produtos ou serviços ou mudanças significativas em seu modo de oferta e produção e, por último, inovações sistemáticas, grandes mudanças resultantes de, por exemplo, introdução de novas tecnologias. (BEKKERS; EDELENBOS; STEIJIN, 2011).

Nesse contexto, o PTT resultante deste projeto pode ser compreendido como uma inovação incremental em relação à prática institucional de estimar prazos, ao passo que representa uma inovação técnica mais profunda por introduzir o uso de redes neurais para fins preditivos — abordagem ainda inédita no Poder Judiciário brasileiro.

Essa nova ferramenta analítica poderá ser utilizada para refinar a gestão processual e auxiliar na otimização da alocação de recursos e no planejamento estratégico do STJ, refletindo a interseção entre ciência de dados, inteligência artificial e gestão de processos judiciais, ao possibilitar que seja demonstrada a duração de um processo recém-ingresso no Tribunal, para que a Administração atue preventivamente naqueles processos que tiverem uma previsão de duração da sua tramitação longa.

Por outro lado, a pesquisa contribui academicamente ao explorar a aplicação de redes neurais para predição do tempo de tramitação de um processo judicial, um território ainda pouco explorado, apesar da inteligência artificial ser um assunto relevante e que tem sido amplamente discutido no meio acadêmico e no contexto sociopolítico brasileiro. Além disso, ela amplia o conhecimento sobre métodos de previsão e análise de dados no contexto jurídico, com a utilização de técnicas de inteligência artificial.

Sob a ótica social, o estudo aborda um problema atual e significativo: a eficiência do sistema judiciário. Ao ampliar o conhecimento acadêmico sobre mecanismos de previsão da duração dos processos judiciais, poderá contribuir para uma justiça mais ágil e efetiva, de forma a suprir a necessidade da sociedade por um sistema judiciário mais eficaz e que atenda de forma mais célere suas demandas.

3.2. Descrição geral do produto

3.2.1 Metodologia de Pesquisa

A pesquisa adota uma abordagem quantitativa exploratória e explicativa, com o objetivo de analisar padrões e prever o tempo de tramitação de processos judiciais no Superior Tribunal de Justiça (STJ) por meio de aprendizado de máquina e redes neurais, o que permite identificar padrões de tramitação processual e prever a duração de processos judiciais, por meio de correlações não lineares.

Sampieri e Collado (2014) e Creswell (2021) afirmam que pesquisas quantitativas podem ser classificadas em descritivas, correlacionais, experimentais e exploratórias, dependendo da forma como os dados são utilizados para responder às perguntas da pesquisa.

No contexto deste estudo, a abordagem exploratória permite mapear padrões ocultos nos dados processuais, enquanto a abordagem explicativa busca evidenciar o grau de influência entre diferentes variáveis judiciais e a duração dos processos, para garantir clareza na formulação das variáveis, replicabilidade do estudo e validade estatística dos resultados. Essa metodologia possibilita que o algoritmo da rede neural treinada possa ser reaplicado e adaptado para diferentes tribunais, consolidando a utilização da inteligência artificial como ferramenta auxiliar na modernização do sistema judiciário.

As variáveis foram extraídas do Painel Justiça em Números do CNJ, disponível para acesso público. A variável dependente do estudo é o tempo de tramitação do processo (medido em dias), enquanto as variáveis independentes incluem código da classe processual, código do órgão responsável, frequência de decisões de cada magistrado no tribunal e características processuais associadas, não sendo consideradas variáveis mediadoras ou moderadoras, pois a análise se concentra na modelagem preditiva direta das variáveis que podem influenciar a duração dos processos.

O procedimento metodológico empregado neste estudo está alinhado com pesquisas anteriores que utilizaram aprendizado de máquina para previsão de tempos processuais e modelos estatísticos, como os trabalhos de Oliveira et al. (2022), Silva et al. (2023) e Araújo et al. (2021). Os quais demonstram que modelos estatísticos aplicados ao aprendizado de máquina são capazes de prever a tramitação processual com alto grau de precisão, fornecendo subsídios para a gestão dos tribunais.

No mesmo sentido, a Pesquisa sobre o Uso de Inteligência Artificial no Poder Judiciário (CNJ, 2023) revelou que 62 tribunais brasileiros já utilizam aprendizado de máquina para previsão processual, sendo que 87% desses projetos aplicam abordagens supervisionadas, como a utilizada nesta pesquisa. Isso reforça a validade da abordagem quantitativa exploratória e explicativa e demonstra que o aprendizado de máquina é um método aceito e aplicado no sistema judiciário brasileiro.

Para responder às questões de pesquisa propostas, a metodologia foi estruturada em três etapas principais: coleta e tratamento dos dados, treinamento e validação da rede neural e análise e interpretação dos resultados.

A primeira etapa consistiu na coleta e tratamento dos dados, extraídos diretamente do Painel Justiça em Números do CNJ, utilizando o conjunto de bibliotecas específicas da linguagem de programação *python*. Seguindo a metodologia similar a utilizada por Cruz et al. (2024), foram analisados 631.437 processos baixados entre janeiro de 2023 e junho de 2024, independentemente da data de entrada desse processo no STJ, garantindo um conjunto de dados numericamente significativo para o treinamento da rede neural. Essa etapa teve como principal objetivo mitigar o risco de que o modelo fosse treinado com dados inconsistente, não padronizados ou insuficientes.

Na segunda etapa, foi realizado o treinamento e validação do modelo preditivo, utilizando uma rede neural do tipo *Multi-Layer Perceptron (MLP)*, treinada por meio de validação cruzada e *RandomizedSearchCV* para ajuste de hiperparâmetros. A escolha da rede neural MLP baseia-se na literatura especializada, que destaca sua eficiência na modelagem de dados complexos e na previsão de séries temporais, como observado nos estudos de Gruginskie & Vaccaro (2018) e Mota et al. (2020).

O processo de validação teve como objetivo permitir, com certo grau de confiabilidade, que o modelo tivesse capacidade de generalização e evitar problemas de sobreajuste (*overfitting*), de forma a possibilitar que as previsões fossem aplicáveis a novos processos judiciais.

Na terceira etapa, foi realizada a análise e interpretação dos resultados, utilizando as métricas estatísticas coeficiente de determinação (R^2) e erro absoluto médio (MAE). O R^2 indica o percentual da variação dos dados explicada pelo modelo, enquanto o MAE mede a diferença média entre as previsões e os valores reais.

Além dessa análise estatística, foi aplicada a técnica de SHAP *values* (*SHapley Additive exPlanations*) para interpretar a influência das variáveis que possuem mais impacto na previsão da tramitação processual. Os SHAP *values* permite quantificar a contribuição individual de cada variável na previsão gerada pelo modelo.

Dessa forma, a abordagem segue a metodologia utilizada por Lundberg e Lee (2017), que propuseram os SHAP *values* como ferramenta para interpretar modelos de aprendizado de máquina de forma compreensível e explicável. No contexto deste estudo, essa análise possibilita entender quais fatores mais impactaram a previsão do tempo de tramitação dos processos judiciais no STJ, dentre as variáveis utilizados para treinamento da rede neural.

3.2.2 Coleta de Dados

Inicialmente, procedeu-se à coleta de dados, no mês de agosto de 2024, através do menu "*downloads*" do Painel Justiça em Números, fornecido pelo Conselho Nacional de Justiça (CNJ), disponível em <https://painel-estatistica.stg.cloud.cnj.jus.br/estatisticas.html>. A seleção deliberada de "STJ" no campo "Tribunal" e de "Baixados" no campo "Indicador" foi realizada para concentrar

a análise em processos com ciclo completo de tramitação, o que permite uma avaliação abrangente da eficiência processual no âmbito do STJ.

A base de dados adquirida abrange um total de 631.437 (seiscentos e trinta e um mil e quatrocentos e trinta e sete) informações de processos judiciais baixados, cujo marco temporal utilizado pelo CNJ foi a data de conclusão nos anos de 2023 ou 2024, independentemente do ano do início de sua tramitação, compreendendo variáveis como: *ano, mês, data de início, data de referência, tribunal, nome do órgão, código do órgão, id do município, município, UF, grau, formato, procedimento, recurso originário, processo, códigos de assuntos, códigos de classes, código da última classe e nome da última classe*.

A inclusão de variáveis multifacetadas, como códigos de assuntos e código de classes, permitiu explorar a complexidade e a variabilidade dos processos, aspectos essenciais ao treinamento de um modelo preditivo com base em técnicas de *deep learning*, especialmente adequado para analisar dados de alta dimensionalidade e extrair padrões latentes (KAPLAN e HAENLEIN, 2019; OLOWONOMI *et al.* 2021).

A utilização de processos baixados, definidos pelo CNJ como aqueles que alcançaram uma conclusão dentro do sistema judiciário, seja por julgamento, arquivamento ou outro mecanismo legal, se deve à sua capacidade de fornecer dados completos sobre o ciclo de vida dos processos. A análise desses dados reflete a relevância do *Big Data* na otimização de processos institucionais, em linha com as transformações impulsionadas pela era da informação no setor judiciário (MAYER-SCHONBERGER e CUKIER, 2013).

3.2.3 Preparação do ambiente de produção

Para a preparação do ambiente de produção do *Google Colaboratory*, com todas as ferramentas necessárias para construção do algoritmo, foram importadas as seguintes bibliotecas:

- ***Pandas***: manipulação e análise de dados, permitindo a criação, leitura e manipulação de estruturas de dados como *DataFrames*;
- ***Numpy***: fornece suporte para *arrays* e matrizes, bem como uma coleção de funções matemáticas de alto desempenho;

- **Matplotlib:** criação de gráficos e visualizações de dados;
- **Scikit-learn (sklearn):** biblioteca para aprendizado de máquina que fornece uma gama de ferramentas para modelagem, pré-processamento, validação e avaliação de modelos, tais como:
 - **train_test_split:** Para dividir os dados em conjuntos de treino e teste;
 - **MLPRegressor:** modelo de rede neural supervisionada;
 - **RandomizedSearchCV:** utilizado para otimizar os hiperparâmetros do modelo;
 - **cross_val_score:** realiza validação cruzada e avalia o desempenho do modelo;
 - **StandardScaler:** normaliza os dados e garante que todas as características contribuam de forma equitativa; e
 - **SimpleImputer:** lida com valores ausentes no conjunto de dados.
- **Joblib:** fornece ferramentas para salvar e carregar objetos em *python* de forma eficiente, especialmente objetos grandes como modelos treinados.

As importações mencionadas estão demonstradas na Figura 3.1.

Figura 3.1 - Preparação do Ambiente de Produção

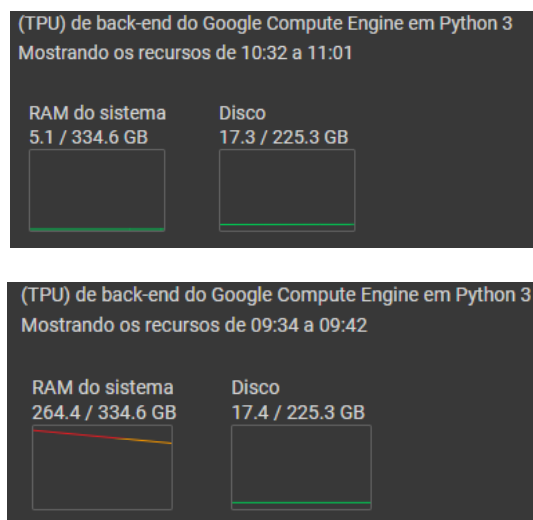
```
1 import pandas as pd
2 from sklearn.model_selection import train_test_split
3 from sklearn.impute import SimpleImputer
4 from sklearn.preprocessing import StandardScaler, OneHotEncoder
5 from sklearn.compose import ColumnTransformer
6 from sklearn.pipeline import Pipeline
7 from sklearn.neural_network import MLPRegressor
8 from sklearn.model_selection import RandomizedSearchCV
9 from sklearn.metrics import r2_score, mean_absolute_error
10 import numpy as np
11 import matplotlib.pyplot as plt
12 import joblib
```

Fonte: elaborado pelo autor

Considerando a necessidade de alto processamento de dados, foi necessária a utilização de uma configuração mais robusta do ambiente de produção em nuvem do

Google Colaboratory e mesmo assim não foi suficiente para utilizar todo o conjunto de dados coletados, conforme Figura 3.2.

Figura 3.2 - Capacidade de Processamento



Fonte: elaborado pelo autor

3.2.4 Carregamento, limpeza e preparação dos dados

Nesta fase, foram implementadas uma série de procedimentos de carregamento, preparação e limpeza dos dados, para assegurar a qualidade e a integridade do conjunto de informações que serão utilizadas para treinar a rede neural. Para as atividades de limpeza dos dados, foram utilizadas as bibliotecas especializadas como *Pandas* e *Numpy*.

Inicialmente, os dados foram carregados a partir de um arquivo *.csv*:

```
data = pd.read_csv('/content/drive/MyDrive/Mestrado/STJ_TBaix.csv',
encoding='ISO-8859-1', delimiter=';', on_bad_lines='skip')
```

A escolha do *encoding='ISO-8859-1'* garante a correta leitura de caracteres especiais, enquanto *delimiter=';'* especifica que os dados são separados por ponto e vírgula. O parâmetro *on_bad_lines='skip'* permite ignorar linhas problemáticas e evitar interrupções no carregamento.

Quanto à capacidade de processamento dos dados do ambiente de produção do *Google Colaboratory*, considerando o limite disponibilizado de 334 GB (trezentos e trinta e quatro gigabytes) de memória *RAM*, foi definido como escopo uma amostra aleatória de 10% dos dados do conjunto:

```
percentual = 0.10 # seleciona 10% da base de dados
```

```
X = X.sample(frac=percentual, random_state=42)
```

```
y = y[X.index] # Manter as mesmas amostras de y
```

Ao longo do treinamento do algoritmo, foram consumidos aproximadamente 300 GB (trezentos *gigabytes*) dessa memória. Para entender e confirmar a estrutura dos dados carregados, foi utilizada a função `print(data.info())`, a qual retornou o resultado constante da Figura 3.3:

Figura 3.3 - Estrutura dos Dados

```
Dados carregados com sucesso.
Total de linhas: 631437
Estrutura inicial dos dados:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 631437 entries, 0 to 631436
Data columns (total 20 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   I»çAno                                631437 non-null  int64
1   Mes                                    631437 non-null  int64
2   Data de inicio                         631437 non-null  object
3   Data de Referencia                     631437 non-null  object
4   Tribunal                              631437 non-null  object
5   Recurso                                631437 non-null  int64
6   Nome orgao                             631437 non-null  object
7  Codigo orgao                             631437 non-null  int64
8   id_municipio                           631437 non-null  int64
9   Municipio                              631437 non-null  object
10  UF                                      631437 non-null  object
11  Grau                                    631437 non-null  object
12  Formato                                631437 non-null  object
13  Procedimento                           631437 non-null  object
14  Recurso Originario                     631437 non-null  object
15  Processo                                631437 non-null  object
16  Codigos assuntos                       631437 non-null  object
17  Codigos Classes                        631437 non-null  object
18 Codigo da Ultima classe CN               631437 non-null  int64
19  Nome da ultima classe CN               631437 non-null  object
dtypes: int64(6), object(14)
memory usage: 96.3+ MB
None
```

Fonte: elaborado pelo autor

Com referência à estrutura dos dados, Figura 3.3, foram encontradas 20 (vinte) colunas, com 631.437 (seiscentos e trinta e um mil e quatrocentos e trinta e sete) linhas, cada linha corresponde a 1 (um) processo. Pontua-se que a coluna “*Data de Referência*” corresponde à data em que o processo mudou seu *status* para “*Baixado*”.

Em seguida, foi feita a conversão das colunas com datas por meio da função `pd.to_datetime`, para permitir a manipulação e o cálculo de intervalos entre essas datas, por exemplo a duração da tramitação de cada processo.

Na sequência, após a análise da estrutura das colunas listadas na Figura 3.3, optou-se pela exclusão de algumas colunas que não continham variabilidade de dados que fornecesse informação útil ao treinamento da rede neural:

- **Tribunal:** nessa coluna consta o mesmo dado em todas as linhas “*Superior Tribunal de Justiça*”;
- **Recurso:** 99% dos processos apresentavam a situação “01”;
- **Município:** nessa coluna consta o mesmo dado em todas as linhas “*Brasília*”;
- **id_município:** nessa coluna consta o mesmo dado em todas as linhas “743”;
- **UF:** nessa coluna consta o mesmo dado em todas as linhas “*DF*”;
- **Grau:** nessa coluna consta o mesmo dado em todas as linhas “*SUP*”;
- **Nome da última classe CN:** essa informação não está disponível no início da tramitação do processo, o que a torna ineficaz para previsão do tempo de tramitação de um processo; e
- **Ano e Mês:** suas informações são derivadas diretamente da coluna “*Data de início*”, que já contém os dados relevantes em um formato mais completo. A manutenção de colunas redundantes pode confundir o modelo e aumentar a dimensionalidade desnecessariamente.

Com referência às colunas “*Código de Assuntos*” e “*Código de Classes*”, sem entrar no mérito do enquadramento e significado de cada código, foi constatado que alguns processos possuem associação com mais de um código em cada uma dessas colunas.

Por isso, essas colunas foram expandidas, utilizando o método *str.get_dummies()*, para criar colunas adicionais que representam cada código de forma independente, com vistas a possibilitar uma análise mais aprofundada, permitindo que cada um seja tratado como uma característica distinta no modelo preditivo.

Assim, após a expansão, as novas colunas codificadas foram concatenadas ao *DataFrame* original e as colunas originais “*Código de Assuntos*” e “*Código de Classes*” foram removidas para evitar redundância, Figura 3.4.

Figura 3.4 - Expansão dos Códigos de Assuntos e de Classes

```

37 # Usar str.get_dummies() para criar as colunas one-hot
38 codigos_assuntos_encoded = data['Codigos assuntos'].str.strip('{}').str.split(',').str.join('|').str.get_dummies()
39 codigos_classes_encoded = data['Codigos Classes'].str.strip('{}').str.split(',').str.join('|').str.get_dummies()
40
41 # Concatenar as colunas codificadas ao DataFrame original
42 data_expanded = pd.concat([data, codigos_assuntos_encoded, codigos_classes_encoded], axis=1)
43
44 # Remover as colunas originais
45 data_expanded.drop(columns=['Codigos assuntos', 'Codigos Classes'], inplace=True)

```

Fonte: elaborado pelo autor

Para calcular a duração do processo e armazená-la numa nova coluna, foi feito o cálculo de dias entre as duas datas presentes nas colunas “*Data de Início*” e “*Data de Referência*”, o resultado foi armazenado numa nova coluna: “*Dias até Data de Referência*”.

Em seguida, com intuito de obter variações da duração com base na sazonalidade, foram extraídas **novas variáveis** da coluna “*Data de Início*”: ***ano do início, mês do início e dia da semana do início***. Essas novas variáveis foram úteis para identificar padrões temporais na duração dos processos e, cada uma, representa uma nova coluna no *DataFrame*, Figura 3.5.

Figura 3.5 - Variáveis de Sazonalidade

```

47 # Calcular a duração do processo
48 data_expanded['Dias até Data de Referência'] = (data_expanded['Data de Referência'] - data_expanded['Data de início']).dt.days
49
50 # Extração de features sazonais e categóricas
51 data_expanded['Ano_Início'] = data_expanded['Data de início'].dt.year
52 data_expanded['Mes_Início'] = data_expanded['Data de início'].dt.month
53 data_expanded['Dia_Semana_Início'] = data_expanded['Data de início'].dt.dayofweek

```

Fonte: elaborado pelo autor

Da mesma forma, para enriquecer o conjunto de dados, foi realizada a combinação entre as colunas “*Código Órgão*” e “*Procedimento*”, criando uma coluna chamada “*Código órgão_Procedimento*”. Para tanto, foi utilizada a função *astype(str)*, garantindo que os valores da coluna “*Código Órgão*” sejam convertidos do formato numérico para textual antes da concatenação.

Isso é importante porque, se o código do órgão fosse um número (por exemplo, um inteiro), a operação de concatenação não funcionaria corretamente. A nova coluna é gerada ao concatenar o código do órgão com o procedimento, separados por um sublinhado “_”. Por exemplo, se código órgão for “12728” e o procedimento for “*Conhecimento Criminal*”, a nova coluna resultaria em “12728_*ConhecimentoCriminal*”.

Adicionalmente, entendeu-se como pertinente a utilização da frequência de um

código de órgão, referente à unidade organização responsável pela decisão de baixa do processo, por exemplo “*Presidência*”, como uma das variáveis relevantes para aperfeiçoar o entendimento sobre a dinâmica dos processos judiciais.

A frequência não só ajuda a entender a distribuição e a carga de trabalho, mas também fornece informações valiosas que podem influenciar as previsões do modelo: órgãos com alta frequência podem ter processos que são resolvidos mais rapidamente, enquanto órgãos com baixa frequência podem lidar com casos que exigem mais tempo ou recursos, por exemplo.

Por isso, com a utilização da função `value_counts(normalize=True)`, foram contadas quantas vezes cada valor aparece na coluna “Codigo orgao” e retorna a proporção (ou frequência relativa) em relação ao total. Isso é útil para entender a distribuição dos dados e verificar se há órgãos que aparecem com mais frequência.

De forma complementar, a função `map` aplica a série de frequências calculadas a cada entrada da coluna “Código Órgão”, preenchendo a **nova coluna** “*Frequencia_Tribunal*” com os valores correspondentes. Assim, cada linha terá a frequência do respectivo código do órgão, permitindo uma análise mais rica e contextualizada dos dados.

A criação das colunas “*Código órgão_Procedimento*” e “*Frequência_Tribunal*” foi implementada com demonstrado na Figura 3.6.

Figura 3.6 - Criação das Colunas “Código órgão_Procedimento” e “Frequência_Tribunal”

```
55 # Criar a nova coluna de combinação
56 data_expanded['Codigo orgao_Procedimento'] = data_expanded['Codigo orgao'].astype(str) + "_" + data_expanded['Procedimento']
57 data_expanded['Frequencia_Tribunal'] = data_expanded['Codigo orgao'].map(data_expanded['Codigo orgao'].value_counts(normalize=True))
```

Fonte: elaborado pelo autor

A partir deste ponto, foi feita a definição das **variáveis preditivas (X)** e a **variável alvo (Y)**. As variáveis preditivas **X** são definidas removendo as colunas que não são necessárias, enquanto a variável alvo **Y** é a duração em dias até a data de referência, conforme Figura 3.7.

Figura 3.7 - Definição das Variáveis X e Y

```
58 # Definir X e y
59 X = data_expanded.drop(columns=['Dias até Data de Referência', 'Nome orgao', 'Processo', 'Data de Referência'])
60 y = data_expanded['Dias até Data de Referência']
```

Fonte: elaborado pelo autor

Após a definição das variáveis preditivas (**X**) e da variável alvo (**Y**), foi

selecionada uma amostra representativa dos dados, a ser utilizada para treinar o modelo. Para isso, considerando a capacidade de processamento do ambiente de produção, foi selecionada uma **amostragem aleatória de aproximadamente 10% do *DataFrame***, o qual passou a contar com 50.515 (cinquenta mil e quinhentos e quinze) linhas e 3.293 (três mil e duzentos e noventa e três) colunas.

A função *sample* foi utilizada para realizar a amostragem, onde *frac=percentual* especifica a proporção dos dados a ser selecionada, e *random_state=42* garante a reprodutibilidade dos resultados. A variável *Y* é ajustada para incluir apenas os índices que correspondem à amostra selecionada em *X*. Em seguida, os dados são divididos em conjuntos de treino e teste, utilizando a função *train_test_split*. Aqui, *test_size=0.2* indica que **20% dos dados serão reservados para teste**, enquanto **80% serão utilizados para treinamento**. O parâmetro *random_state=42* novamente assegura que a divisão dos dados seja reprodutível, Figura 3.8.

Figura 3.8 - Escolha e Divisão da Amostra

```
66 # Selecionar aleatoriamente um percentual do DataFrame para treinar o modelo
67 percentual = 0.10 # Por exemplo, 10%
68 X = X.sample(frac=percentual, random_state=42)
69 y = y[X.index] # Manter as mesmas amostras de y
70
71 # Dividir os dados em conjuntos de treino e teste
72 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Fonte: elaborado pelo autor

3.2.5 Pré-processamento dos Dados

Após a divisão dos dados nas variáveis *X* e *Y*, deu-se início à etapa de pré-processamento, que visa garantir que os dados estejam em um formato adequado para o treinamento do modelo. Para isso, foram identificadas as características numéricas e textuais (categóricas) do conjunto de dados. As listas *numeric_features* e *categorical_features* contêm os nomes das colunas correspondentes a dados numéricos e categóricos, respectivamente, Figura 3.9.

Figura 3.9 - Definição do Pipeline

```
78 # Pipeline para tratamento de dados
79 numeric_features = X.select_dtypes(include=[np.number]).columns.tolist()
80 categorical_features = X.select_dtypes(include=['object']).columns.tolist()
```

Fonte: elaborado pelo autor

Um pré-processador é então definido, utilizando *ColumnTransformer* para aplicar diferentes transformações às colunas de acordo com seu tipo, Figura 3.10.

Figura 3.10 - Pré-Processamento

```

82 # Definir o pré-processador
83 preprocessor = ColumnTransformer(
84     transformers=[
85         ('num', SimpleImputer(strategy='mean'), numeric_features),
86         ('cat', OneHotEncoder(handle_unknown='ignore'), categorical_features + ['Codigo orgao_Procedimento']) # Adicionando a nova coluna
87     ])

```

Fonte: elaborado pelo autor

No código apresentado na Figura 3.10, as colunas numéricas são tratadas com *SimpleImputer*, que substitui valores ausentes pela média da coluna e as colunas textuais (categóricas) são convertidas em variáveis *dummy* (*one-hot encoding*) com *OneHotEncoder*, o que é essencial para transformar dados textuais em um formato que a máquina possa processar. O parâmetro *handle_unknown='ignore'* assegura que categorias que não estavam presentes durante o treinamento não gerem erros.

Para exemplificar o funcionamento do *One-Hot Encoding*, considere a Tabela 3.1.

Tabela 3.1 – Dados Categóricos Originais

Codigo orgao	Tribunal	Procedimento
12728	Primeira Turma	Conhecimento Criminal
12728	Segunda Turma	Recurso Especial
12728	Primeira Turma	Conhecimento Criminal
12580	Terceira Turma	Recurso Ordinário
12580	Terceira Turma	Conhecimento Criminal

Fonte: elaborado pelo autor

Ao aplicar a funcionalidade do *One-Hot Encoding*, nas colunas “Tribunal” e “Procedimento”, obtém-se a Tabela 3.2:

Tabela 3.2 - Aplicação do One-Hot Encoding

Co-digo orgao	Tribu-nal_Pri-meira_Turma	Tribunal_Se-gunda_Turma	Tribu-nal_Ter-ceira_Turma	Procedi-mento_Co-nheci-mento Cri-minal	Procedi-mento_Re-curso Es-pacial	Procedi-mento_Re-curso Or-dinário
12728	1	0	0	1	1	0
12728	0	1	0	0	1	0
12728	1	0	0	1	0	0
12580	0	0	1	0	0	1
12580	0	0	1	1	0	0

Fonte: elaborado pelo autor

Como é possível perceber, *One-Hot Encoding* é utilizado para transformar variáveis textuais em um formato que pode ser processado em linguagem computacional. Cria-se uma coluna para cada categoria, inserindo os valores numéricos **0** ou **1**, na qual:

- **1** indica que a categoria correspondente está presente na amostra; e
- **0** indica que a categoria correspondente não está presente na amostra.

3.2.6 Criação do Pipeline e Treinamento do Modelo

O pipeline definido na Figura 3.9 é criado para integrar o pré-processamento e o modelo em um único objeto, Figura 3.11.

Figura 3.11 - Criação do Pipeline

```
89 # Criar um pipeline que inclui o pré-processador e o modelo
90 pipeline = Pipeline(steps=[
91     ('preprocessor', preprocessor),
92     ('scaler', StandardScaler()), # Normalizar os dados após o pré-processamento
93     ('model', MLPRegressor(random_state=42))
94 ])
```

Fonte: elaborado pelo autor

O *pipeline*, portanto, realiza as seguintes etapas:

1. **Pré-processamento dos dados** (tratamento de valores ausentes e codificação das variáveis categóricas);
2. **Normalização dos dados numéricos** com *StandardScaler*, que padroniza as características removendo a média e escalando para a unidade de variância; e
3. **Treinamento do modelo** utilizando *MLPRegressor*, que implementa a rede neural do tipo *Multi-Layer Perceptron*.

3.2.7 Ajuste de Hiperparâmetros do Modelo MLP

O *MLPRegressor* dispõe de diversos hiperparâmetros passíveis de serem utilizados e ajustados conforme a especificidade do modelo a ser desenvolvido. Abaixo, conforme documentação da referida biblioteca, foi elaborada a Tabela 3.3 contendo o hiperparâmetro, uma breve explicação e o valor padrão (*default*) de cada parâmetro disponível no modelo:

Tabela 3.3 - Definição de Hiperparâmetros do MLPRegressor

Hiperparâmetro	Descrição	Padrão
----------------	-----------	--------

hidden_layer_sizes	Número de camadas ocultas e de neurônios por camada oculta	100
activation	Função de ativação para a camada oculta: identity : ativação no-op, útil para implementar gargalo linear, retorna $f(x) = x$. logistic : a função sigmóide logística, retorna $f(x) = 1 / (1 + \exp(-x))$. tanh : a função tan hiperbólica, retorna $f(x) = \tanh(x)$. relu : a função de unidade linear retificada, retorna $f(x) = \max(0, x)$.	relu
solver	O solucionador para otimização de peso: lbfgs : um otimizador da família de métodos quase-Newton. sgd : refere-se à descida gradiente estocástica. adam : refere-se a um otimizador estocástico baseado em gradiente proposto por Kingma, Diederik e Jimmy Ba.	adam
alpha	Força do termo de regularização L2. O termo de regularização L2 é dividido pelo tamanho da amostra quando somado à perda.	1
batch_size	Tamanho dos minilotes para otimizadores estocásticos. Se o solucionador for 'lbfgs', o regressor não usará minibatch. Quando definido como "auto", $\text{batch_size} = \min(200, n_samples)$.	auto
learning_rate	Cronograma de taxa de aprendizagem para atualizações de peso: constant : uma taxa de aprendizagem constante. invscaling : diminui a taxa de aprendizagem a cada passo de tempo usando um expoente de escala inverso. adaptive : mantém a taxa de aprendizagem constante enquanto a perda de treinamento continuar diminuindo.	constant
learning_rate_init	A taxa de aprendizagem inicial usada. Controla o tamanho do passo na atualização dos pesos.	1
power_t	O expoente para escala inversa da taxa de aprendizagem. Usado na atualização da taxa de aprendizagem efetiva quando o learning_rate está definido como 'invscaling'.	0.5
max_iter	Número máximo de iterações. O solucionador itera até a convergência ou este número de iterações. Para solucionadores estocásticos, isso determina o número de épocas.	200
shuffle	Se as amostras devem ser embaralhadas em cada iteração. Usado apenas quando solver='sgd' ou 'adam'.	TRUE
random_state	Determina a geração de números aleatórios para inicialização de pesos e bias, divisão de teste de treino, e amostragem em lote. Passar um int para resultados reproduzíveis.	None
tol	Tolerância para a otimização. Quando a perda ou pontuação não melhora pelo menos tol em iterações consecutivas, a convergência é considerada alcançada e o treinamento é interrompido.	1,00E-04
verbose	Se deve imprimir mensagens de progresso.	FALSE
warm_start	Quando definido como True, reutiliza a solução da chamada anterior como inicialização.	FALSE
momentum	Momentum para atualização de descida gradiente. Deve estar entre 0 e 1. Usado somente quando solver='sgd'.	0.9
nesterovs_momentum	Seja para usar o impulso de Nesterov. Usado apenas quando solver='sgd' e momentum > 0.	TRUE
early_stopping	Se deve usar a parada antecipada para encerrar o treinamento quando a pontuação de validação não está melhorando.	FALSE
validation_fraction	A proporção de dados de treinamento a serem reservados como conjunto de validação para parada antecipada. Deve estar entre 0 e 1.	0.1
beta_1	Taxa de decaimento exponencial para estimativas do primeiro momento vetor no adam. Deve estar em [0, 1).	0.9
beta_2	Taxa de decaimento exponencial para estimativas do segundo momento vetor no adam. Deve estar em [0, 1).	0.999

<i>epsilon</i>	Valor para estabilidade numérica no adam.	1,00E-08
<i>n_iter_no_change</i>	Número máximo de épocas sem melhoria para interromper o treinamento. Usado apenas quando solver='sgd' ou 'adam'.	10
<i>max_fun</i>	Usado apenas quando solver='lbfgs'. Número máximo de chamadas de função.	15000

Fonte: elaborado pelo autor

A coluna “*Padrão*” da Tabela 3.3 corresponde às configurações pré-estabelecidas do modelo, isso significa que qualquer usuário ao desenvolver o algoritmo não altere expressamente esses valores, serão utilizados os parâmetros pré-estabelecidos pelo modelo. Assim, optou-se por utilizar a configuração padrão do modelo para a maioria dos hiperparâmetros, exceto para os constantes na Tabela 3.4 e Figura 3.12, nos quais foi aplicado o método ***RandomizedSearchCV***, que realiza uma busca aleatória para encontrar as melhores combinações de hiperparâmetros.

Tabela 3.4 - Seleção de Hiperparâmetros do *MLPRegressor* com *RandomizedSearchCV*

Hiperparâmetro	GridSearchCV	Descrição
<i>hidden_layer_sizes</i>	(50), (100), (50, 50), (50, 100), (100,100)	Define o número e o tamanho das camadas ocultas. Variações comuns incluem uma ou duas camadas ocultas com 50 a 100 neurônios cada.
<i>alpha</i>	0.001 e 0.01	Força da regularização L2 para evitar <i>overfitting</i> . Valores menores podem ser adequados para dados complexos, enquanto valores maiores ajudam a simplificar o modelo.
<i>learning_rate_init</i>	0.001 e 0.01	A taxa de aprendizagem inicial quando solver é 'sgd' ou 'adam'. Valores maiores aceleram o treinamento, mas podem ultrapassar o mínimo.
<i>max_iter</i>	200, 300, 500 e 1000	Número máximo de iterações para o treinamento convergir. Um número maior permite mais ajustes, mas aumenta o tempo de treinamento.

Fonte: elaborado pelo autor

O dicionário *param_dist*, Figura 3.12, linha 106, contém as variáveis a serem testadas, como o número de neurônios nas camadas ocultas, a taxa de aprendizado, a regularização L2 e o número máximo de iterações. Aqui, o modelo é treinado com diferentes combinações de hiperparâmetros, utilizando validação cruzada (*scoring= 'r2'*) para avaliar o desempenho de cada combinação.

Figura 3.12 - Parametrização do RandomizedSearchCV

```

96 # Parametros a serem testados pelo RandomizedSearchCV
97 param_dist = {
98     'model__hidden_layer_sizes': [(50,), (100,), (50, 50), (50, 100), (100, 100)], # Algumas combinações de camadas ocultas
99     'model__alpha': [0.001, 0.01], # Regularização L2
100     'model__learning_rate_init': [0.001, 0.01], # Taxa de aprendizado inicial
101     'model__max_iter': [200, 300, 500, 1000] # Iterações
102 }
103
104 # RandomizedSearchCV: Testa um número fixo de combinações aleatórias
105 random_search = RandomizedSearchCV(estimator=pipe,
106                                     param_distributions=param_dist,
107                                     n_iter=10, # Número de combinações a serem testadas
108                                     cv=5, # Validação cruzada
109                                     n_jobs=-1, # Usar todos os núcleos disponíveis
110                                     scoring='r2',
111                                     random_state=42)

```

Fonte: elaborado pelo autor

Com a aplicação dos parâmetros elencados como mais precisos, a previsão da variável alvo no conjunto de teste é realizada, Figura 3.13.

Figura 3.13 - Treinamento do Modelo com os Hiperparâmetros mais precisos

```

113 # Executar o RandomizedSearchCV
114 random_search.fit(X_train, y_train)

127 # Modelo otimizado
128 best_model = random_search.best_estimator_
129 y_pred_optimized = best_model.predict(X_test)

```

Fonte: elaborado pelo autor

3.2.8 Avaliação do Modelo

Na sequência, o modelo treinado foi avaliado usando o conjunto de teste constantes na Tabela 3.5 e implementados com o código constante na Figura 3.14, para verificar sua capacidade de generalização e de prever dados novos.

Tabela 3.5 - Validação do Desempenho do Modelo

Métrica	Descrição	Como Funciona
Coeficiente de Determinação (R²)	Mede a proporção da variação da variável dependente que é previsível a partir das variáveis independentes.	Varia de 0 a 1, onde 1 indica explicação total da variabilidade dos dados pelo modelo.
Erro Absoluto Médio (MAE)	Média da soma dos valores absolutos das diferenças entre previsões e valores reais.	Valores mais baixos indicam previsões mais precisas, fornecendo uma medida direta do erro médio.

Fonte: elaborado pelo autor

Figura 3.14 - Avaliação do Modelo

```
131 # Avaliação do modelo otimizado
132 r2_optimized = r2_score(y_test, y_pred_optimized)
133 mae_optimized = mean_absolute_error(y_test, y_pred_optimized)
134 print(f'R² Otimizado: {r2_optimized}')
135 print(f'MAE Otimizado: {mae_optimized}')
```

Fonte: elaborado pelo autor

3.2.9 Resultados

Os resultados obtidos com o treinamento do algoritmo sugerem eficácia e assertividade, no que diz respeito à previsão da duração dos processos judiciais no Superior Tribunal de Justiça (STJ), considerando o conjunto de dados testados. A seguir, são apresentadas as principais métricas de desempenho e os resultados das combinações de hiperparâmetros testadas.

3.2.9.1. Descrição da Estrutura do Conjunto de Dados

Para treinamento do modelo, foram utilizados 50.515 (cinquenta mil quinhentos e quinze) registros, abrangendo um total de 3.293 (três mil duzentas e noventa e três) colunas. Essa amostra representa aproximadamente 10% do conjunto de dados original coletado, garantindo que o modelo pudesse ser treinado em um ambiente controlado e manejável, levando em consideração as limitações de capacidade de processamento do ambiente de produção utilizado.

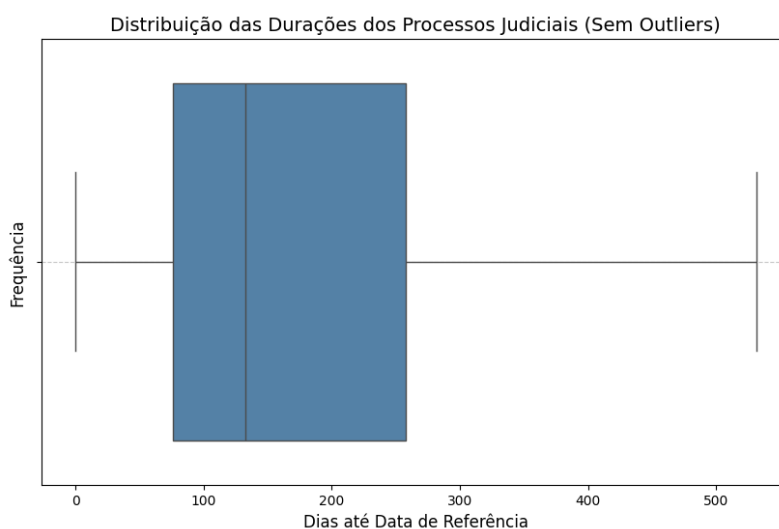
As principais estatísticas descritivas da duração dos processos judiciais em dias são apresentadas a seguir:

- **Média:** 251,43 dias;
- **Desvio Padrão:** 407,06 dias;
- **25º Percentil:** 76 dias;
- **Mediana (50º Percentil):** 133 dias;
- **75º Percentil:** 259 dias;
- **Mínimo:** 0 dias; e
- **Máximo:** 11.221 dias.

Essas métricas fornecem uma visão da distribuição das durações dos processos, indicando que a maioria dos casos é resolvida em média de cerca de 251 dias, enquanto a minoria dos casos apresenta duração excepcionalmente longa.

Para ilustrar a análise, os dados foram aplicados à Figura 3.15, constituindo um *boxplot* da distribuição das durações dos processos judiciais, excluindo-se os *outliers*, podendo ser interpretada à luz das estatísticas descritivas apresentadas.

Figura 3.15 - BOXPLOT Distribuição da Duração dos Processos Judiciais



Fonte: elaborado pelo autor

Observa-se que a **mediana**, situada em **133 dias**, está deslocada para a parte inferior da caixa, confirmando que a maior parte dos processos possui uma tramitação relativamente curta em relação à média de **251,43 dias**. Essa diferença entre a mediana e a média sugere a presença de assimetria positiva, ou seja, um conjunto de processos com durações muito superiores à mediana está impactando a média, tornando-a mais elevada.

O **intervalo interquartil (IQR)**, delimitado pelo **25º percentil (76 dias)** e **75º percentil (259 dias)**, representa a faixa onde estão concentrados **50% dos processos** analisados. Isso significa que, na maior parte dos casos, a tramitação processual varia entre cerca de **dois meses e meio a oito meses e meio**. Essa informação é relevante para modelos preditivos, pois indica que previsões dentro desse intervalo possuem maior probabilidade de acerto.

O **desvio padrão de 407,06 dias** sugere uma dispersão elevada na duração dos processos, indicando que há uma grande variabilidade nos tempos de tramitação. Essa

dispersão é evidenciada pelos valores **mínimo de 0 dias** (provavelmente relacionado a processos que foram resolvidos imediatamente) e **máximo de 11.221 dias** (o equivalente a mais de 30 anos), reforçando a existência de um subconjunto de processos com duração extremamente longa.

A presença dessa **assimetria positiva**, com casos de tramitação excessivamente longa, pode ser explicada por fatores como **complexidade processual, recursos interpostos e variação na carga de trabalho dos magistrados e tribunais**. Esse comportamento é fundamental para a interpretação dos resultados do modelo preditivo, pois sugere que as previsões devem considerar diferentes perfis de tramitação e não apenas a média dos dados.

Além disso, a remoção de outliers foi essencial para obter um retrato mais fiel da distribuição da duração processual sem a interferência de casos extremos, permitindo uma melhor calibração do modelo de previsão e garantindo maior precisão na estimativa do tempo de tramitação dos processos no **Superior Tribunal de Justiça (STJ)**.

3.2.9.2. Desempenho de Hiperparâmetros

A Tabela 3.6 resume o desempenho de cada combinação de hiperparâmetros testada durante o processo de otimização, ordenadas conforme o seu grau de assertividade:

Tabela 3.6 – Desempenho das Combinações de Hiperparâmetros

Combinação de Hiperparâmetros	Média do R²	Desvio Padrão do R²
{'model__max_iter': 200 , 'model__learning_rate_init': 0.001 , 'model__hidden_layer_sizes': (50,), 'model__alpha': 0.001 }	0.927150	0.003116
{'model__max_iter': 200 , 'model__learning_rate_init': 0.01 , 'model__hidden_layer_sizes': (50, 100), 'model__alpha': 0.001 }	0.922999	0.003186
{'model__max_iter': 200 , 'model__learning_rate_init': 0.01 , 'model__hidden_layer_sizes': (100,), 'model__alpha': 0.001 }	0.922152	0.004866
{'model__max_iter': 500 , 'model__learning_rate_init': 0.001 , 'model__hidden_layer_sizes': (100,), 'model__alpha': 0.001 }	0.915588	0.001064
{'model__max_iter': 500 , 'model__learning_rate_init': 0.001 , 'model__hidden_layer_sizes': (50, 50), 'model__alpha': 0.001 }	0.896269	0.004947

{'model__max_iter': 500 , 'model__learning_rate_init': 0.001 , 'model__hidden_layer_sizes': (50 , 100), 'model__alpha': 0.001 }	0.882373	0.005589
{'model__max_iter': 1000 , 'model__learning_rate_init': 0.01 , 'model__hidden_layer_sizes': (50 , 100), 'model__alpha': 0.001 }	0.882373	0.005589
{'model__max_iter': 500 , 'model__learning_rate_init': 0.01 , 'model__hidden_layer_sizes': (50 , 100), 'model__alpha': 0.001 }	0.882373	0.005589
{'model__max_iter': 500 , 'model__learning_rate_init': 0.01 , 'model__hidden_layer_sizes': (50 , 50), 'model__alpha': 0.001 }	0.880978	0.005060
{'model__max_iter': 500 , 'model__learning_rate_init': 0.01 , 'model__hidden_layer_sizes': (50 , 100), 'model__alpha': 0.01 }	0.878742	0.006214

Fonte: elaborado pelo autor

Após a realização da busca por hiperparâmetros com o método *RandomizedSearchCV*, os melhores hiperparâmetros encontrados foram:

- Número de iterações máximas (***max_iter***): 200;
- Taxa de aprendizado inicial (***learning_rate_init***): 0.001;
- Tamanho das camadas ocultas (***hidden_layer_sizes***): (50,); e
- Regularização L2 (***alpha***): 0.001.

Esses parâmetros foram escolhidos com base em sua capacidade de maximizar o desempenho do modelo, refletindo uma otimização adequada para a complexidade dos dados.

3.2.9.3. Descrição das Variáveis (Features)

Para estimar a duração dos processos judiciais no STJ, diversas *features* (variáveis) foram utilizadas. A seguir, detalha-se cada uma delas, considerando sua relevância e impacto no desempenho do modelo.

A variável “*Data de Início do Processo*” captura padrões sazonais e tendências temporais que podem influenciar a duração dos processos. As variáveis derivadas da data de início, como “*Ano_Início*”, “*Mes_Início*” e “*Dia_Semana_Início*”, adicionam camadas de informação temporal relevantes para capturar tendências específicas. Por exemplo, processos iniciados em dias próximos a feriados ou em determinados dias da

semana podem ter durações diferentes devido a fatores operacionais. Como essas informações estão disponíveis no início do processo, seu uso é apropriado.

Já as variáveis “*Códigos de Assuntos*” e “*Códigos de Classes Processuais*” representam os temas jurídicos abordados nos processos. Ao converter esses códigos em variáveis *dummy* por meio da codificação *one-hot*, o modelo captura a influência específica de cada assunto e natureza na duração do processo. Essa granularidade é importante, pois diferentes assuntos e naturezas processuais podem ter complexidades e procedimentos distintos.

A variável “*Código do Órgão*” identifica a unidade judiciária responsável pelo processo. Diferentes órgãos podem ter cargas de trabalho, eficiência e práticas processuais variadas, influenciando o tempo de tramitação. Da mesma forma, a variável “*Procedimento*” permite que o modelo diferencie entre processos que, embora possam ser da mesma classe ou assunto, seguem fluxos processuais diferentes. A codificação adequada dessa variável assegura que o modelo capture essas nuances sem introduzir ruído desnecessário.

A criação da variável combinada “*Código do Órgão_Procedimento*” busca capturar interações específicas entre o órgão e o procedimento adotado, com intuito de revelar padrões que não seriam evidentes ao considerar as variáveis isoladamente. Com lógica similar, a variável “*Frequência do Tribunal*”, calculada como a proporção de processos associados a cada órgão julgador, tem o potencial de refletir a carga de trabalho e a eficiência de cada unidade judiciária.

3.2.9.4. Desempenho do Modelo

O desempenho do modelo foi avaliado por meio de duas métricas principais:

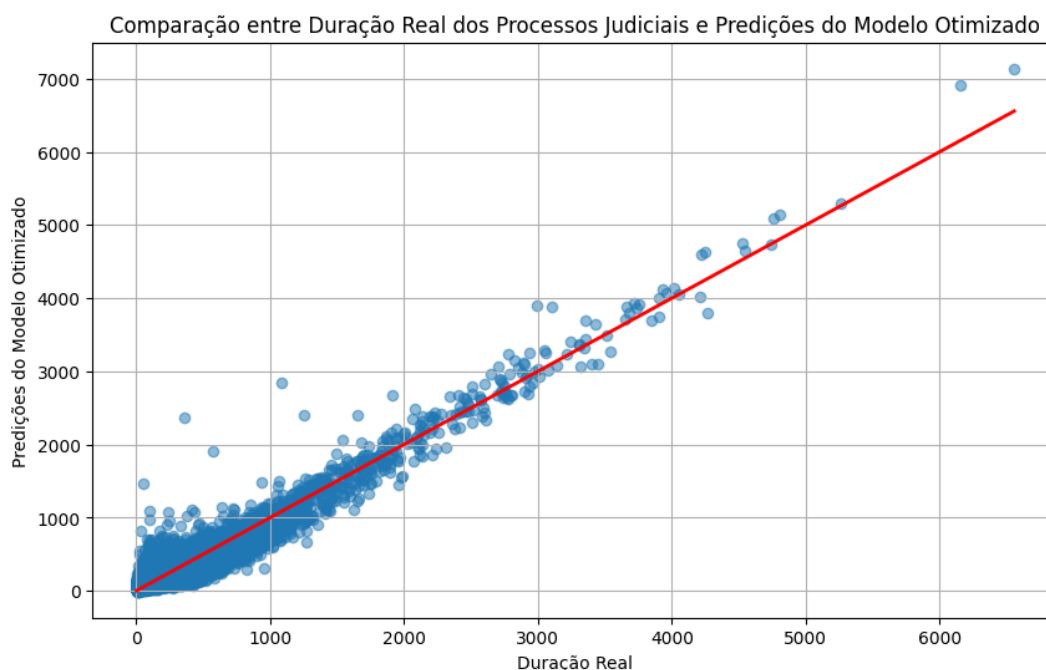
- **R² Otimizado:** 0.9243; e
- **Erro Absoluto Médio (MAE):** 71.61.

O valor do **R²** indica que aproximadamente **92.43% da variabilidade** na duração dos processos judiciais **pode ser explicada pelas variáveis preditivas** selecionadas, o que demonstra uma excelente capacidade de previsão do modelo. O **MAE**, por sua vez, sugere que, em média, **as previsões do modelo estão a cerca de 71 dias da duração**

real dos processos, o que é deve ser ponderado com cautela uma vez que a média de duração dos processos encontrada no conjunto de dados foi de 259 dias, contudo, considerando o contexto judicial, o resultado demonstra-se satisfatório.

A partir dos resultados, foi gerado um gráfico de dispersão contendo como variáveis a predição do modelo com as configurações otimizadas e a duração real dos processos judiciais, conforme Figura 3.16.

Figura 3.16 - Gráfico de Dispersão dos Resultados



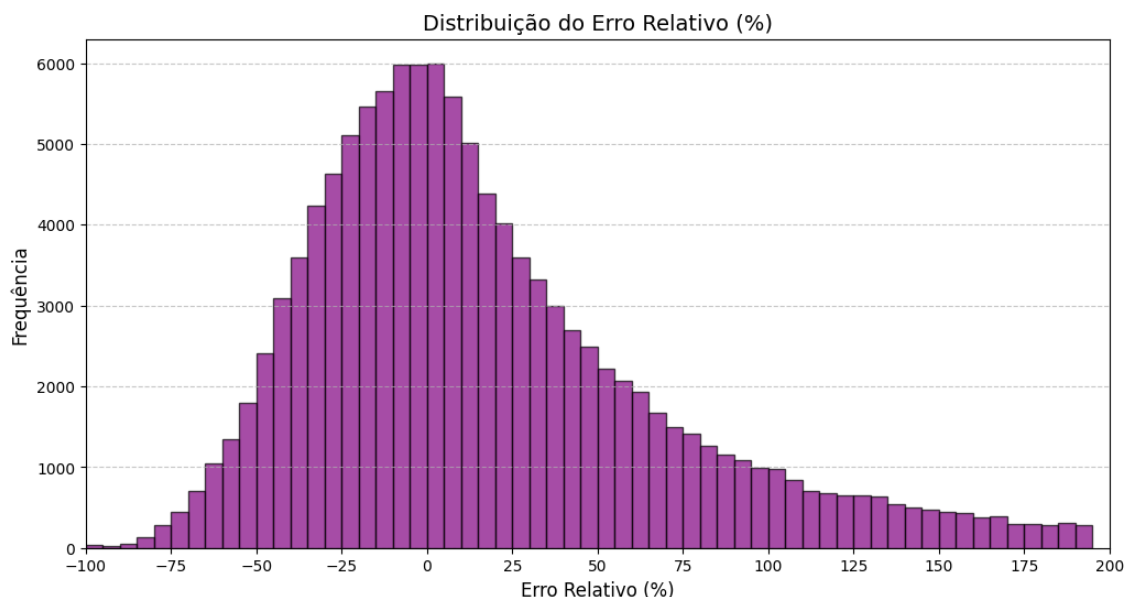
Fonte: elaborado pelo autor

Ao analisar da Figura 3.16, percebe-se que a maioria dos pontos se aproxima da linha vermelha (duração real do processo), isso indica que o algoritmo tem uma boa capacidade de prever a duração dos processos na maior parte dos casos. Apesar disso, ao observar os dados em longas durações - especialmente acima de 3000 dias, nota-se uma dispersão maior dos pontos, sugerindo que o modelo pode ter dificuldades em prever com precisão nesses casos.

O mesmo pode ser percebido na Figura 3.17, ao analisar a distribuição do erro relativo (%) percebe-se uma distribuição assimétrica à direita, indicando que, na maioria dos casos, o modelo apresenta erros pequenos ou moderados, concentrados em torno do zero, o que significa que as previsões foram, em grande parte, próximas dos valores reais.

A distribuição tem um pico próximo de 0%, evidenciando que o modelo tem uma boa acurácia na maioria das previsões.

Figura 3.17 - Gráfico de Distribuição de Erros



Fonte: elaborado pelo autor

No entanto, observa-se uma cauda longa à direita, sugerindo que há uma quantidade menor, mas significativa, de casos com erros elevados. Isso indica que, para alguns processos, o modelo superestima a duração da tramitação, possivelmente devido a variáveis não capturadas adequadamente ou eventos atípicos no fluxo processual.

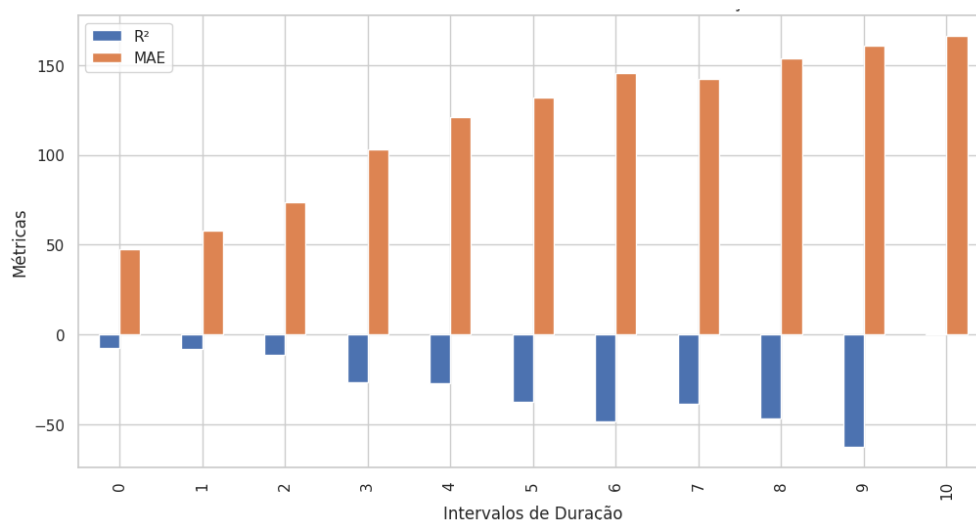
Por outro lado, a assimetria menor à esquerda indica que o modelo também subestima a duração de alguns processos, mas em menor grau. Isso pode estar relacionado a características específicas de determinados processos ou à variabilidade inerente ao tempo de tramitação.

Isso também pode ser constatado na Figura 3.18, na qual exemplifica-se a análise do desempenho do modelo em diferentes intervalos de duração dos processos judiciais, indicando que a eficácia da previsão varia conforme a extensão temporal dos casos analisados.

O coeficiente de determinação (R^2), representado pelas barras azuis, apresenta valores negativos em praticamente todos os intervalos, o que sugere que o modelo não consegue explicar bem a variabilidade dos dados nesses grupos. Esse comportamento indica que, em determinados casos, as previsões podem ser menos precisas do que o uso

da média histórica dos dados.

Figura 3.18 - Desempenho do Modelo por Intervalo de Duração dos Processos Judiciais



Fonte: elaborado pelo autor

O erro absoluto médio (MAE), representado pelas barras laranjas, cresce conforme aumenta a duração do processo, sugerindo que o modelo apresenta um erro médio mais baixo para os processos de curta duração e maior dificuldade para prever corretamente os casos que se estendem por períodos mais longos. Nos primeiros intervalos de duração, o MAE se mantém em níveis reduzidos, indicando que as previsões são mais confiáveis nesses cenários. No entanto, à medida que a duração dos processos aumenta, o MAE cresce significativamente, refletindo um desempenho menos eficiente do modelo preditivo para processos com durações mais longas.

De forma complementar, foram extraídas duas tabelas do conjunto de dados de teste: uma contendo as 50 melhores previsões - Tabela 3.7, caracterizadas pelos menores erros absolutos, e outra com as 50 piores previsões, onde os erros absolutos são mais elevados - Tabela 3.8. O erro absoluto, que mede a diferença entre a duração real dos processos judiciais e a previsão gerada pelo modelo, é uma métrica importante, pois fornece uma visão clara sobre a média das discrepâncias entre as previsões e os valores reais.

Tabela 3.7 – 50 Melhores Resultados

Nº	Data de Início	Data da Baixa	Duração Real (dias)	Previsão da Duração (dias)	Erro Absoluto
1	10/04/2023	30/06/2023	81	81	0
2	23/04/2024	14/05/2024	21	21	0

3	26/05/2023	23/08/2023	89	89	0
4	06/12/2023	04/03/2024	89	89	0
5	27/02/2024	28/05/2024	91	91	0
6	05/04/2022	20/06/2023	441	441	0
7	30/05/2023	30/08/2023	92	92	0
8	03/08/2023	17/11/2023	106	106	0
9	27/02/2024	10/05/2024	73	73	0
10	27/02/2024	10/05/2024	73	73	0
11	14/07/2023	26/02/2024	227	227	0
12	07/11/2022	28/02/2023	113	113	0
13	15/09/2022	10/02/2023	148	148	0
14	17/11/2023	29/02/2024	104	104	0
15	26/12/2023	19/02/2024	55	55	0
16	15/12/2022	14/09/2023	273	273	0
17	16/08/2023	15/12/2023	121	121	0
18	02/08/2023	20/11/2023	110	110	0
19	11/12/2023	19/02/2024	70	70	0
20	30/04/2024	21/05/2024	21	21	0
21	26/07/2023	27/09/2023	63	63	0
22	07/12/2023	25/06/2024	201	201	0
23	01/08/2023	10/10/2023	70	70	0
24	01/12/2023	21/03/2024	111	111	0
25	19/01/2022	16/05/2023	482	482	0
26	18/12/2023	09/04/2024	113	113	0
27	02/02/2024	18/04/2024	76	76	0
28	02/06/2023	09/08/2023	68	68	0
29	20/04/2023	17/08/2023	119	119	0
30	16/03/2023	11/04/2023	26	26	0
31	11/07/2022	26/09/2023	442	442	0
32	09/02/2024	01/07/2024	143	143	0
33	13/05/2023	15/08/2023	94	94	0
34	25/07/2023	29/09/2023	66	66	0
35	27/03/2023	31/05/2023	65	65	0
36	19/03/2024	05/07/2024	108	108	0
37	29/08/2023	08/02/2024	163	163	0
38	14/04/2023	21/07/2023	98	98	0
39	14/04/2023	21/07/2023	98	98	0
40	24/04/2024	26/06/2024	63	63	0
41	08/02/2024	05/07/2024	148	148	0
42	29/11/2022	09/08/2023	253	253	0
43	04/03/2024	18/04/2024	45	45	0
44	15/09/2023	18/03/2024	185	185	0
45	26/10/2023	19/02/2024	116	116	0
46	11/12/2023	19/02/2024	70	70	0
47	21/06/2023	24/08/2023	64	64	0
48	03/03/2023	22/05/2023	80	80	0
49	24/03/2023	06/10/2023	196	196	0
50	25/11/2022	27/02/2023	94	94	0

Fonte: elaborado pelo autor

Tabela 3.8 – 50 Piores Resultados

Nº	Data de Início	Data da Baixa	Duração Real (dias)	Previsão da Duração (dias)	Erro Absoluto
1	11/04/2005	24/03/2023	6556	7	6549
2	25/04/2006	03/03/2023	6156	7	6149
3	18/10/2012	27/06/2024	4270	4	4266
4	27/02/2015	24/08/2023	3100	4	3096
5	20/02/2015	05/05/2023	2996	4	2992
6	27/09/2018	07/02/2024	1959	1	1958
7	05/09/2018	28/11/2023	1910	3	1907
8	07/10/2019	12/04/2024	1649	2	1647
9	07/11/2019	15/04/2024	1621	1	1620
10	04/07/2019	21/09/2023	1540	2	1538
11	07/10/2019	09/03/2023	1249	2	1247
12	07/07/2020	27/06/2023	1085	3	1082
13	04/02/2021	30/08/2023	937	1	936
14	22/11/2022	07/03/2023	105	980	875
15	25/01/2024	27/02/2024	33	825	792
16	13/07/2023	22/03/2024	253	956	703
17	26/05/2022	09/01/2023	228	898	670
18	15/07/2021	26/02/2024	956	313	643
19	09/09/2022	01/03/2023	173	813	640
20	16/09/2022	17/06/2024	640	1	639
21	25/09/2023	15/12/2023	81	707	626
22	10/03/2020	05/09/2023	1274	663	611
23	22/02/2023	30/05/2023	97	696	599
24	19/07/2023	13/11/2023	117	704	587
25	28/04/2022	21/11/2023	572	2	570
26	30/09/2022	12/04/2023	194	754	560
27	29/11/2023	08/03/2024	100	660	560
28	27/06/2023	27/05/2024	335	890	555
29	04/08/2023	07/11/2023	95	632	537
30	07/12/2022	16/06/2023	191	727	536
31	29/11/2022	27/02/2023	90	626	536
32	16/12/2021	18/03/2024	823	292	531
33	02/03/2023	27/03/2023	25	554	529
34	17/08/2022	23/03/2023	218	743	525
35	21/06/2023	13/12/2023	175	694	519
36	11/10/2023	06/03/2024	147	663	516
37	23/05/2023	23/08/2023	92	608	516
38	16/02/2023	18/04/2023	61	570	509
39	29/05/2023	01/09/2023	95	600	505
40	06/07/2023	17/10/2023	103	596	493
41	09/11/2022	20/03/2023	131	618	487
42	23/08/2022	23/05/2023	273	750	477
43	18/04/2023	18/05/2023	30	499	469
44	22/02/2022	03/05/2023	435	1	434
45	18/01/2022	15/03/2023	421	1	420
46	08/02/2023	20/02/2024	377	1	376
47	20/04/2023	11/04/2024	357	2	355
48	20/10/2023	14/06/2024	238	1	237
49	23/06/2023	28/09/2023	97	1	96
50	19/07/2023	11/09/2023	54	2	52

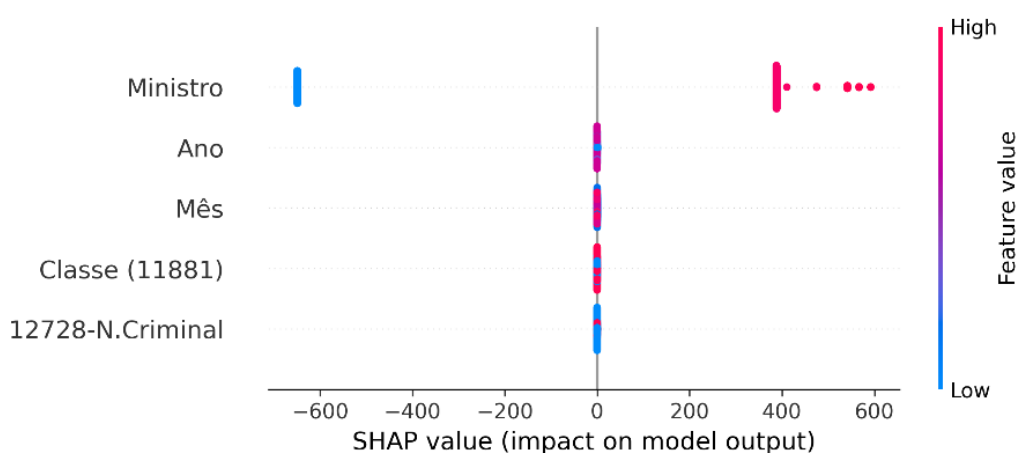
Fonte: elaborado pelo autor

As melhores previsões, que apresentam erros absolutos reduzidos, refletem a

eficácia do modelo em capturar a dinâmica dos processos judiciais em situações normais. Por outro lado, as piores previsões revelam casos em que o modelo falhou em oferecer estimativas precisas, destacando a importância de investigar esses *outliers* para aprimorar a aprendizagem do algoritmo, considerando fatores adicionais que possam impactar a duração dos processos.

Da mesma forma, a análise dos resultados utilizando *SHAP* (*SHapley Additive exPlanations*) permitiu visualizar as 5 variáveis utilizadas que mais impactaram as previsões do modelo, conforme Figura 3.19.

Figura 3.19 - Gráfico SHapley Additive exPlanations (SHAP)

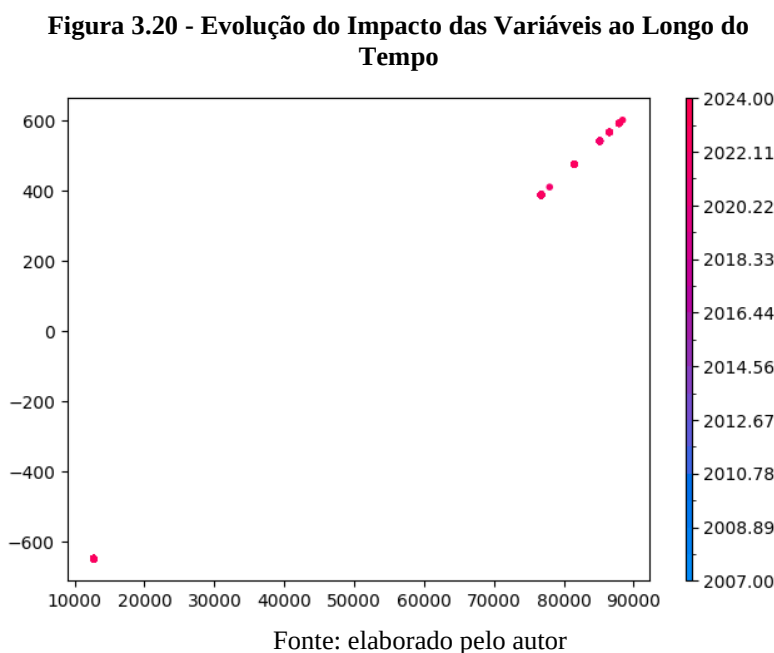


Fonte: elaborado pelo autor

A avaliação da importância das variáveis indicou que o fator "Ministro" exerce uma influência significativa na previsão da duração dos processos, com valores SHAP elevados e bem distribuídos, sugerindo que o magistrado responsável pelo julgamento é um elemento determinante para a celeridade processual. Em contrapartida, variáveis temporais, como "Ano" e "Mês", demonstraram menor impacto direto nas previsões do modelo, apresentando valores SHAP próximos de zero. Isso sugere que esses fatores, isoladamente, não são preditores primários da duração dos processos, podendo estar mais associados a outras variáveis interligadas no sistema judiciário.

A partir desse resultado, considerando as 5 variáveis mais relevantes apontadas na Figura 3.19, o gráfico da Figura 3.20 mostra a relação entre o impacto das variáveis no modelo preditivo e a evolução temporal dos processos judiciais. O eixo horizontal representa uma métrica associada ao tempo, enquanto o eixo vertical indica a magnitude do impacto das variáveis no resultado do modelo. A escala de cores, variando do azul ao

rosa, representa os anos dos processos, evidenciando padrões temporais.



Isso sugere que, ao longo do tempo, os fatores que influenciam a tramitação processual podem ter sofrido mudanças, com um aumento do impacto das variáveis mais recentes. Isso pode indicar a modernização do Judiciário, mudanças na composição dos tribunais ou alterações na legislação e nos procedimentos internos do STJ. Além disso, a presença de um ponto discrepante na parte inferior esquerda sugere um caso atípico, que pode ter sido influenciado por uma variável externa ou um erro na base de dados.

3.3. Limitações da Pesquisa

Uma das principais limitações refere-se à capacidade de processamento do ambiente de produção, em que pese a utilização de 50.515 registros, representando aproximadamente 10% do conjunto de dados original. Embora essa amostra tenha sido selecionada para garantir a eficiência do processamento, sua restrição pode impactar a capacidade de generalização do modelo.

Por outro lado, há variáveis que não foram consideradas nesta pesquisa, por não estarem disponíveis na base de dados utilizadas, e que poderiam aprimorar o modelo. Por exemplo, características específicas dos processos, como a natureza do delito, a presença de múltiplas partes envolvidas ou a complexidade jurídica do caso, podem influenciar significativamente a duração do processo.

Outras variáveis, como o histórico de decisões do Ministro responsável pelo caso e a carga de trabalho atual desses Magistrados, bem como de sua equipe de apoio, também são fatores importantes que podem afetar a tramitação processual e, consequentemente, a previsão de duração. Além disso, a introdução de variáveis temporais, como mudanças na legislação, feriados judiciais ou recessos, poderia contribuir para uma modelagem mais precisa, considerando o impacto que essas questões têm sobre os prazos de tramitação.

Outro ponto a ser considerado é a influência de fatores externos, como crises de saúde pública (por exemplo, a pandemia de COVID-19), que podem alterar significativamente a dinâmica do sistema judicial e a movimentação processual. A inclusão de dados sobre a interação entre diferentes variáveis, como o tipo de ação, o tempo médio de resolução de casos semelhantes e as condições socioeconômicas das partes envolvidas, talvez possa fornecer um entendimento mais profundo sobre o comportamento dos processos e permitir que o modelo capture nuances que não foram consideradas inicialmente.

3.4. Produto Técnico Tecnológico

O Produto Técnico Tecnológico (PTT) desenvolvido neste trabalho consiste em um modelo preditivo treinado para estimar a duração de processos judiciais no Superior Tribunal de Justiça (STJ), utilizando técnicas avançadas de aprendizado de máquina, especificamente redes neurais artificiais. A entrega deste PTT ao STJ inclui:

- algoritmo preditivo treinado (modelo_otimizado.pkl); e
- código-fonte completo: scripts em python que abrangem todas as etapas do processo, desde a preparação dos dados até o treinamento e validação do modelo.

Para que o STJ possa fazer uso eficaz deste PTT, o modelo pode ser incorporado aos seus sistemas de gestão processual, permitindo que as previsões sejam acessadas em tempo real. A integração é facilitada pelo uso de tecnologias amplamente adotadas, como APIs RESTful desenvolvidas em Flask ou Django.

Para previsão da duração, os dados de novos processos (como datas de início, códigos de assuntos, classes processuais, entre outros) devem ser fornecidos ao modelo. O pipeline de pré-processamento assegura que os dados estejam no formato adequado para a previsão. Dessa forma, o modelo retorna a estimativa de duração em dias para cada

processo.

Assim, com previsões mais precisas, o STJ pode planejar melhor a distribuição de processos entre ministros, juízes auxiliares e servidores; identificar gargalos e necessidade de reforço em determinadas áreas ou classes processuais; realizar o monitoramento proativo de processos com risco de exceder prazos legais, permitindo intervenções tempestivas; gerenciar o cumprimento das metas estabelecidas pelo Conselho Nacional de Justiça (CNJ), dentre outros usos.

Quanto ao impacto social, há a possibilidade de informar às partes envolvidas estimativas mais realistas sobre o tempo de tramitação de seus processos, o que pode resultar numa melhor percepção de eficiência e comprometimento do STJ com a celeridade processual.

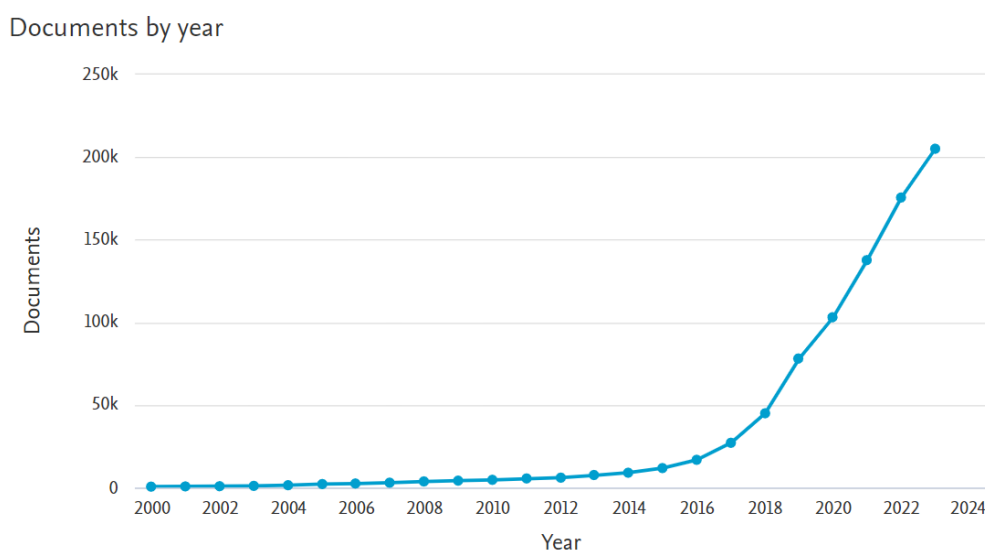
O STJ pode utilizar os *scripts* fornecidos para retreinar o modelo periodicamente, incorporando novos dados e adaptando-se a mudanças nos procedimentos judiciais ou legislações. O modelo pode ser ajustado para incluir novas variáveis ou para focar em áreas específicas de interesse do tribunal, bem como ser estendido para outros tribunais ou instâncias, adaptando o modelo às particularidades de cada contexto, dado o seu alto potencial de replicabilidade.

3.5. Base teórica utilizada

Inicialmente, com o intuito de demonstrar a relevância acadêmica deste estudo, foi realizado um levantamento bibliométrico da literatura na base de dados *SCOPUS* em fevereiro de 2024. Seguindo o método bibliométrico, o foco foi quantificar, analisar e avaliar estudos acadêmicos relativos à aplicação de redes neurais a partir do ano 2000 até 2023.

Foram utilizadas as palavras-chaves “*Machine Learning*” e “*Deep Learning*” para obter o quantitativo de publicações que possuísse uma das duas palavras-chaves e, como resultado, foram encontrados 852.887 mil (oitocentos e cinquenta e dois mil oitocentos e oitenta e sete) documentos, distribuídos ao longo dos anos conforme consta na Figura 3.21.

Figura 3.21 - Levantamento Bibliométrico



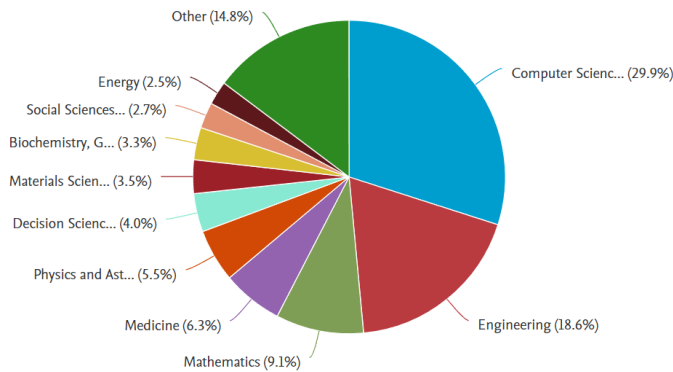
Fonte: Scopus Análise (2024)

Após refinar a busca concatenando as palavras-chaves: “*Machine Learning*” OR “*Deep Learning*” AND “*Justice*” OR “*Judiciary*” OR “*Judge*”, obteve-se como resultado 2.951 (duas mil novecentas e cinquenta e uma) publicações.

Assim, depreende-se, a partir do levantamento realizado, que as redes neurais, no contexto de *Deep Learning*, vem sendo aplicadas para atender às diversas necessidades humanas; tais como: segurança da informação, ciência médica, negócios, finanças, bancária, seguros, mercado de ações, geração de eletricidade, gestão, indústria nuclear, exploração mineral, mineração, previsão de qualidade, previsão de rendimento de colheitas, tratamento de água e política, dentre outros, sendo possível visualizar essa distribuição entre diferentes áreas do conhecimento na Figura 3.22.

Figura 3.22 - Publicações por Área do Conhecimento

Documents by subject area



Fonte: Scopus Análize (2024)

Portanto, essa análise inicial demonstra a importância crescente e a aplicabilidade diversificada das redes neurais em múltiplas disciplinas. A abordagem interdisciplinar e o volume substancial de literatura encontrada ressaltam a contribuição significativa que este estudo pode proporcionar para o campo acadêmico, expandindo o entendimento das redes neurais além de suas aplicações tradicionais, especialmente no contexto jurídico, que é um território ainda pouco explorado no Brasil.

Dessa forma, para uma compreensão abrangente e organizada da literatura pertinente ao estudo das redes neurais aplicadas ao contexto jurídico, especialmente na previsão da duração de processos no Superior Tribunal de Justiça (STJ), a revisão de literatura a seguir foi desenvolvida a partir de uma estruturação em eixos temáticos. Esta abordagem permite uma análise detalhada e sistematizada, cobrindo desde os fundamentos teóricos até as aplicações práticas.

A seguir, é apresentada a Tabela 3.9 que resume os eixos temáticos, seus objetos de estudo específicos, a relevância de cada um para a pesquisa e os tópicos abordados:

Tabela 3.9 - Eixos Temáticos

Eixo Temático	Objeto de Estudo	Relevância para o Projeto	Tópicos Abordados
Eixo 1: Inteligência Artificial	Fundamentos e aplicações de soluções de Inteligência Artificial	Compreender as bases teóricas e práticas para desenvolvimento de soluções de <i>Machine Learning</i>	- Inteligência Artificial; - Aprendizado de Máquina e Aprendizado Profundo; - Modelo Multi-layer Perceptron;

			• Bibliotecas Python: Numpy, Pandas e Scikit-learn; • Google Colaboratory
Eixo 2: Análise e Gestão de Dados	Como os dados judiciais são coletados, processados e utilizados	Entender a gestão de dados no contexto jurídico brasileiro para fundamentar o uso de redes neurais	• Gestão de Dados pelo Poder Judiciário; • Despesas com Informática e Novos Processos; • Casos Pendentes e Tempo de Giro do Acervo; • Processos de Conhecimento e de Execução
Eixo 3: Inovação	Inovação como motor de transformação econômica, social e tecnológica	Explorar a evolução do papel da inovação na modernização do serviço público	• Conceitos e dinâmicas relacionadas à inovação no Setor Público;
Eixo 4: Utilização de IA pelo Sistema Jurídico	Soluções de IA em uso pelo Sistema Judiciário	Investigar na literatura os resultados obtidos com a utilização de IAs pelo Sistema Judiciário	• Automação no Poder Judiciário; • Previsão de prazos processuais; • Modelos baseados em redes neurais

Fonte: elaborado pelo autor

3.5.1 Eixo 1: Inteligência Artificial

Inteligência Artificial (IA) é o ramo da ciência da computação que se dedica a criar sistemas capazes de realizar tarefas que, até então, exigiriam inteligência humana para serem executadas. Isso inclui atividades como aprender, raciocinar, reconhecer padrões, compreender a linguagem e tomar decisões. Conforme destacado no Teste de *Turing*, a IA busca alcançar uma funcionalidade que possa ser considerada inteligente dentro de um contexto específico, sendo avaliada por sua capacidade de produzir respostas ou ações que se alinhassem com aquelas esperadas de um ser humano sob condições controladas (TURING, 1950).

Esse conceito abarca a ideia de que a IA não precisa emular a forma humana ou expressar emoções para ser considerada "inteligente". Em vez disso, deve ser capaz de processar informações, aprender com experiências, e responder a perguntas ou realizar tarefas de maneira que sejam indistinguíveis das de um ser humano, quando observadas através de uma interface que não revela sua natureza não humana, refletindo uma abordagem funcional e pragmática para o desenvolvimento e avaliação da IA.

Ampliando essa visão, *John McCarthy*, em 2007, reforçou que a IA deve ser entendida como "*a ciência e engenharia de fazer máquinas inteligentes, especialmente*

programas de computador inteligentes", destacando que ela não se limita a emular a inteligência humana, mas também explora capacidades que podem não ser biologicamente observáveis (MCCARTHY, 2007).

Além disso, a IA abrange diversos sub-ramos, incluindo aprendizado de máquina (*machine learning*) e aprendizado profundo (*deep learning*), que segundo Kaplan e Haenlein (2019, pág. 3), revelam a “*capacidade do sistema para interpretar corretamente dados externos, aprender a partir desses dados e utilizar essas aprendizagens para atingir objetivos e tarefas específicos através de adaptação flexível*”; processamento de linguagem natural (PLN), que permite às máquinas entender e responder em linguagem humana; e visão computacional, que dota as máquinas da capacidade de interpretar e entender o mundo visual.

3.5.1.1. Aprendizado de Máquina e Aprendizado Profundo

A imensa produção de dados impulsiona o avanço de métodos de aprendizado de máquina, que se baseiam na alimentação de volumosas quantidades de informações para que um sistema possa adquirir conhecimento sobre uma variável específica ou realizar uma tarefa. Segundo Olowonomi *et al.* (2021), uma interpretação frequente do aprendizado de máquina sugere que se refere à habilidade de sistemas computacionais em efetuar escolhas inteligentes sem necessidade de programação direta, sendo tais sistemas fortemente influenciados pela análise de dados.

Quanto a essa massiva utilização de dados, os autores Viktor Mayer-Schönberger e Kenneth Cukier (2013, p. 66) ressaltam a relevância atual do *Big Data*:

[...] a dataficação representa um essencial enriquecimento da compreensão humana. Com a ajuda do *Big Data*, não mais veremos o mundo como uma sequência de acontecimentos explicados como fenômenos naturais ou sociais, e sim como um Universo composto essencialmente por informações. [...]

Os métodos de *machine learning* podem ser agrupados em duas vertentes: com supervisão e sem supervisão. Na abordagem supervisionada, os modelos computacionais são desenvolvidos com base em conjuntos de dados que já contêm as respostas desejadas. Para os modelos sob supervisão, os problemas são categorizados em dois tipos: classificação e regressão. Já em métodos de aprendizado não supervisionado, os algoritmos operam sem guia ou conhecimento prévio das respostas exatas, aplicando-se

primordialmente em atividades de classificação e identificação de padrões recorrentes.

Nesse contexto, *Deep Learning* (aprendizado profundo), um subconjunto do *Machine Learning*, faz uso de Redes Neurais Artificiais (**RNAs**), as quais são definidas por Haykin (2017, p. 28) como sendo um “*processador maciçamente e paralelamente distribuído constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso*”.

Sua utilização oferece as seguintes propriedades úteis e capacidades (HAYKIN, 2017, p. 29 a 31):

- Não linearidade;
- Mapeamento de entrada e saída (aprendizagem supervisionada): envolve a modificação dos pesos sinápticos de uma rede neural pela aplicação de um conjunto de amostras de treinamento rotuladas ou exemplos de tarefas;
- Adaptabilidade;
- Resposta à evidência;
- Informação contextual;
- Tolerância a falhas;
- Implementação em VLSI (*very-large-scale-integration*);
- Uniformidade de análise e projeto; e
- Analogia neurobiológica.

Por isso, as técnicas de RNAs têm encontrado amplo uso ao demonstrar eficiência na abordagem de desafios práticos através da fase inicial de aprendizado com dados etiquetados, seguido por uma operação independente (KARAR *et al.* 2020).

Apesar de se basearem nas redes neurais do cérebro humano, que são formadas por unidades de processamento trabalhando em conjunto e cujas operações são definidas pelas ligações entre estas unidades (WANG *et al.* 2021; HAYKIN, 2017; BRAGA *et al.* 2011), os neurônios artificiais ainda são primitivos em comparação com aqueles encontrados no cérebro humano (HAYKIN, 2017, p. 36).

3.5.1.2. Modelo Multi-layer Perceptron (MLP)

As RNAs são baseadas em redes neurais do cérebro humano, sendo compostas por elementos denominados neurônios. McCulloch e Pitts (1943) propuseram um modelo de neurônio, conhecido como Neurônio de McCulloch-Pitts, uma simplificação altamente

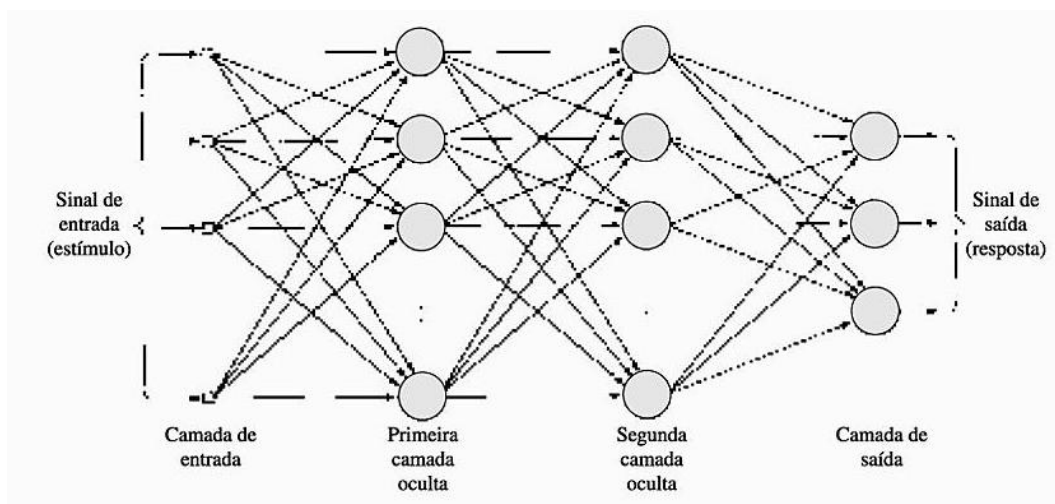
abstrata do neurônio biológico, que o representa como um dispositivo lógico binário capaz de realizar uma soma ponderada das entradas e emitir uma saída com base em um limiar. Segundo Haykin (2017, p. 39), eles foram os pioneiros na modelagem de redes neurais.

Posteriormente, como um desdobramento do modelo de McCulloch-Pitts, Frank Rosenblatt propôs, em 1957, o conceito de *Perceptron* — um modelo computacional inspirado no sistema nervoso, voltado para a aprendizagem com reforço, similar ao aprendizado supervisionado, e capaz de realizar classificações binárias com base em estímulos sensoriais.

O Perceptron é uma unidade de processamento que realiza a soma ponderada de suas entradas sinápticas e, em seguida, aplica uma função de ativação binária baseada em limiar — equivalente a uma função degrau — produzindo uma saída do tipo tudo-ou-nada. Essa simplicidade o torna adequado para tarefas de classificação supervisionada com categorias bem separáveis. Como descreve o próprio Rosenblatt (1958, p. 389, tradução nossa), “se a soma algébrica das intensidades dos impulsos excitatórios e inibitórios for igual ou superior ao limiar (θ) da unidade A, então a unidade A dispara, novamente de forma tudo-ou-nada”.

Na sequência, dada a limitação do modelo anterior e constituindo uma evolução desse conceito, foi proposto o modelo *Mult Layer Perceptron (MLP)* ou *Perceptron Multicamadas*, o qual pode ser definido como uma rede neural, supervisionada, composta por múltiplos neurônios entre a camada de entrada e a de saída, Figura 3.23.

Figura 3.23 - Diagrama Funcionamento do Modelo MLP



Fonte: Haykin, 2017, p. 186.

As camadas e a quantidade de neurônios podem ser ajustadas de acordo com a complexidade esperada da saída (RAMCHOUN *et al* 2016; SILVA *et al* 2010; BRAGA *et al* 2011).

Assim, a inclusão de uma ou mais camadas intermediárias, também conhecidas como camadas ocultas, possibilita que a rede MLP implemente qualquer função contínua, superando assim a restrição de linearidade do modelo anterior. Posto isso, um desafio fundamental na implementação de redes MLP diz respeito ao treinamento das camadas intermediárias.

Diferentemente de redes de uma única camada, onde a resposta desejada e o erro de cada nodo são claramente definidos, em uma rede MLP a determinação do erro para os nodos das camadas intermediárias não é direta. Isso se deve à ausência de uma resposta desejada específica para esses nodos, o que levanta a questão de como seu erro deve ser calculado (BRAGA *et al* 2011; HAYKIN, 2017).

Braga *et al* (2011) exploram essa questão, discutindo a complexidade de avaliar o erro em redes que utilizam funções de ativação do tipo limiar (geralmente retornam 0 ou 1, ou -1 e 1). Para superar esses desafios, ele elenca o método baseado em **gradiente descendente** como uma estratégia viável para o treinamento de redes MLP.

Esse método requer que as funções de ativação sejam contínuas e diferenciáveis, preferencialmente não-decrescentes, para que os erros cometidos pela rede possam ser comunicados às camadas anteriores com precisão.

Por isso, a relevância das redes MLP reside na sua capacidade de tratar dados não linearmente separáveis, expandindo significativamente o universo de problemas que podem ser abordados por redes neurais artificiais. A introdução do algoritmo **back-propagation** e suas variações têm como objetivo aprimorar o desempenho do treinamento dessas redes, o que evidencia a relevância do desenvolvimento no campo das RNAs.

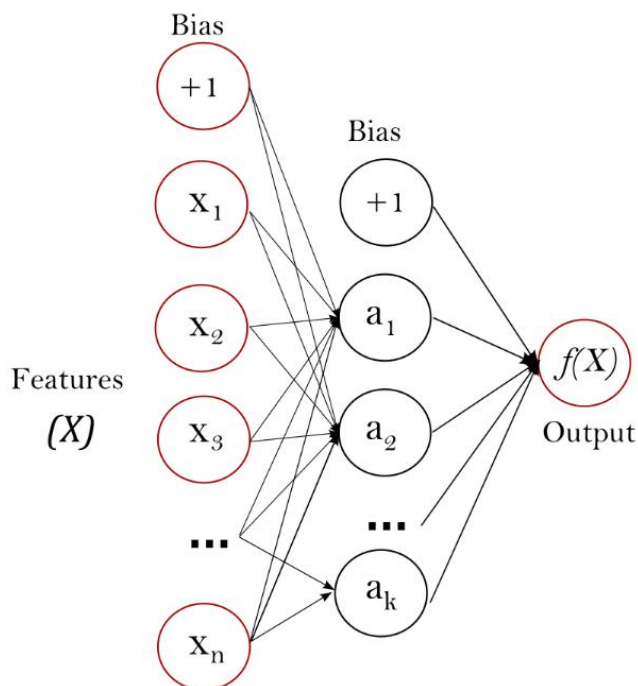
3.5.1.3. Scikit-learn: Machine Learning com Python

Uma abordagem para aplicar o MLP (*Perceptron Multicamada*) envolve o uso do *Python*, especificamente através da biblioteca *Scikit-learn*. Essa ferramenta de aprendizado de máquina facilita a criação de modelos de redes neurais artificiais para

regressão, classificação e outras funções.

A Figura 3.24 é um exemplo de MLP com uma camada oculta, representando graficamente um algoritmo de aprendizado supervisionado que aprende uma função $f(\cdot): \mathbf{R}^m \rightarrow \mathbf{R}^0$ treinando em um conjunto de dados, onde “ m ” é o número de dimensões para entrada e “ 0 ” é o número de dimensões para saída (SCIKIT-LEARN, 2024).

Figura 3.24 - MLP de uma camada oculta



Fonte: Scikit-Learn, 2024.

Dessa forma, a Figura 3.23 pode ser interpretada da seguinte forma:

1. **Features (X):** são os dados de entrada do neurônio, representa um recurso ou característica que se deseja que o modelo aprenda;
2. **Bias (+1):** é um termo adicional na função de ativação do neurônio que ajusta a saída ao longo do eixo y, pode ser pensado como uma forma de iniciar o modelo. Mesmo se todos os recursos de entrada X fossem zero, ainda haveria uma ativação (ou saída) do neurônio devido ao **Bias**;
3. **Pesos e Somas Ponderadas:** cada ligação entre uma entrada X_i e o neurônio a_i na camada oculta tem um peso associado. O neurônio calcula a soma ponderada das entradas multiplicando cada recurso X_i pelo seu peso correspondente e somando-os todos juntos, incluindo o **Bias**;

4. **Neurônios (a_i):** os neurônios na camada oculta $a_1, a_2, \dots, a_k$ processam as somas ponderadas das entradas. Cada neurônio na camada oculta tem sua própria soma ponderada e aplica sua própria função de ativação, que pode ser ReLU, tanh, sigmóide etc;
5. **Função de Ativação ($f(X)$):** após calcular a soma ponderada, o neurônio aplica uma função de ativação não linear, como tanh ou ReLU. A função de ativação determina se e como o sinal processado pelo neurônio será passado adiante; e
6. **Output ($f(X)$):** a saída final $f(X)$ do neurônio é o resultado da função de ativação. Em uma MLP, essa saída seria usada como entrada para os neurônios na próxima camada.

Segundo Pedregosa *et al* (2011), a biblioteca *Scikit-Learn* torna o *machine learning* acessível para não especialistas, enfatizando a facilidade de sua utilização, desempenho e consistência, com poucas dependências e distribuição sob a licença de código aberto, o que incentiva sua utilização tanto em ambientes acadêmicos quanto comerciais, bem como promove uma consistência e uma padronização nas práticas de modelagem, contribuindo significativamente para o avanço da pesquisa e desenvolvimento em inteligência artificial.

Possui, dentre outras, as seguintes ferramentas:

- ***StandardScaler*:** utilizado na etapa de pré-processamento de dados, normaliza as características (*features*) de um conjunto de dados antes de aplicá-las ao modelo;
- ***train_test_split*:** separa de forma randômica o conjunto de dados em dois grupos distintos: um para teste e outro para treinamento; e
- ***RandomSearchCV*:** ajuda a percorrer hiperparâmetros predefinidos e ajustar o estimador (modelo) em um conjunto de treinamento específico. Ao final, são apresentados os melhores parâmetros dentre os hiperparâmetros listados.

Portanto, conforme indicado pela documentação do *Scikit-learn*, a configuração de uma rede MLP é viabilizada por meio do uso de um construtor, que permite a personalização de uma série de hiperparâmetros. Esses hiperparâmetros vêm com valores padrão, possibilitando a criação de uma rede MLP mesmo sem a especificação explícita

de qualquer um deles, divide-se em dois tipos: *MLPRegressor* e *MLPClassifier*.

Por isso, é difícil estabelecer qual o melhor parâmetro para treinar o modelo. Contudo, para ajustar os hiperparâmetros pode-se utilizar o método do *RandomSearchCV* e validação cruzada. Esta função ajuda a percorrer hiperparâmetros predefinidos e ajustar o estimador (modelo) em um conjunto de treinamento específico. Ao final, são apresentados os melhores parâmetros dentre os hiperparâmetros listados.

Além disso, a integração com o *Numpy*, biblioteca do *Python* que suporta a criação e manipulação de *arrays* e matrizes multidimensionais, oferece uma ampla gama de operações matemáticas e numéricas. Conforme destacado por Pedregosa et al. (2011), a escolha estratégica do scikit-learn em utilizar *arrays* do *Numpy* para representação de dados de entrada, garante eficiência e integração com o ecossistema de ciência de dados em *Python*.

Esse tipo de abordagem não só facilita a análise e manipulação de grandes volumes de dados, mas também assegura a compatibilidade com outras bibliotecas científicas, otimizando assim o desenvolvimento e implementação de modelos complexos de aprendizado de máquina.

3.5.1.4. Numpy, Pandas e Google Colaboratory

Numpy e *Pandas* são bibliotecas de análise e manipulação de dados, ambas de código aberto, construídas sobre linguagem de programação *Python* (PANDAS, 2024; NUMPY, 2024). Com as ferramentas é possível ler arquivos em formato CSV, SQL, XLSX, entre outros. De forma complementar, o *Colab* é um produto da *Google* que possibilita a qualquer pessoa escrever e executar código *Python* dos servidores em nuvem da *Google* (GOOGLE, 2024).

Posto isso, a utilização conjunta dessas ferramentas com o *Scikit-learn* permite a aplicação de uma ampla gama de algoritmos de aprendizado supervisionado e não supervisionado. A integração do *Scikit-learn* com o *Pandas* e o *Numpy* facilita a implementação de *pipelines* de *machine learning*, desde a fase de pré-processamento de dados até a modelagem e avaliação de algoritmos, em um ambiente propício à experimentação e à inovação acadêmica (PEDREGOSA et al. 2011).

3.5.2 Eixo 2: Dados do Poder Judiciário Brasileiro

Nesta seção, serão abordados tópicos referentes à gestão de dados no Poder Judiciário brasileiro, com base no relatório Justiça em Números (SALOMÃO *et al.* 2019; 2021; 2023). A produção de dados é inerente à função jurisdicional e sua transformação em informação é tema central na gestão do Judiciário, tanto na circulação dessas informações quanto na gestão administrativa.

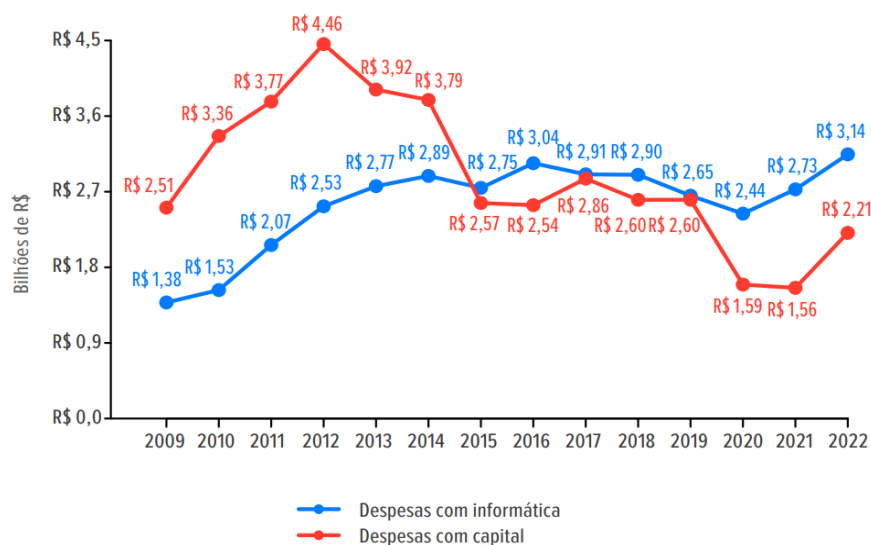
Por isso, a produção de dados estatísticos confiáveis é a base para o planejamento estratégico e o monitoramento de ações voltadas à melhoria da prestação jurisdicional, bem como reforça que a existência e a publicidade desses dados permitem ao público compreender as atividades do Judiciário, que os órgãos de cúpula monitorem seu desempenho e que o governo promova políticas públicas de reforma (OLIVEIRA; CUNHA, 2020).

Sobre o assunto, a Estratégia Brasileira de Inteligência Artificial (EBIA) ressalta a importância dos dados abertos, os quais fomentam o desenvolvimento de aplicações e soluções, gerando benefícios econômicos ao setor público e à economia como um todo, que possibilitam a tomada de decisão baseada em evidências, com benefícios significantes e transversais, na medida em que atingem áreas inicialmente não previstas (MCTI, 2021, p.44).

Assim, um dos marcos na transformação da gestão de dados no Poder Judiciário brasileiro, foi a implementação do sistema *DataJud*. Este sistema proporciona uma abordagem unificada e automatizada para a coleta e processamento de dados processuais, que, até então, dependiam de métodos manuais suscetíveis a erros, podendo ser acessados por meio do Painel de Estatísticas do *DataJud*, uma ferramenta pública que permite o acesso dinâmico às informações processuais.

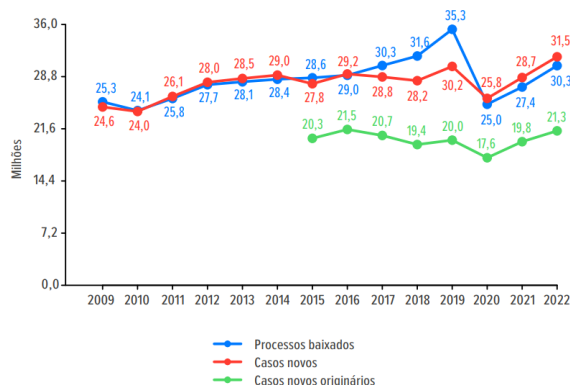
3.5.2.1. Despesas com Informática e Novos Processos

A evolução das despesas com informática no Poder Judiciário brasileiro, conforme ilustrado nos dados constantes do relatório Justiça em Números (SALOMÃO *et al.*, 2023), revela um aumento significativo de gastos até o ano de 2015, seguido de uma estabilização e posteriormente uma recuperação moderada (Figura 3.25).

Figura 3.25 - Série Histórica da Despesa com Informática e com Capital

Fonte: SALOMÃO et al. 2023.

Com relação ao volume de processos, houve um aumento dos casos novos originários em 2022, destacando-se também no contexto de aceleração digital, desencadeado pelas necessidades impostas pela pandemia de COVID-19. A digitalização forçada, resultante das medidas de isolamento, impulsionou a utilização de sistemas eletrônicos para a tramitação processual, o que pode ter contribuído para o crescimento na entrada de novos casos (Figura 3.26).

Figura 3.26 - Série Histórica dos Casos Novos e Processos Baixados

Fonte: SALOMÃO et al. 2023

A análise cruzada dos dados financeiros e processuais do Judiciário brasileiro sugere uma correlação entre o investimento em tecnologia da informação e o manejo do acervo processual. O período subsequente à pandemia evidenciou uma retomada nos

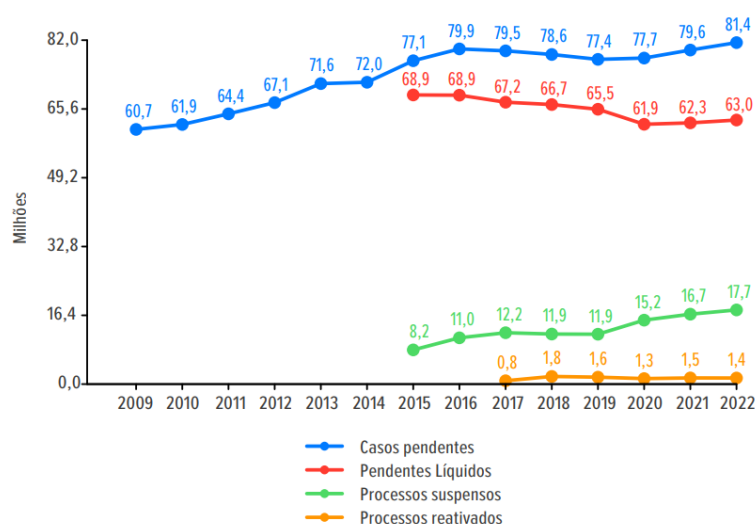
gastos com informática, alinhando-se com o incremento nos casos novos originários.

3.5.2.2. Casos Pendentes e Tempo de Giro do Acervo

Os casos pendentes no Judiciário brasileiro, definidos como processos que foram iniciados e nunca concluídos ou que foram reativados após uma baixa inicial, representam uma medida crítica do acúmulo processual enfrentado pelo sistema. A análise da Figura 3.27 indica um crescimento constante nos casos pendentes ao longo dos anos, culminando em um total expressivo de 81,4 milhões em 2022. Este volume reflete não apenas os desafios operacionais do judiciário, mas também a necessidade de estratégias efetivas de gestão de casos para lidar com o crescente *backlog* processual.

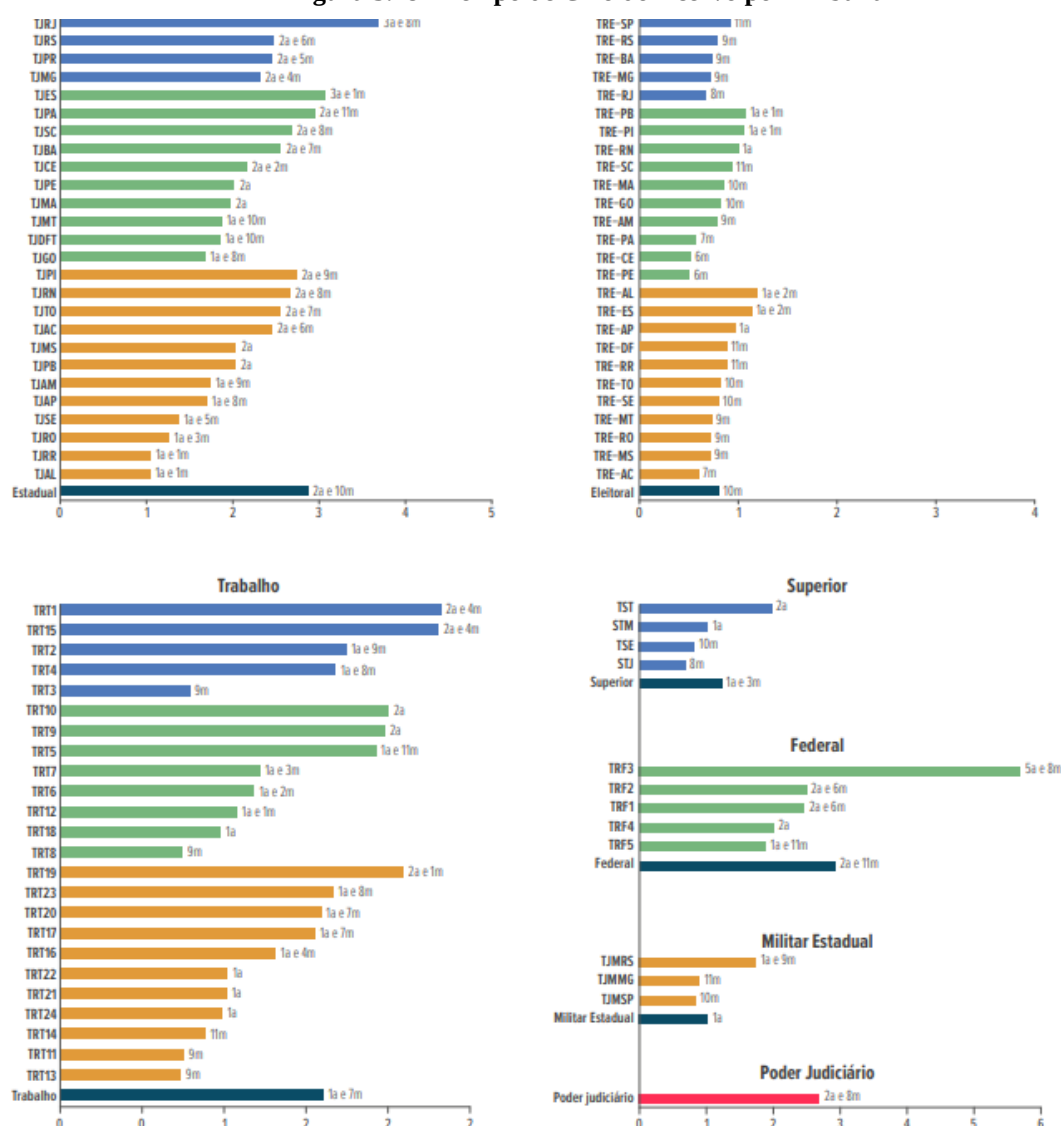
Figura 3.27 - Série histórica dos casos pendentes

Fonte: SALOMÃO et al. 2023



O indicador “*Tempo de Giro do Acervo*” emerge como um indicador chave para compreender a eficiência do Judiciário na resolução de casos pendentes. Este indicador, que reflete o tempo necessário para liquidar o acervo existente mantendo a produtividade atual, varia entre os ramos da justiça. Enquanto na Justiça Estadual e Federal são necessários quase três anos para zerar o acervo, na Justiça do Trabalho e nos Tribunais Superiores esse tempo é substancialmente menor (Figura 3.28).

Figura 3.28 - Tempo de Giro do Acervo por Tribunal

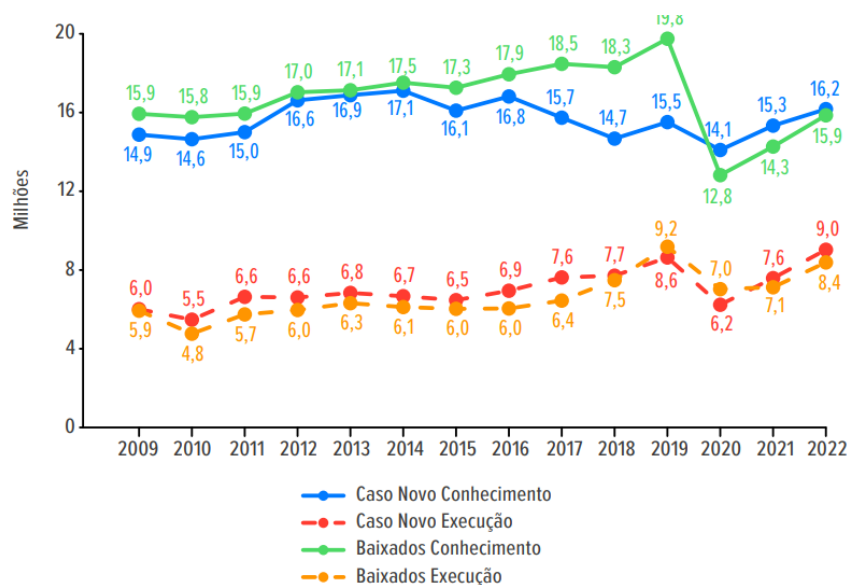


Fonte: SALOMÃO et al. 2023

3.5.2.3. Processos de Conhecimento e de Execução

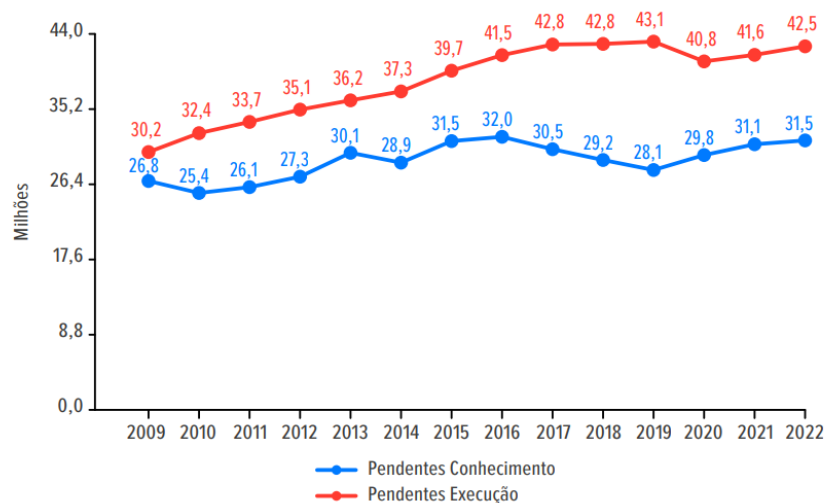
As Figuras 3.29 e 3.30 indicam um aumento consistente nos casos novos de conhecimento até 2019, seguido por uma diminuição em 2020 e um pico subsequente em 2021, sugerindo uma influência direta da pandemia na litigiosidade. A fase de execução, notoriamente mais morosa, também reflete essa tendência, embora de forma mais gradual.

Figura 3.30 - Série Histórica dos Casos Novos e Baixados nas Fases de Conhecimento e Execução



Fonte: SALOMÃO et al. 2023

Figura 3.29 - Série Histórica dos Casos Pendentes nas Fases de Conhecimento e Execução



Fonte: SALOMÃO et al. 2023

O número de casos baixados, tanto de conhecimento quanto de execução, mostrou flutuações ao longo dos anos, com um declínio notável nos casos de conhecimento após 2019, possivelmente devido ao aumento da complexidade e da quantidade de casos.

O gráfico da Figura 3.30 demonstra que os casos pendentes de execução se mantiveram consistentemente superiores aos de conhecimento durante todo o período observado, de 2009 a 2022. Observa-se uma tendência de crescimento contínuo e estável nos casos pendentes de execução, sem grandes oscilações, indicando um volume sempre

maior em comparação aos pendentes de conhecimento.

3.5.3 Eixo 3: Inovação

O conceito de inovação remete à antiguidade, conforme SCHWAB (2016), a inovação inicialmente estava ligada à introdução de novas ideias ou métodos, especialmente no contexto de mudanças sociais e culturais significativas. Nesse período, a inovação era vista como uma forma de questionar tradições e propor estratégias de reforma globais e disruptivas (CAVALCANTE; CAMÕES, 2017).

Com o advento da Primeira Revolução Industrial no século XVIII, a inovação passou a ser diretamente associada ao progresso tecnológico. Segundo SCHWAB (2016), invenções como a máquina a vapor não apenas transformaram processos produtivos, mas também reconfiguraram estruturas sociais e econômicas. Essa fase marcou o início de uma visão mais instrumental e econômica da inovação.

Durante o século XX, especialmente após a Segunda Revolução Industrial, a inovação tornou-se sinônimo de eficiência e competitividade. Já nas décadas de 1960 a 1990, a chamada Terceira Revolução Industrial introduziu a computação, a automação e as telecomunicações, expandindo o alcance e a velocidade da inovação. Schwab (2016) ressalta que, nesse período, a inovação deixou de ser exclusiva do setor produtivo e passou a transformar serviços, comunicação e relações sociais.

A Quarta Revolução Industrial marca uma convergência entre os mundos físico, digital e biológico. Tecnologias como inteligência artificial, blockchain e biotecnologia ampliaram a complexidade e os impactos da inovação, que agora molda desde padrões de consumo até sistemas de governança (SCHWAB, 2016).

Conforme Cavalcante e Camões (2017), o conceito de inovação evoluiu para incluir dimensões sociais e organizacionais, além da tecnologia. Isso reflete uma visão mais ampla, em que a inovação não se restringe a avanços técnicos, mas também envolve práticas que promovem inclusão, sustentabilidade e resolução de problemas coletivos.

Embora historicamente associada ao setor privado, a inovação ganhou relevância no setor público, onde busca-se aumentar a eficiência e melhorar serviços para a sociedade. Cavalcante e Camões (2017) destacam que a inovação pública é frequentemente impulsionada por crises e demandas sociais, como nas áreas de saúde e

educação.

Segundo BEKKERS et al (2011), a inovação é o principal motor da mudança econômica, constituindo-se numa condição necessária para a criação de uma economia competitiva, na qual pressupõe-se um ambiente institucional onde empresas, organizações não governamentais e governos possam aprender, desenvolver e compartilhar conhecimentos avançados.

No cenário atual, a inovação é cada vez mais vista como resultado de esforços colaborativos. Cavalcante e Camões (2017) discutem a tríplice hélice, modelo que integra governo, empresas e academia para fomentar ambientes de inovação. Essa abordagem reflete uma mudança cultural em que a inovação é cocriada, e não apenas desenvolvida de forma isolada ou como processo natural da evolução humana.

Por isso, o conceito de inovação está intrinsecamente ligado à transformação digital e à ética. Cavalcante e Camões (2017) apontam que, enquanto as tecnologias digitais continuam redefinindo o mundo, é fundamental garantir que os benefícios sejam amplamente distribuídos, evitando desigualdades e exclusões, consolidando-se como um processo contínuo e orientado para o bem-estar coletivo.

Para que as economias inovadoras prosperem, é essencial que o setor público também se transforme em um setor inovador, apto a enfrentar uma série de desafios sociais, como a melhoria do sistema educacional, o combate à criminalidade, a revitalização de cidades e regiões social e economicamente desfavorecidas, o avanço de uma economia sustentável ou a criação de infraestruturas de trânsito eficientes e sem congestionamentos.

Ao abordar esses desafios, os governos são levados a repensar sua organização interna e sua estrutura de gestão. Dessa forma, um setor econômico competitivo depende igualmente de um setor público inovador, uma vez que ambos são interligados e interdependentes. (BEKKERS; EDELENBOS; STEIJN, 2011).

A necessidade de inovação pública pode ser definida como a busca de novas ideias e conceitos, tecnologias, técnicas e métodos, formas, sistemas e procedimentos para criar interações significativas entre o governo e a sociedade, a fim de lidar com uma série de problemas sociais, constituindo-se como uma condição necessária para a modernização

do governo, a fim de enfrentar novos desafios sociais (BEKKERS; EDELENBOS; STEIJN, 2011, p. 4 e 10).

Observa-se que as estratégias de reforma globais parecem ter sido superadas, com foco voltado para pequenas alterações na administração pública, que proporcionam resultados em curto prazo e com custos transacionais reduzidos. Essa abordagem reflete a necessidade de adaptações ágeis e eficientes diante de demandas emergentes. Contudo, essas iniciativas enfrentam limitações impostas pelo aumento das expectativas dos cidadãos, pela complexidade dos problemas e por restrições orçamentárias, exigindo estratégias cada vez mais inovadoras e sustentáveis (CAVALCANTE; CAMÕES, 2017).

As agências do setor público frequentemente produzem ou adotam iniciativas em resposta às constantes mudanças econômicas, políticas, sociais e tecnológicas, características de um mundo mais globalizado e interconectado. Nesse contexto, apresenta-se como uma construção multidimensional, variando conforme tipos, objetivos, resultados e fases envolvidas. Ações que visam eficácia e eficiência não apenas contribuem para melhorar a qualidade e a eficiência dos serviços públicos, mas também podem resultar em maior satisfação dos cidadãos (CAVALCANTE; CAMÕES, 2017; MCTI, 2021, p. 52).

Em contraponto, segundo FREITAS e MEDEIROS (2015) a inovação que é constituída por artefatos não legitimados pela organização como usuais ou “normais” em práticas comuns, tende a ser recepcionada por atitudes, em relação à inovação, com uma tendência a ver práticas incomuns como quase inaceitáveis, resultando assim em pressões sociais contra seu uso.

O uso mais intenso e inovador das tecnologias em rede ajuda as administrações públicas a modernizarem suas estruturas e funções para melhorar seu desempenho global, como o fortalecimento dos processos de governança eletrônica para promover maior transparência, responsabilização e compromissos entre o governo e os seus cidadãos. Os governos devem também se adaptar ao fato de que o poder também está passando dos atores estatais para os não estatais e de instituições estabelecidas para redes mais abertas. As novas tecnologias e os agrupamentos sociais e interações que elas promovem permitem que praticamente qualquer pessoa exerça influência de maneira que teria sido inconcebível há apenas alguns anos (SCHWAB, 2016, p. 75).

A capacidade de adaptação dos governos irá determinar sua sobrevivência. Eles resistirão se abraçarem um mundo de mudanças exponencialmente disruptivas e se submeterem suas estruturas aos níveis de transparência e eficiência que podem ajudá-los a manter suas vantagens competitivas (SCHWAB, 2016, p. 76).

Por fim, conforme proposto por BEKKERS et al (2011, p. 10), as inovações públicas podem ser classificadas da seguinte forma:

- Inovações de produtos ou serviços: focadas na criação de novos serviços ou produtos públicos;
- Inovações tecnológicas: surgem através da criação e utilização de novas tecnologias, como a utilização de dispositivos de mensagens de texto e de radiodifusão celular para alertar os cidadãos em caso de emergência;
- Inovações de processos: focadas na melhoria da qualidade e eficiência dos processos de negócios internos e externos, como o redesenho de processos de prestação de serviços como a apuração digital de impostos;
- Inovações organizacionais e de gestão: focadas na criação de novas formas organizacionais, na introdução de novos métodos e técnicas de gestão e em novos métodos de trabalho;
- Inovações conceituais: ocorrem em relação à introdução de novos conceitos, quadros de referência ou mesmo novos paradigmas que ajudam a ressignificar a natureza de problemas específicos, bem como suas possíveis soluções;
- Inovações de governança: são direcionadas ao desenvolvimento de novas formas e processos de governança para resolver problemas sociais específicos, como as práticas de governança que tentam melhorar as capacidades de autorregulação e auto-organização das redes de políticas;
- Inovações institucionais: transformações fundamentais nas relações institucionais entre organizações, instituições e outros atores do setor público e, mais especificamente, da administração pública.

3.5.4 Eixo 4: Utilização de soluções de IA pelo Sistema Jurídico

A aplicação da Inteligência Artificial (IA) na gestão do Poder Judiciário tem demonstrado avanços significativos na eficiência e transparência processual. O uso de modelos de aprendizado de máquina e redes neurais possibilita uma abordagem preditiva, permitindo que tribunais otimizem a alocação de recursos e antecipem gargalos na tramitação dos processos. Estudos exploram a utilização da IA para prever prazos processuais, automatizar análises jurídicas e facilitar a tomada de decisão por magistrados e servidores.

O CNJ, por meio da Pesquisa sobre o Uso de Inteligência Artificial no Poder Judiciário (2023), revelou que 62 tribunais já possuem projetos de IA em desenvolvimento, sendo que 63 estão em produção e 46 ainda em fase de implementação. A Justiça Estadual lidera em número de projetos (48,6%), seguida pela Eleitoral (16,4%) e pela Justiça do Trabalho (14,3%). A pesquisa também destacou que as principais aplicações da IA no Judiciário incluem classificação de texto (49,3%), busca semântica (21,4%) e detecção de similaridade textual (35%), demonstrando a relevância da IA para automatização de processos judiciais.

3.5.4.1. Modelos Preditivos para a Duração de Processos

A previsibilidade processual é um dos principais desafios do Judiciário. Oliveira et al. (2022) propuseram um modelo baseado em aprendizado de máquina para prever a duração de julgamentos no Tribunal Regional do Trabalho da 5ª Região (TRT-5). Utilizando uma base de dados de mais de 330.000 processos, os autores avaliaram sete técnicas de aprendizado de máquina, sendo o modelo *Adaboost* o mais eficaz, com um coeficiente de determinação (R^2) de 0,819.

Essa pesquisa destacou que magistrados e órgãos julgadores são variáveis determinantes na previsão da duração dos processos, indicando que o comportamento dos juízes influencia diretamente na celeridade judicial.

O impacto do uso de redes neurais também foi evidenciado por Silva et al. (2023), que analisaram modelos *BiLSTM* e mecanismos de autoatenção para prever o tempo restante até a conclusão de processos. O estudo demonstrou que essas abordagens superaram métodos tradicionais, como regressão linear e *Random Forest*, sugerindo que arquiteturas de aprendizado profundo são mais eficazes na identificação de padrões

processuais.

A pesquisa do CNJ (2023) revelou que 87 tribunais utilizam aprendizado supervisionado, enquanto outros 27,1% utilizam aprendizado não supervisionado. Entre os modelos mais utilizados, as redes neurais são predominantes (56,4%), seguidas por árvores de decisão e florestas aleatórias (21,1%). Esses achados reforçam que técnicas avançadas de aprendizado profundo estão sendo cada vez mais aplicadas na tramitação processual.

3.5.4.2. Aplicações da IA na Triagem e Classificação de Processos

Além da previsão de prazos, a IA tem sido empregada na classificação e triagem automática de processos judiciais. O Projeto Victor, implementado pelo Supremo Tribunal Federal (STF) em parceria com a Universidade de Brasília (UnB), utiliza redes neurais para identificar processos com repercussão geral, reduzindo drasticamente o tempo de análise dos recursos extraordinários.

Antes do Projeto Victor, a triagem manual desses processos levava cerca de 40 a 60 minutos; com a IA, esse tempo foi reduzido para apenas 5 segundos, otimizando a gestão processual e permitindo que os magistrados foquem em atividades de maior complexidade (ANDRADE et al., 2021; DAL PONT et al., 2021; MAIA FILHO; JUNQUILHO, 2018).

De forma similar, Castro Júnior et al. (2020) apresentaram a ferramenta Berna, um modelo baseado em Processamento de Linguagem Natural (PLN) e aprendizado por similaridade para identificar petições iniciais com fatos e teses jurídicas semelhantes. A implementação desse sistema no Judiciário de Goiás resultou em uma precisão de 96%, facilitando a vinculação de processos repetitivos e a gestão de precedentes.

O RAFA 2030, estudado por Fux et al. (2022), foi desenvolvido para classificar automaticamente processos segundo os Objetivos de Desenvolvimento Sustentável (ODS) da ONU. Utilizando aprendizado de máquina e PLN, a ferramenta obteve alta acurácia na classificação de processos, permitindo uma melhor gestão processual e priorização de julgamentos.

Além disso, Mota et al. (2020) investigaram o uso de redes neurais convolucionais multimodais para classificação de páginas de processos judiciais, combinando

informações textuais e visuais para segmentação automática de documentos. Essa abordagem pode contribuir significativamente para a automação da gestão documental no Judiciário.

3.5.4.3. Impacto da IA na Uniformização e Aplicação de Precedentes

A aplicação da IA para garantir a uniformidade na observância de precedentes também tem sido explorada. Sabo e Rover (2020) investigaram como o aprendizado de máquina pode auxiliar na aplicação do sistema de precedentes do Código de Processo Civil de 2015 (CPC/2015).

O estudo revelou que muitos tribunais ainda enfrentam dificuldades na identificação e aplicação de precedentes vinculantes, levando à necessidade de reclamações ao Superior Tribunal de Justiça (STJ) para garantir a uniformização jurisprudencial.

Macohin (2019) analisou a jurisprudência do Conselho Nacional de Justiça (CNJ) utilizando redes complexas. O estudo demonstrou que a métrica de centralidade de autovetor pode ser aplicada para identificar precedentes mais influentes, permitindo que tribunais priorizem decisões relevantes e aprimorem a consistência dos julgamentos.

A pesquisa do CNJ (2023) mostrou que tribunais já utilizam IA para recomendação de precedentes, com 23,1% dos projetos focados em reconhecimento de entidades jurídicas e classificação de jurisprudência. Contudo, um dos desafios levantados pela pesquisa foi a falta de padronização entre tribunais, uma vez que cada instituição desenvolve soluções próprias, muitas vezes sem integração com os sistemas eletrônicos nacionais, como o Processo Judicial Eletrônico (PJe). Apenas 36,4% dos projetos declararam integração com o PJe, indicando um problema de interoperabilidade e compartilhamento de informações jurídicas entre diferentes órgãos.

3.6. Relevância do Produto

3.6.1 Complexidade e Aderência

O Produto Técnico Tecnológico (PTT) desenvolvido neste trabalho representa um alto nível de complexidade técnica e computacional. A construção de um modelo preditivo baseado em redes neurais artificiais para estimar a duração dos processos judiciais no Superior Tribunal de Justiça (STJ), baseou-se inicialmente no tratamento e

na preparação de um extenso conjunto de dados, composto por 631.437 (seiscentos e trinta e um mil e quatrocentos e trinta e sete mil) registros de processos judiciais baixados no STJ.

Desse total, 50.515 (cinquenta mil e quinhentos e quinze mil) registros e 3.293 (três mil e duzentos e noventa e três) características (features). Esse volume de dados exigiu conhecimento técnico para manipulação e pré-processamento, incluindo imputação de valores ausentes, codificação de variáveis categóricas e normalização desses dados.

A implementação do modelo de rede neural Multilayer Perceptron (*MLPRegressor*) demandou a otimização de hiperparâmetros, como o número de neurônios na camada oculta, a taxa de aprendizado e a regularização. O uso do método *RandomizedSearchCV* permitiu a exploração da variabilidade dos hiperparâmetros disponíveis. Além disso, cada neurônio processou aproximadamente 33,2 bilhões de entradas ao longo de 200 iterações, evidenciando a magnitude computacional envolvida.

Quanto à aderência do PTT às áreas de concentração e linhas de atuação deste programa de pós-graduação, esta pesquisa alinha-se diretamente com várias das linhas de pesquisa propostas. O PTT contribui para a Gestão e governança baseada em evidências, ao utilizar dados concretos e análises criteriosas para melhorar a tomada de decisões judiciais. A incorporação de técnicas avançadas de aprendizado de máquina reflete as Políticas de governança digital e algorítmica, demonstrando como algoritmos podem otimizar processos públicos.

Além disso, o modelo está intimamente relacionado ao Desenvolvimento e transferência de instrumentos e tecnologias de gestão e governança, fornecendo uma ferramenta inovadora que pode ser adotada por outras instituições. Ao promover maior eficiência e transparência no sistema judicial, o PTT também aborda a Accountability e governança corporativa no poder judiciário.

Dessa forma, o trabalho evidencia uma forte aderência às áreas de concentração e linhas de atuação do programa, integrando inovação tecnológica com práticas de governança e gestão pública.

3.6.2 Potencial inovador

O PTT possui um elevado potencial inovador, caracterizado pela introdução de um modelo preditivo baseado em redes neurais artificiais no contexto do sistema judiciário brasileiro. Tradicionalmente, a previsão da duração de processos judiciais tem sido um desafio devido à complexidade e à variedade de fatores que influenciam o tempo de tramitação. A aplicação de aprendizado de máquina, especialmente redes neurais, representa uma abordagem nova e tecnologicamente aprimorada em relação aos métodos anteriormente utilizados.

A inovação está presente tanto no desenvolvimento técnico quanto na aplicação prática. O modelo alcançou um coeficiente de determinação (R^2) de 0,9243, indicando que é capaz de explicar aproximadamente 92,43% da variabilidade na duração dos processos judiciais. Além disso, o conhecimento gerado é fruto da pesquisa original deste discente, atendendo ao critério de uso intensivo de conhecimento inovador estabelecido pela CAPES.

3.6.3 Aplicabilidade

A aplicabilidade do PTT é ampla e apresenta tanto potencial quanto realização prática. Em termos de aplicabilidade potencial, o modelo foi desenvolvido utilizando linguagens e bibliotecas amplamente adotadas, como *Python*, *Scikit-learn*, *NumPy* e *Pandas*, facilitando sua integração nos sistemas existentes do STJ. A modularidade do código e a documentação detalhada permitem que o modelo seja adaptado e expandido conforme necessário, inclusive para outros tribunais ou órgãos que desejem implementar soluções similares.

Quanto à aplicabilidade realizada, o modelo foi testado com sucesso em um ambiente controlado, demonstrando sua viabilidade técnica. A integração do modelo aos sistemas do STJ pode ser realizada sem grandes dificuldades, graças ao uso de tecnologias compatíveis e à disponibilização de scripts de implantação. A facilidade de uso é reforçada pela inclusão de pipelines automatizados para pré-processamento dos dados e pela capacidade de atualização contínua do modelo com novos dados, garantindo sua relevância ao longo do tempo.

3.6.4 Impacto Potencial.

O impacto potencial do PTT é significativo, tanto para o STJ quanto para outras

organizações do sistema judiciário. Em termos de impacto potencial, o modelo aborda uma questão de alta relevância: a necessidade de otimizar a gestão dos processos judiciais e reduzir a morosidade do sistema. Ao fornecer previsões precisas sobre a duração dos processos, o PTT permite que o STJ e outros tribunais melhorem o planejamento estratégico, a alocação de recursos e a transparência com o público.

Outras organizações podem se beneficiar do PTT ao adaptá-lo para suas próprias necessidades. Por exemplo, tribunais regionais e estaduais podem utilizar o modelo como base para desenvolver suas próprias soluções preditivas, promovendo a eficiência em diferentes níveis do sistema judiciário. A metodologia e as técnicas empregadas podem servir de referência para projetos futuros, estimulando a inovação e a modernização em todo o setor.

Em termos de impacto realizado, embora o PTT ainda não tenha sido implementado operacionalmente no STJ, os resultados obtidos durante o desenvolvimento indicam melhorias potenciais significativas. A capacidade de prever a duração dos processos com alto grau de precisão pode levar a melhorias sociais, como o aumento da confiança pública no sistema judiciário, e econômicas, como a redução de custos operacionais associados a atrasos e ineficiências. Além disso, o PTT contribui para a disseminação de práticas inovadoras no ambiente jurídico, incentivando uma cultura de melhoria contínua e uso inteligente da tecnologia em prol da justiça.

3.6.5 Documentos comprobatórios e evidências.

Para atender às exigências do Comitê da área interdisciplinar da CAPES e assegurar a adequada classificação do Produto Técnico Tecnológico (PTT), foram reunidos documentos comprobatórios que evidenciam o desenvolvimento, a aplicação e o potencial impacto do modelo preditivo para estimar a duração dos processos judiciais no Superior Tribunal de Justiça (STJ):

- Código-Fonte completo do modelo; e
- Descrição Detalhada do Produto Técnico Tecnológico (*vide Cap. 3*).

Salienta-se que foram utilizados dados públicos, respeitando as diretrizes éticas e legais aplicáveis. A origem dos dados e o processo de coleta estão descritos no trabalho. Da mesma forma, a metodologia seguida está documentada de forma detalhada,

permitindo a replicação e validação por outros pesquisadores ou instituições interessadas.

O **código-fonte**, a base de dados em **CSV** e modelo treinado em **.pkl**, utilizados como base para a rede neural, encontram-se disponíveis no repositório público no GITHUB: <https://github.com/DjacyJunior/Defesamestrado>.

Por fim, registra-se que não foram realizadas entrevistas, questionários estruturados, atas de reuniões ou outros documentos de interação direta com o STJ para este trabalho. O desenvolvimento do PTT foi baseado em pesquisas bibliográficas, análise de dados públicos e aplicação de técnicas de aprendizado de máquina.

REFERÊNCIAS

- ALMEIDA NETO, M. A. DE et al. **Um modelo de inferência para a classificação de resultados processuais da Justiça Estadual**. Revista de Engenharia e Pesquisa Aplicada, v. 3, n. 3, 30 ago. 2018.
- ANDRADE, M. D. et al. **Inteligência artificial para o rastreamento de ações com repercussão geral: o projeto victor e a realização do princípio da razoável duração do processo**. Revista Eletrônica de Direito Processual – REDP, v. v. 21, n. 1, p. 312–335, 2021.
- BEKKERS, V.; EDELENBOS, J.; STEIJN, B. **Innovation in the Public Sector: Linking Capacity and Leadership**. New York: Palgrave Macmillan: [s.n.].
- BRAGA, A. P.; LUDERMIR, T. B.; CARVALHO, A. **Redes Neurais Artificiais: Teoria e Aplicações**. LTC Editora, 2011.
- BRASIL. **Constituição (1988). Constituição da República Federativa do Brasil**. Brasília, DF: Senado Federal, 1988.
- CASTRO, M. P.; GUIMARÃES, T. A. **Dimensões da inovação em organizações da justiça: proposição de um modelo teórico-metodológico**. Cadernos EBAPE.BR, v. 17, n. 1, p. 173–184, mar. 2019.
- CAVALCANTE, P. L. C.; CAMÕES, M. R. DE S. **Inovação pública no Brasil: uma visão geral de seus tipos, resultados e indutores**. Em: <https://repositorio.ipea.gov.br/handle/11058/8794>. [s.l: s.n.]. p. 119–143.
- CONSELHO NACIONAL DE JUSTIÇA. **Justiça em números**. Disponível em: <https://www.cnj.jus.br/pesquisas-judiciarias/justica-em-numeros/>. Acesso em: 10 fev. 2024.
- CONSELHO NACIONAL DE JUSTIÇA. **Relatório do XIII Encontro Nacional do Poder Judiciário**. Brasília: CNJ, 2020.
- CONSELHO NACIONAL DE JUSTIÇA. **Liões CNJ: Laboratório de Inovação, Inteligência e ODS**. Disponível em: <https://www.cnj.jus.br/programas-e-acoes/agenda-2030/lioes-cnj-laboratorio-de-inovacao-inteligenica-e-ods/>. Acesso em: 5 jan. 2024.
- CONSELHO NACIONAL DE JUSTIÇA; PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO. **Pesquisa uso de inteligência artificial (IA) no Poder Judiciário: 2023**. Brasília: [s.n.].
- CRESWELL, John W.; CRESWELL, J. David. **Projeto de pesquisa: métodos qualitativo, quantitativo e misto**. Tradução por Sandra Maria Mallmann da Rosa. Revisão técnica de Dirceu da Silva. 5. ed. Porto Alegre: Penso, 2021.
- DAL PONT, T. R. et al. **Classification and Association Rules in Brazilian Supreme Court Judgments on Pre-trial Detention**. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in

Bioinformatics). **Anais...** Springer Science and Business Media Deutschland GmbH, 2021.

FREITAS, C. S. DE; MEDEIROS, J. J. **Organizational impacts of the electronic processing system of the brazilian superior tribunal of justice.** Journal of Information Systems and Technology Management, v. 12, n. 2, p. 317–332, 31 ago. 2015.

FUX, L. et al. **Classificação de processos judiciais segundo Objetivos de Desenvolvimento Sustentável da Agenda ONU 2030.** Revista da CGU, v. 14, n. 26, 20 dez. 2022.

GIACOMAZZO, G.; LEITE, D. **O mestrado profissional no âmbito das políticas públicas na educação: reflexões sobre a produção do conhecimento científico.** p. 475–493, 2014.

GIL-GARCIA, J. R.; DAWES, S. S.; PARDO, T. A. **Digital government and public management research: finding the crossroads.** Public Management Review, v. 20, n. 5, p. 633–646, 4 maio 2018.

GOOGLE COLAB. **Google Colaboratory.** Disponível em: <https://colab.research.google.com/>. Acesso em: 02 fev. 2024.

HAYKIN, S. **Redes neurais: princípios e prática.** 2. ed. Porto Alegre: Bookman Editora, 2017.

KAPLAN, A.; HAENLEIN, M. Siri, Siri, in my hand: **Who’s the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence.** Business Horizons, v. 62, n. 1, p. 15-25, 2019.

KARAR, M. E.; AL-RASHEED, M. F.; AL-RASHEED, A. F.; REYAD, O. **IoT and Neural Network-Based Water Pumping Control System For Smart Irrigation.** Information Sciences Letters, v. 9, n. 2, 2020.

LACERDA, M. G. P. DE; D’AMORIM, C. B. C.; ALVIM, A. F. M. **Prevenção Decisões Judiciais em Tribunais Brasileiros de Primeira Instância com Aprendizagem de Máquina.** Revista de Engenharia e Pesquisa Aplicada, v. 6, n. 3, p. 81–89, 5 abr. 2021.

LERMAN, P. M. **Fitting Segmented Regression Models by Grid Search.** Journal of the Royal Statistical Society, 1980.

MACOHIN, A. **Análise da Jurisprudência do Conselho Nacional de Justiça através de Redes Complexas.** Revista CNJ, v. v. 3, n. 2, p. 19–26, 2019.

MAIA FILHO, M. S.; JUNQUILHO, T. A. **Projeto Victor: perspectivas de aplicação da inteligência artificial ao direito.** Revista de Direitos e Garantias Fundamentais, v. 19, n. 3, p. 218–237, 29 dez. 2018.

MAYER-SCHÖNBERGER, V.; CUKIER, K. **Big Data: A Revolution That Will**

- Transform How We Live, Work, and Think.** Elsevier, 2013.
- MCCARTHY, J. **What is Artificial Intelligence.** Stanford: Stanford University, 2007.
- McCULLOCH, W. S.; PITTS, W. **A logical calculus of the ideas immanent in nervous activity.** The bulletin of mathematical biophysics, v. 5, n. 4, p. 115–133, 1943.
Reimpresso em Bulletin of Mathematical Biology, v. 52, p. 99–115, 1990.
- MCTI. **Estratégia Brasileira de Inteligência Artificial (EBIA).** [s.l: s.n.].
- MELATI C, JANISSEK-MUNIZ R. **A inteligência na gestão pública: uma análise sob a perspectiva institucional.** Rev Adm Pública, v. 56, n. 6, 721–44, 2022.
- MORGAN, Gareth. **Imagens da organização.** 2. ed. São Paulo: Atlas, 2006.
- MOTA, C. C. R. et al. **Classification of Court Lawsuits Pages using Multimodal Convolution Neural Networks.** [s.l: s.n.].
- NAÇÕES UNIDAS BRASIL. **Objetivos de Desenvolvimento Sustentável.** Disponível em: <https://brasil.un.org/pt-br/sdgs>. Acesso em: 22 fev. 2024.
- NUMPY. **NumPy.** Disponível em: <https://numpy.org/>. Acesso em: 02 fev. 2024.
- OLOWONONI, F. O.; RAWAT, D. B.; LIU, C. **Resilient Machine Learning for Networked Cyber Physical Systems: A Survey for Machine Learning Security to Securing Machine Learning for CPS.** IEEE Communications Surveys & Tutorials, vol. 23, n. 1, 2021.
- OLIVEIRA, F. L. de; CUNHA, L. G. **Os indicadores sobre o Judiciário brasileiro: limitações, desafios e o uso da tecnologia.** Revista Direito GV, v. 16, n. 1, e1948, 2020.
- OLIVEIRA, R. S.; REIS, A. S.; NASCIMENTO, E. G. S. **Predicting the number of days in court cases using artificial intelligence.** PLoS ONE, v. 17, n. 5 5, 1 maio 2022.
- ORGANIZAÇÃO DAS NAÇÕES UNIDAS. **Transformando nosso mundo: a Agenda 2030 para o desenvolvimento sustentável.** Nova York: ONU, 2015.
- PANDAS. **Getting started with pandas.** Disponível em: https://pandas.pydata.org/docs/getting_started/index.html#getting-started. Acesso em: 15 jan. 2024.
- PEDREGOSA, F. et al. **Scikit-learn: Machine Learning in Python.** Journal of Machine Learning Research, v. 12, p. 2825-2830, 2011.
- PIRES, A. et al. **Aplicação da Inteligência Artificial na identificação de conexões pelo fato e tese jurídica nas petições iniciais e integração com o Sistema de Processo Eletrônico.** Revista CNJ, v. v. 4, n. 1, 2020.

RAMCHOUN, H.; IDRISSE, M. A. J.; GHANOU, Y.; ETTAOUIL, M. **Multilayer Perceptron: Architecture Optimization and Training**. International Journal of Interactive Multimedia and Artificial Intelligence, v. 4, n. 1, 2016.

ROSENBLATT, F. **The perceptron: a probabilistic model for information storage and organization in the brain**. Psychological Review, v. 65, n. 6, p. 386-408, 1958.

SABO, I. C.; ROVER, A. J. **Compliance to legal precedents and repetitive legal cases management with machine learning**. Revista Opinião Jurídica, v. 18, n. 28, p. 69-93, 1 maio 2020.

SALOMÃO, L. F. et al. **Tecnologia aplicada à gestão dos conflitos no âmbito do Poder Judiciário brasileiro – 1ª fase**. Rio de Janeiro: Centro de Inovação, Administração e Pesquisa do Judiciário da Fundação Getúlio Vargas, 2019.

SALOMÃO, L. F. et al. **Tecnologia aplicada à gestão dos conflitos no âmbito do Poder Judiciário brasileiro – 2ª fase**. Rio de Janeiro: Centro de Inovação, Administração e Pesquisa do Judiciário da Fundação Getúlio Vargas, 2021.

SALOMÃO, L. F. et al. **Tecnologia aplicada à gestão dos conflitos no âmbito do Poder Judiciário brasileiro – 3ª fase**. Rio de Janeiro: Centro de Inovação, Administração e Pesquisa do Judiciário da Fundação Getúlio Vargas, 2023.

SCOPUS. **Scopus** - Pesquisa. Disponível em: <https://www.scopus.com/search/form.uri?display=basic#basic>. Acesso em: 02 mar. 2024.

SCIKIT-LEARN. **Neural network models (supervised)**. Disponível em: https://scikit-learn.org/stable/modules/neural_networks_supervised.html. Acesso em: 20 jan. 2024.

SCIKIT-LEARN. **sklearn.neural_network.MLPRegressor**. Disponível em: https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPRegressor.html#sklearn.neural_network.MLPRegressor. Acesso em: 23 jan. 2024.

SCHWAB, K. **A quarta revolução industrial**. 1ª Edição 2016 ed. São Paulo: Edipro, 2016.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes neurais artificiais para engenharia e ciências aplicadas**. São Paulo: Artliber, p. 33-111, 2010.

SILVA, R. O. et al. **Predição de Tempo Restante para Conclusão de Processos de Negócio Utilizando Aprendizado Profundo**. Proceedings of the 38th Brazilian Symposium on Databases, 2023.

TURING, A. M. **Computing machinery and intelligence**. Mind, v. 59, p. 433-460, 1950.

WANG, P.; HAFSHEJANI, B. A.; WANG, D. **An improved multilayer perceptron**

approach for detecting sugarcane yield production in IoT based smart agriculture.
Microprocessors and Microsystems, v. 82, pp. 1-7, 2021.